

---

# **Toward Accurate Galaxy Cluster Masses with Gravitational Weak Lensing**

**Maria Paulus**

---



München 2021



---

# **Toward Accurate Galaxy Cluster Masses with Gravitational Weak Lensing**

**Maria Paulus**

---

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

vorgelegt von  
Maria Paulus  
aus Wien, Österreich

München, den 16.3.2021

Erstgutachter: Prof. Joseph Mohr

Zweitgutachter: Prof. Jochen Weller

Tag der mündlichen Prüfung: 11.5.2021

*“I haven’t been everywhere, but it’s on my list”*  
– Susan Sontag



# Contents

<b>Zusammenfassung</b>	<b>xiii</b>
<b>Abstract</b>	<b>xiv</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Galaxy Clusters as Cosmological Probes . . . . .	1
1.1.1 Dark Matter Candidates . . . . .	2
1.2 Observational Signatures of Galaxy Clusters . . . . .	2
1.2.1 Optical . . . . .	3
1.2.2 X-ray . . . . .	4
1.2.3 Millimetre regime . . . . .	4
1.2.4 Gravitational Lensing . . . . .	5
1.3 Surveys . . . . .	8
1.3.1 Dark Energy Survey . . . . .	8
1.3.2 Sloan Digital Sky Survey . . . . .	13
1.3.3 Kilo-Degree Survey . . . . .	15
1.3.4 Rubin . . . . .	15
1.3.5 Euclid . . . . .	18
1.4 Outline and Motivation . . . . .	20
<b>2 Cluster Member Contamination of SPT and MARD-Y3 clusters</b>	<b>21</b>
2.1 Abstract . . . . .	21
2.2 Introduction . . . . .	22
2.3 Data . . . . .	24
2.3.1 DES Y1 shear and photo-z catalogues . . . . .	24
2.3.2 MARD-Y3 cluster catalogue . . . . .	24
2.3.3 SPT cluster catalogue . . . . .	26
2.3.4 SPT + MARD-Y3 cluster sample . . . . .	28
2.4 Cluster member contamination . . . . .	28
2.4.1 Weak lensing . . . . .	31
2.4.2 Source background distribution . . . . .	32
2.4.3 Cluster contamination and correction . . . . .	33
2.4.4 P(z) decomposition method . . . . .	35

2.4.5	Likelihood . . . . .	38
2.4.6	Validation with mock catalogues . . . . .	39
2.5	Cluster contamination measurements . . . . .	39
2.5.1	Validation of the cluster contamination results . . . . .	40
2.5.2	Evaluating the adequacy of the model . . . . .	48
2.5.3	Redshift variation of cluster contamination . . . . .	50
2.5.4	Quantifying the impact on cluster mass . . . . .	52
2.6	Summary and Conclusions . . . . .	54
2.7	Acknowledgments . . . . .	55
<b>3</b>	<b>Cluster Member Contamination of redMaPPer clusters</b>	<b>59</b>
3.1	Introduction . . . . .	59
3.2	Data . . . . .	61
3.2.1	DES Y1 . . . . .	61
3.2.2	redMaPPer Cluster Catalogue . . . . .	62
3.3	Cluster member contamination . . . . .	63
3.3.1	Source background distribution . . . . .	63
3.3.2	P(z) decomposition method . . . . .	64
3.4	Cluster contamination measurements . . . . .	67
3.4.1	Model-free estimation of the cluster contamination . . . . .	71
3.4.2	Redshift variation of cluster contamination . . . . .	72
3.5	Comparison to the literature . . . . .	72
3.6	Impact on Cluster Mass . . . . .	78
3.6.1	Lensing signal . . . . .	78
3.6.2	Miscentering correction . . . . .	79
3.6.3	Mean mass in redshift-richness bins . . . . .	80
3.6.4	Impact on cluster mass estimation . . . . .	82
3.7	Summary and Conclusions . . . . .	83
<b>4</b>	<b>Conclusions and Outlook</b>	<b>85</b>
<b>A</b>	<b>Appendix to Chapter 2</b>	<b>89</b>
A.1	Mock catalogue . . . . .	89
A.2	Model-free estimation of the cluster contamination . . . . .	89
<b>Acknowledgements</b>		<b>102</b>

# List of Figures

1.1	Coma cluster in optical, X-ray and millimetre wavelengths . . . . .	3
1.2	Representation of a typical gravitational lensing diagram . . . . .	6
1.3	Comparing survey footprints . . . . .	9
1.4	DES footprint . . . . .	12
1.5	SDSS DR7 footprint . . . . .	14
1.6	KiDS data release 4 . . . . .	16
1.7	Rubin footprint . . . . .	17
1.8	Euclid's wide and deep surveys footprints . . . . .	19
2.1	Mass versus redshift for MARD-Y3 . . . . .	27
2.2	Mass versus redshift for SPT . . . . .	29
2.3	Comparing SPT and MARD-Y3 richnesses . . . . .	30
2.4	Example source background distribution . . . . .	34
2.5	$P(z)$ of a cluster shown in logarithmic annuli from the cluster centre . . . . .	36
2.6	Binned results: Best fit parameters for SPT, MARD-Y3 and SPT+MARD-Y3	41
2.7	Gaussianized interpolation results: Best fit parameters for SPT, MARD-Y3 and SPT+MARD-Y3 . . . . .	42
2.8	Different normalisation parameters of the cluster contamination for SPT, MARD-Y3 and SPT+MARD-Y3 . . . . .	45
2.9	Field-subtracted redshift distribution along with the best fit model both stacked in the same cluster redshift range. . . . .	47
2.10	Cluster contamination fraction plotted as a function of radius, comparing data and model prediction, stacked in redshift bins . . . . .	48
2.11	Cluster contamination fraction plotted as a function of radius, comparing data and model prediction, stacked in richness bins . . . . .	49
2.12	Best fit parameters for individual redshift bins . . . . .	51
3.1	Contour plot of the best fit model parameters for redMaPPer . . . . .	66
3.2	Different normalisation parameters of the cluster contamination versus cluster redshift for the redMaPPer sample. . . . .	68
3.3	Comparing different normalisation parameters of the cluster contamination for the redMaPPer, SPT and MARD-Y3 sample. . . . .	69
3.4	Best fit parameters for different richness cuts . . . . .	70

---

3.5	Comparing field-subtracted redshift distribution with the best fit model stacked in $0.2 < z_{\text{cl}} < 0.3$	73
3.6	Cluster contamination fraction as a function of radius	74
3.7	Field-subtracted redshift distributions along with the best fit models stacked in various cluster redshift - richness bins	76
3.8	Cluster contamination fraction as a function of radius for 12 redshift - richness bins	77
3.9	Impact on cluster mass for different contamination fractions	81
A.1	Mock simulation for the combined SPT + MARD-Y3 cluster sample	90
A.2	Field-subtracted redshift distribution stacked in various cluster redshift bins along with the average best-fit model shown for the first radial bin (0.25-0.38 Mpc)	91
A.3	Same as Figure A.2 but for the second radial bin (0.38-0.57 Mpc).	91
A.4	Same as Figure A.2 but for the third radial bin (0.57-0.87 Mpc).	92
A.5	Same as Figure A.2 but for the fourth radial bin (0.87-1.31 Mpc).	92
A.6	Field-subtracted redshift distribution stacked in various cluster richness bins along with the average best-fit model shown for the first radial bin (0.25-0.38 Mpc)	93
A.7	Same as Figure A.6 but for the second radial bin (0.38-0.57 Mpc).	94
A.8	Same as Figure A.6 but for the third radial bin (0.57-0.87 Mpc).	95
A.9	Same as Figure A.6 but for the fourth radial bin (0.87-1.31 Mpc).	96

# List of Tables

2.1	Binned results: Best fit parameters for SPT, MARD-Y3 and SPT+MARD-Y3 in numbers . . . . .	43
2.2	Gaussianized interpolation results: Best fit parameters for SPT, MARD-Y3 and SPT+MARD-Y3 in numbers . . . . .	44
2.3	Systematic mass uncertainty implied by the parameter uncertainties on the cluster member contamination constraints . . . . .	52



# Zusammenfassung

Galaxienhaufen gelten als leistungsstarke kosmologische Sonden, da ihre Häufigkeit als Funktion der Rotverschiebung und Masse direkt von der zugrunde liegenden Kosmologie abhängt. Ihre Beobachtungsmerkmale ermöglichen eine Auswahl von Galaxienhaufen in verschiedenen Wellenlängen (optisch, Röntgen und Millimeter) und schließlich eine Schätzung ihrer Massen durch Kalibrierung ihrer Beziehung zwischen Masse und einer beobachtbaren Variable. Eine der vielversprechendsten Wege zur Bestimmung von Galaxienhaufenmassen ist der schwache Gravitationslinseneffekt. Die induzierte tangentiale Ausrichtung von Hintergrundgalaxien um einen Vordergrundgalaxienhaufen ist direkt empfindlich gegenüber der Galaxienhaufenmasse. Massenschätzungen von Galaxienhaufen durch den schwachen Gravitationslinseneffekt unterliegen jedoch systematischen Unsicherheiten, die sich auf die Kosmologie auswirken. In dieser Arbeit konzentrieren wir uns auf die sogenannte Galaxienhaufen-Kontaminationskorrektur, eine wichtige systematische Unsicherheit, die sich auf die Analyse von schwachen Linsen auswirkt. Diese Kontamination kommt zustande da Haufeneigene Galaxien aufgrund von photometrischen Unsicherheiten bei der Rotverschiebungsschätzungen in die Hintergrundverteilung von Galaxien streuen. Infolgedessen wird die Schermessung der schwachen Gravitationslinse unterschätzt, da wir über Galaxien ohne Schersignal mitteln. Dies macht es notwendig, das gemessene Signal zu verstärken, um das wahre Signal wiederherzustellen. In Kapitel 1 stellen wir das Grundgerüst vor, um die Physik einer Massenschätzung durch den schwachen Gravitationslinseneffekt zu verstehen. Des Weiteren stellen wir verschiedene bildgebende Vermessungen von Regionen des Himmels vor und vergleichen sie. In Kapitel 2 stellen wir eine neue und verbesserte Methode für die Kontaminationsmessung von Haufenmitgliedern vor, die wir auf 250 SPT Sunyaev-Zel'dovich selektierte und auf 208 MARD-Y3-röntgenselektierte Galaxienhaufen anwenden. In Kapitel 3 wiederholen wir die Galaxienhaufen Kontaminationsanalyse auf 5750 redMaPPer optisch ausgewählten Galaxienhaufen. Darüber hinaus messen wir die Auswirkung auf die mittleren Glaxienhaufenmassen für einzelne Bins in Rotverschiebung,  $z$ , und Anzahl der Tochtergalaxien,  $\lambda$ , wenn wir die Ergebnisse dieser Analyse im Vergleich zu denen einer vergleichbaren Arbeit betrachten. In Kapitel 4 werden unsere Ergebnisse zusammengefasst und wir schließen mit einer Diskussion über neue Probleme, die sich für die zukünftige Arbeit ergeben haben.



# Abstract

Galaxy clusters are regarded as powerful cosmological probes, as their abundance as a function of redshift and cluster mass is directly sensitive to the underlying cosmology. Their observational features allow a selection of galaxy clusters in various wavelengths (optical, X-ray and millimetre) and finally an estimation of their masses by calibrating their mass-observable relation. A sought after way to determine cluster masses is via gravitational weak lensing. The induced tangential alignment of background galaxies around a foreground cluster is directly sensitive to the cluster mass. But weak lensing cluster mass estimates are subject to systematic uncertainties, which impact the resulting cosmological constraining power from clusters. In this thesis we focus on the so called cluster contamination correction, an important systematic uncertainty impacting the weak lensing analysis. This contamination results from cluster member galaxies scattering into the background sample of weak lensing source galaxies due to photometric redshift uncertainties. As a result, the weak lensing shear measurement is diluted, as we are averaging over galaxies with no shear signal at all. This makes it necessary to boost the measured signal in order to recover the true signal. In Chapter 1 we introduce the basic framework to understand the physics of a weak lensing mass estimation and introduce and compare various imaging surveys. In Chapter 2 we establish a new and improved method for the cluster member contamination estimation, which we apply to 250 South Pole Telescope Sunyaev-Zel'dovich selected and to 208 MARD-Y3 X-ray selected cluster samples. In Chapter 3 we repeat the cluster contamination analysis on the 5750 redMaPPer optically selected clusters. Furthermore, we estimate the impact on the mean cluster masses for individual redshift-richness bins when considering the results from this analysis versus those from a comparable work. In Chapter 4 our findings are summarised and we conclude with a discussion on new problems that have emerged for future work.



# Chapter 1

## Introduction

### 1.1 Galaxy Clusters as Cosmological Probes

Galaxy clusters can be found at the high-mass end of collapsed objects. Typical galaxy clusters contain between hundreds to thousands of galaxies with masses ranging between  $10^{14}$  and  $10^{15}$  solar masses. The clusters mass is dominated by dark matter, which makes up roughly 80% of the total cluster mass. This fact was first discovered in the 1930s by Fritz Zwicky who estimated the mass of the Coma cluster using the virial theorem, by measuring the velocity dispersions of the galaxies within the cluster (Zwicky, 1933). When comparing this mass estimate to the mass measured from the total stellar luminosity he found that the member galaxies were moving too fast within the cluster for the amount of visible mass. To account for the random speed of the galaxies, a much larger luminous mass would be required. This discrepancy could not be explained by adding the intra-cluster gas, which was found to account for roughly 10% of the total cluster mass (e.g., Allen et al., 2002). The cluster member galaxies themselves only account for about 2% of the total mass. Thus, a large amount of unseen matter is required to bind the fast moving galaxies in the cluster and to account for the missing mass that is not accounted for by the visible baryonic mass. Zwicky called this matter “dark matter”.

Due to their unique properties galaxy clusters are considered as powerful cosmological probes. Their abundance as a function of redshift and cluster mass is directly sensitive to the underlying cosmology and can be used to probe the amount of structure in the Universe and its growth over cosmic time. By considering individual cluster systems, significant progress was made in studying the baryonic and dark matter content of clusters as well as providing first evidence for a low matter density Universe (White et al., 1993). By combining cluster mass estimates one can constrain the average matter density,  $\Omega_m$ , and the amplitude of density perturbations,  $\sigma_8$ , of the Universe (recently by Dietrich et al., 2018; Costanzi et al., 2019; Bocquet et al., 2020).

In order to extract accurate cosmological constraints from a sample of galaxy clusters one first needs to create a cluster sample that has been observed in either optical wavelengths, in X-ray and through their Sunyaev-Zel'dovich effect, and second needs to estimate their

masses by calibrating their mass-observable relation (MOR), that links cluster mass to an observable quantity. Generally, there is no obvious mass proxy in the galaxy distribution and a scaling between mass and an observable can only be calibrated empirically. A lot of different mass proxies have been used in the literature, often dependent on how the cluster sample was selected. For an X-ray selected cluster sample, the X-ray luminosity is generally used as the mass proxy (e.g. Klein et al., 2019), whereas for an optically selected cluster sample the mass proxy is based on the number of (red) galaxies inside a given radius, its richness (e.g. Rykoff et al., 2016a). A reliable way to determine cluster masses is via gravitational weak lensing, as the induced tangential alignment of background galaxies around a foreground cluster is directly sensitive to the cluster mass. Furthermore, the resulting weak lensing signal is sensitive to the total cluster mass, not just the baryonic component, and is independent of the physical state of the cluster.

### 1.1.1 Dark Matter Candidates

Over the years many dark matter candidates have been hypothesised and proposed but up to this date it is still not clear what constitutes dark matter. Dark matter can refer to any substance which interacts predominantly via gravity with visible matter. Furthermore, as it is not observed to shine, the dark matter particles must have very weak electromagnetic interactions. This would suggest that it is non-baryonic or even, at least partially, baryonic in nature. The main baryonic candidates are so called massive compact halo objects (MACHOs), which include non-primordial black holes, neutron stars, faint old white dwarfs and brown dwarfs (see e.g. Alcock et al., 1997; Fields et al., 1998). Most of these possibilities were eventually ruled out by measurements of the abundance of light elements produced in the primordial nucleosynthesis (Dar, 1995; Sarkar, 1996; Olive, 1997), favouring a non-baryonic, electromagnetically not interacting dark matter. This result has been confirmed with higher precision by measurements of the CMB (e.g. Hinshaw et al., 2013) as well as searches for gravitational microlensing effects (e.g. Freese et al., 2003). This leaves the non-baryonic dark matter candidates. These are for example hypothetical particles such as axions, sterile neutrinos, weakly interacting massive particles (WIMPs), gravitationally-interacting massive particles (GIMPs) or supersymmetric particles. Many experiments are attempting to directly and indirectly detecting such particles (see e.g. Klasen et al., 2015; Roszkowski et al., 2018), where WIMPS are the most popular candidates, closely followed by axions. The axion is a low mass, weakly interacting particles with zero spin, emerging from a theory that was proposed to solve the strong charge-parity (CP) problem (e.g. Du et al., 2018).

## 1.2 Observational Signatures of Galaxy Clusters

In this section we aim to discuss the main observational signatures of galaxy clusters at different wavelengths, namely in optical, X-ray and millimetre. Figure 1.1 shows images of the Coma cluster in optical wavelengths from the Sloan Digital Sky Survey (left panel),

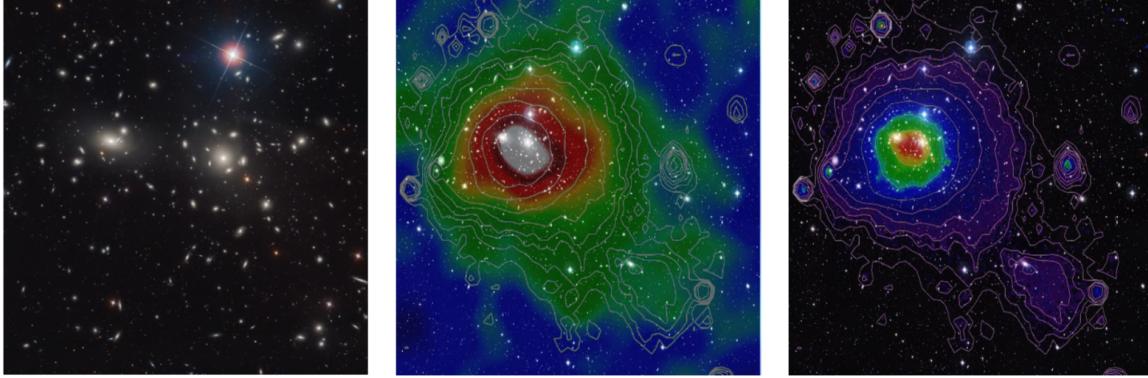


Figure 1.1: The Coma galaxy cluster in optical data from the Sloan Digital Sky Survey (left panel), as seen in by Planck through the SZ effect (central panel, shaded colour), and by ROSAT in X-rays (right panel, shaded colour and contours in both panels). The images in the central and right panels are overlaid on optical images obtained by DSS. Credits: ESA / LFI and HFI Consortia (Planck image); MPI (ROSAT image); NASA/ESA/DSS2/SDSS (optical image). Acknowledgement: Davide De Martin (ESA/Hubble).

in millimetre wavelength (SZE effect) as seen by the Planck (centre panel) and in X-ray produced by the ROSAT satellite (right panel). The colours in the centre and right images correspond to the intensity of the measured signal. Also these two images are superimposed onto the optical image as a visual aid. These observational features are dominated by the depth of the gravitational potential of the cluster (Kravtsov & Borgani, 2012, , for a review) and therefore the strength of their signatures correlates with the clusters mass (see e.g. Pratt et al., 2019). As mentioned above, calibrating the relationship between an observable and the cluster mass, is deemed as a powerful cosmological probe.

### 1.2.1 Optical

Optical cluster finder algorithms often take advantage of the fact that a large quantity of the cluster population is made up of old, red galaxies. Hence, in the optical regime these galaxy clusters appear as overdensities of red-sequence galaxies. Focusing on this specific galaxy population, you have galaxies occupying the so called red sequence, which is defined by its narrow colour range as a function of redshift. Photometric data can be filtered by calibrating this relationship as a colour filter on spectroscopic data. By adding a spatial filter, one can select over-densities of galaxies with colours consistent with a given redshift. This permits the accurate estimation of photometric redshifts. The amplitude in the colour-spatial filter is called richness,  $\lambda$ , (Rozo et al., 2009), which roughly scales like the number of galaxies in the cluster above a given magnitude. A further observational feature is the Brightest Central Galaxy (BCG), on which galaxy clusters are preferentially centred. Several cluster finder algorithms have been introduced over the years with the most recent ones being for example the MaxBCG algorithm (Koester et al., 2007), CAMIRA

(Oguri, 2014) and redMaPPer (Rykoff et al., 2016a) as well as AMICO (Bellagamba et al., 2017), which have been applied to surveys like the Sloan Digital Sky Survey Data Release 8 (SDSS DR8; York et al., 2000; Aihara et al., 2011), the Dark Energy Survey (DES; DES Collaboration, 2016) and the Kilo Degree Survey (KiDS, de Jong et al., 2017).

### 1.2.2 X-ray

Galaxy clusters form from the gravitational collapse of over-dense regions in the matter density distribution in the Universe, as deep potential wells cause the in-falling baryonic matter to heat up due to the gravitational contraction. Finally, they reach a state of virial equilibrium and the resulting hot plasma, called intracluster medium (ICM), emits the bulk of the thermal energy in the regime of soft X-rays as bremsstrahlung (Sarazin, 1986; Böhringer & Werner, 2010). If the gravitational potentials of clusters of different mass have a self-similar shape, as implied by numerical simulations of gravitational collapse (e.g. Moore et al., 1996; Navarro et al., 1996), then one finds the following self-similar relation between cluster mass and ICM temperature:  $T \propto M^{2/3}$ , where the mass,  $M$ , refers to the total mass of the galaxy clusters including the dark matter. Most X-ray cluster catalogues are produced using the data from the ROSAT satellite which includes 6 months of observations of the ROSAT All Sky Survey (RASS) followed by 8 years of pointed observations. These pointed observations may cover a much smaller area than RASS but are twice as deep. Examples of RASS based X-ray selected cluster catalogues include REFLEX (Böhringer et al., 2004), NEP (Henry et al., 2006), MARD-Y3 (Klein et al., 2019) and eRosita (Predehl et al., 2021).

### 1.2.3 Millimetre regime

Galaxy clusters show a clear signature in the millimetre regime, known as the (thermal) Sunyaev-Zel'dovich effect (SZE, Sunyaev & Zel'dovich, 1972a). This effect is created by Cosmic Microwave background (CMB) photons passing through a cluster. A small fraction of roughly 1% of these photons will inverse-Compton scatter off the energetic electrons in the Intracluster Medium (ICM) to higher frequencies, producing a characteristic distortion of the CMB spectrum at the location of the cluster. This leads to a frequency dependent change in the CMB temperature, where the observed temperature difference  $\Delta T$  to the mean CMB temperature,  $T_{\text{CMB}}$ , is given by

$$\begin{aligned} \Delta T &= T_{\text{CMB}} f_{\text{SZ}}(x) \int n_e \frac{k_B T_e}{m_e c^2} \sigma_T dl \\ &\equiv T_{\text{CMB}} f_{\text{SZ}}(x) y_{\text{SZ}} . \end{aligned} \quad (1.1)$$

where the integral is along the line of sight,  $x \equiv h\nu/k_B T_{\text{CMB}}$ ,  $k_B$  is the Boltzmann constant,  $c$  is the speed of light,  $n_e$  the electron density,  $m_e$  the electron mass,  $T_e$  the electron temperature,  $\sigma_T$  the Thomson cross-section,  $f_{\text{SZ}}(x)$  encodes the frequency dependence of the thermal SZE effect and  $y_{\text{SZ}}$  is the Compton y-parameter which denotes the amplitude of the SZE effect.

This frequency dependence leads to a decrease in the intensity of the CMB black-body spectrum at frequencies  $\lesssim 217$  GHz (or decrement of observed photons) and to an increase at higher frequencies  $\gtrsim 217$  GHz (or increment of photons). At  $\sim 217$  GHz there is no shift in the CMB intensity. The surface brightness of the thermal SZE effect is independent of redshift, and the integrated thermal SZE signal is expected to be a low-scatter proxy for the cluster mass, as it is proportional to the total thermal energy of the ICM (Motl et al., 2005). These properties make cluster samples produced by SZE surveys attractive for cosmological analyses (Carlstrom et al., 2011a).

Due to the small amplitude of the observed temperature distortions in the CMB caused by the thermal SZE effect, only over the last decade has it become possible to observe clusters through the SZE effect in wide and deep surveys. After the first discovery of a previously unknown cluster through its SZE signature in 2009 (Staniszewski et al., 2009), it has become routine with cluster catalogues being produced by e.g. the South Pole Telescope (SPT, Bleem et al., 2015a, 2020), the Atacama Cosmology Telescope (Hasselfield et al., 2013; Hilton et al., 2018) and the Planck satellite (e.g. Planck Collaboration et al., 2016a).

### 1.2.4 Gravitational Lensing

Einstein's General Theory of Relativity predicts that light rays coming from background sources are deflected by the gravitational field of massive foreground objects along the line of sight. This phenomenon is commonly referred to as gravitational lensing (see Bartelmann, 2010, for a recent review). This phenomenon happens on various scales, with photons being emitted by distant galaxies being bent by galaxies, galaxy clusters or even planets and stars. The images of these background sources that reach us will appear slightly magnified and distorted due to the lensing effect. Putting this effect in a mathematical context an equation is required that relates the true position of the background object to its observed position in the sky. Figure 1.2 shows the typical geometry of a gravitational lens system. A light ray emitted from a source  $S$  is deflected at the lens plane by an angle  $\vec{\alpha}$  before it reaches the observer  $O$ .  $\theta$  denotes the angle between the optical axis and the image  $I$  and the angle between the optical axis and the true source position is  $\beta$ .  $D_s$ ,  $D_d$  and  $D_{ds}$  are the angular diameter distances between observer and source, observer and lens and source and lens, respectively. The deflection angle  $\vec{\alpha}$  can be related to the projected gravitational potential of the lens in a way such that

$$\vec{\alpha} = 2 \int \nabla_{\perp} \Phi ds, \quad (1.2)$$

where  $ds$  is the photon path. This angle is connected to the reduced or observed deflection angle by

$$\vec{\alpha} = \frac{D_{ds}}{D_s} \vec{\alpha}. \quad (1.3)$$

From Fig. 1.2 one can see that  $\theta D_s = \beta D_s - \vec{\alpha} D_{ds}$ . Therefore, the positions of the observed image and the true source are related through the so called lens equation:

$$\vec{\beta} = \vec{\theta} - \vec{\alpha}(\vec{\theta}). \quad (1.4)$$

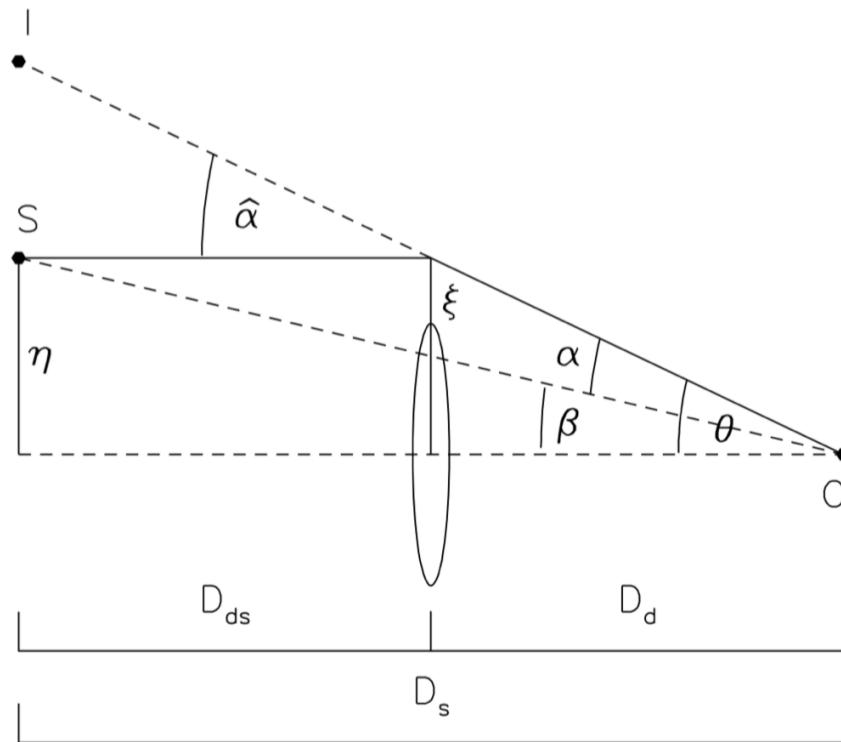


Figure 1.2: Illustration of a gravitational lens system. The light ray propagates from a source  $S$  to the observer  $O$  and is deflected by an angle  $\vec{\alpha}$  along the way.  $\theta$  denotes the angle between the optical axis and the image  $I$  and the angle between the optical axis and the true source position is  $\beta$ . The distances between the observer and the source, the observer and the lens, and the lens and the source are  $D_s$ ,  $D_d$  and  $D_{ds}$ , respectively. Credits: Narayan & Bartelmann (1997)

where  $\vec{\alpha}(\vec{\theta})$  represents the scaled deflection angle. The lens equation (Equ. 1.4) tells us that a source with true position  $\vec{\beta}$  will be seen by an observer at an angular position  $\vec{\theta}$ . If Equ. 1.4 has more than one solution for a fixed  $\vec{\beta}$ , a source at  $\vec{\beta}$  will have images at multiple positions in the sky. This means that the lens is producing multiple images which leads us to the ‘strong’ lensing regime. This can be quantified by the dimensionless surface mass density:

$$\kappa(\vec{\theta}) = \frac{\Sigma(D_d\vec{\theta})}{\Sigma_{\text{crit}}} \quad \text{with} \quad \Sigma_{\text{crit}} = \frac{c^2}{4\pi G} \frac{D_s}{D_d D_{ds}}, \quad (1.5)$$

where  $\kappa(\vec{\theta})$  is the convergence and  $\Sigma_{\text{crit}}$  is called the critical surface mass density (depends on the redshifts of source and lens). Hence, according to the positions of the source, lens and observer, and the mass and shape of the lens, we can distinguish between two different regimes: strong lensing and weak lensing. A mass distribution which has  $\kappa \geq 1$ , i.e.  $\Sigma \geq \Sigma_{\text{crit}}$  (deflection angle  $\vec{\alpha}$  is comparable to the angular size of the source) leads to a ‘strong’ lensing effect with strong distortions, where multiple images of the source are produced. These images can even blend together to form arcs. Otherwise, if  $\kappa < 1$ , i.e.  $\Sigma < \Sigma_{\text{crit}}$  ( $\vec{\alpha}$  is much smaller than the angular size of the source), only small deformations are produced and we are in the regime of weak lensing. In this case, the source will still appear distorted, magnified (convergence) and stretched (shear). Hence,  $\Sigma_{\text{crit}}$  is a characteristic value for the surface mass density which distinguishes between ‘weak’ and ‘strong’ lensing.

One of the most accurate ways to determine galaxy cluster masses is via gravitational weak lensing. By measuring the tangential shear around a cluster, one can accurately infer the cluster mass, as the tangential alignment of background galaxies around a foreground cluster is directly sensitive to the mass of the cluster. One cannot directly measure shear or convergence, but rather a combination of both, known as the reduced tangential shear,  $g_t(\vec{\theta})$ :

$$g_t(\vec{\theta}) = \frac{\gamma(\vec{\theta})}{1 - \kappa(\vec{\theta})}, \quad (1.6)$$

where

$$\gamma(\vec{\theta}) = \frac{\Sigma(< D_d\vec{\theta}) - \Sigma(D_d\vec{\theta})}{\Sigma_{\text{crit}}}. \quad (1.7)$$

Here,  $\gamma(\vec{\theta})$  refers to the tangential shear, whereas  $\Sigma(< D_d\vec{\theta})$  represents the average surface mass density within a projected distance and  $\Sigma(D_d\vec{\theta})$  is the average surface mass density at that projected distance.

The more massive the foreground cluster, the stronger its gravitational field and hence the greater the bending of light rays. Therefore, the amplitude of the distortion provides a direct measure of the projected gravitational potential. The cluster mass can only be measured statistically over an ensemble of galaxy shapes due to the weakness of the weak lensing effect. The projected mass distribution of a cluster estimated via gravitational weak lensing does neither depend on the physical state of the matter nor on the nature of it and no assumptions on the dynamical state of the matter need to be made. This is an

advantage over more traditional ways of determining cluster masses. Dynamical methods, where radial velocity measurements are interpreted in terms of the gravitating mass of the cluster, require the cluster to be in virial equilibrium in order to be reliable. This is not guaranteed. Furthermore, projection effects and the anisotropy of galaxy orbits in clusters also affect the mass determination. An X-ray analysis, that investigates the diffuse X-ray emission from the hot intra-cluster gas residing in the cluster potential well (see, e.g., Sarazin 1986), relies on the assumption that the intra-cluster gas is in hydrostatic equilibrium. Additionally, also here the consequences of projections effects are difficult to assess.

The most common approach to estimate weak lensing masses is based on fitting a mass profile to the observed reduced shear profile. This is possible, as numerical simulations with collisionless, cold dark-matter particles show that, on average, the density profiles of clusters are well described by the Navarro-Frenk-White profile (NFW, Navarro et al., 1997). This NFW profile is defined as:

$$\rho(r) = \frac{\rho_0}{r_s(1 + \frac{r}{r_s})^2}, \quad (1.8)$$

where  $\rho_0$  and  $r_s$  are scaling factors in density and radius. The profile can also be parameterised by the total  $M_\Delta$  and the concentration  $c_\Delta = r_\Delta/r_s$ , where  $\Delta$  refers to a specified spherical overdensity. By making appropriate assumptions about the concentration parameter, one can fit a mass that is commonly referred to as the weak lensing mass  $M_{WL}$ . This  $M_{WL}$  is related to the true mass of the cluster by a bias factor that needs to be accounted for.

Weak lensing mass estimates are subject to various systematic and statistical uncertainties. Due to sky surveys becoming larger and deeper as well as wielding a larger statistical power, the correct characterisation of systematic uncertainties is becoming more important. McClintock et al. (2019) has recently shown that stacked weak lensing mass estimates are for the first time dominated by systematic uncertainties. This means, in order to improve the cosmological constraining power from clusters, one has to first gain a better understanding of the systematic uncertainties affecting the cluster mass calibration.

## 1.3 Surveys

This section mainly focuses on the Dark Energy Survey (DES) and compares it to ongoing and future photometric surveys like SDSS, KiDS, Rubin and Euclid, used or designed for weak lensing analysis. Figure 1.3 shows a comparison between the Euclid wide survey footprint and various complimentary ground based surveys such as DES, KiDS and HSC.

### 1.3.1 Dark Energy Survey

The Dark Energy Survey (DES; DES Collaboration, 2005, 2016) is a photometric survey utilising the Dark Energy Camera (DECam; Flaugher et al., 2015) on the Blanco 4m

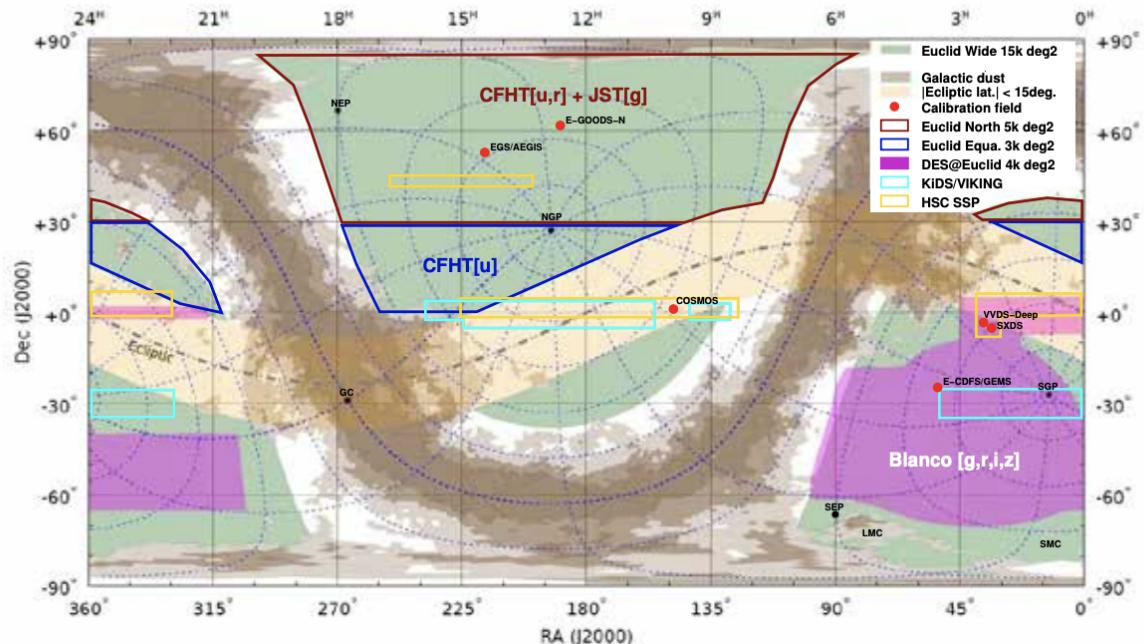


Figure 1.3: The Euclid wide survey footprint is shown in green is compared to various complimentary ground based surveys, such as DES in magenta, KiDS in turquoise and Hyper SuprimeCam (HSC) in yellow. Credit: The Euclid Consortium Newsletter

telescope at the Cerro Tololo Inter-American Observatory (CTIO) in Chile to observe  $5000 \text{ deg}^2$  of the southern sky in five broadband filters, g, r, i, z, Y, ranging from  $\sim 400 \text{ nm}$  to  $\sim 1060 \text{ nm}$  (Li et al., 2016; Burke et al., 2018). The primary goal of DES is to constrain the distribution of dark matter in the Universe, study the origin of cosmic acceleration and the amount and properties of dark energy through four key probes: weak lensing, large-scale structure, galaxy clusters, and Type Ia supernovae. Owing to its large area, depth, and image quality, at its conclusion DES will support optical identification of roughly 100,000 galaxy clusters and groups up to a redshift of  $z \approx 1$ . DES started its main survey operations in 2013, with the Year One (Y1) observational season running from August 31, 2013 to February 9, 2014 (Drlica-Wagner et al., 2018). During this period  $1839 \text{ deg}^2$  of the southern sky were observed in three to four tilings in each of the four DES bands g, r, i, z, as well as  $1800 \text{ deg}^2$  in the Y-band. These observations were processed via a variety of photometric data reduction steps into the Y1 GOLD catalogue which is the main science quality catalogue of DES. Using the fiducial multi-epoch, multi-object fitting algorithm (MOF) DES finds the  $10\sigma$  limiting magnitudes of this data set for  $2''$  apertures to be  $g \approx 23.7$ ,  $r \approx 23.5$ ,  $i \approx 22.9$  and  $z \approx 22.2$ . Due to its low depth and calibration uncertainty, Y band photometry is not used for shape measurements or photometric redshift estimations. Approximately  $1,500 \text{ deg}^2$  of the main survey were utilised, divided into two large non-contiguous areas. The reduction in the area is due to a series of survey masks. These masks are applied to avoid bright stars, satellite tracks, the Large Magellanic Cloud, among others. The two non-contiguous areas are the “SPT” area ( $1,321$  square degrees), which overlaps the footprint of the  $2,500$  square degrees South Pole Telescope Sunyaev-Zel'dovich Survey (Carlstrom et al., 2011a), and the “S82” area ( $116$  square degrees), which overlaps the Stripe-82 deep field of the Sloan Digital Sky Survey (SDSS; Annis et al., 2014). The data processing for Y5 has already been completed and the final observing season, Y6, was finalised on January 9th, 2019.

## Shear Catalogue

The  $1500 \text{ deg}^2$  of the Y1 GOLD catalogue are further processed by the METACALIBRATION algorithm (Huff & Mandelbaum, 2017; Sheldon & Huff, 2017) to define the DES Y1 weak lensing galaxy shape catalogue (Zuntz et al., 2018). The fiducial shear estimates are obtained from a single Gaussian fit by using the ngmix model-fitting algorithm (Sheldon, 2015). The produced DES Y1 METACALIBRATION catalogue has an effective source density of  $6.28 \text{ arcmin}^{-2}$ . The main systematic effect in this shape estimation is a multiplicative bias, i.e. an over- or underestimation of the gravitational shear inferred from the mean tangential ellipticity of the lensed galaxies. To characterise and correct for this bias, METACALIBRATION uses the galaxy images themselves to “de-bias” the shear estimates. This method uses the galaxy images of background sources, giving the advantage that no prior information about galaxy properties or a calibration from simulations are necessarily required. The METACALIBRATION code utilises images taken in riz bands to measure the ellipticities of the galaxies. The algorithm works by distorting deconvolving the original galaxy images from their Point Spread Function (PSF). Then a

small known positive and negative shear is applied to those deconvolved images in both ellipticity directions,  $e_1$  and  $e_2$ . The new images are then reconolved, but this time with a representation of the PSF, and new ellipticities are estimated for these images. We can use the new measurements to get an estimate of the response of a shear estimator,  $R_\gamma$ , on the applied shear by forming finite-difference central derivatives:

$$R_\gamma = \frac{\partial \mathbf{e}}{\partial \boldsymbol{\gamma}}. \quad (1.9)$$

We account for selection effects by examining the response of the selections to the shear. When calculating the mean shear over an ensemble, a weight is applied, which is effectively a kind of smooth selection. It is accounted for in the same manner. This effect is described via a selection response term  $R_s$ , which gives us a response-corrected mean shear estimate

$$\langle \boldsymbol{\gamma} \rangle \approx \langle \mathbf{R} \rangle^{-1} \langle \mathbf{R} \cdot \boldsymbol{\gamma} \rangle \approx \langle \mathbf{R} \rangle^{-1} \langle \mathbf{e} \rangle \quad (1.10)$$

with a joint response  $\mathbf{R} \approx R_\gamma + R_s$ .

This shear response term,  $\mathbf{R}$  is a  $2 \times 2$  Jacobian matrix in a celestial coordinate system for the two ellipticity components  $e_1, e_2$ . In this work we are interested in the mean *tangential* shear on the mean tangential ellipticity. Recent weak lensing shear analysis have assumed  $\mathbf{R}$  to be a scalar due to it being close to isotropic on average (Gruen & Brimioule, 2017; Chang et al., 2018; Prat et al., 2018; Troxel et al., 2018). However, like in McClintock et al. (2019) and Varga et al. (2019), we account for the fact that when considering larger tangential shear measurements on smaller scales around clusters, the response might not be completely isotropic. Hence, we explicitly rotate it to the tangential frame. Tangential ellipticity  $e_T$  is related to  $e_1, e_2$  (and likewise  $\gamma_T$  to  $\gamma_1$  and  $\gamma_2$ ) by

$$e_T = -e_1 \cos(2\phi) - e_2 \sin(2\phi), \quad (1.11)$$

where  $\phi$  is the polar angle of the source in a coordinate system centred on the lens. For the shear response, the corresponding rotation is derived from Equation 1.9 and Equation 1.11 as

$$R_{\gamma,T} = R_{\gamma,11} \cos^2(2\phi) + R_{\gamma,22} \sin^2(2\phi) + (R_{\gamma,12} + R_{\gamma,21}) \sin(2\phi) \cos(2\phi). \quad (1.12)$$

For the selection response term, such rotation cannot be performed as it is only meaningful if you consider an ensemble of galaxies. As in this case the orientation of the source galaxies should be random relative to the clusters, this suggests a symmetrised version of the response in the tangential frame:

$$\langle R_{\text{sel}}^{(T)} \rangle \approx \frac{1}{2} \text{Tr} \langle R_{\text{sel}} \rangle \text{ where } \langle R_{\text{sel}} \rangle_{i,j} \approx \frac{\langle e_i \rangle^{S+} - \langle e_i \rangle^{S-}}{\Delta \gamma_j}. \quad (1.13)$$

The uncertainties introduced from this approximation are sub-dominant due to the already small bias associated with source galaxy selection. In the above equation,  $\langle e_i \rangle^{S\pm}$  is the

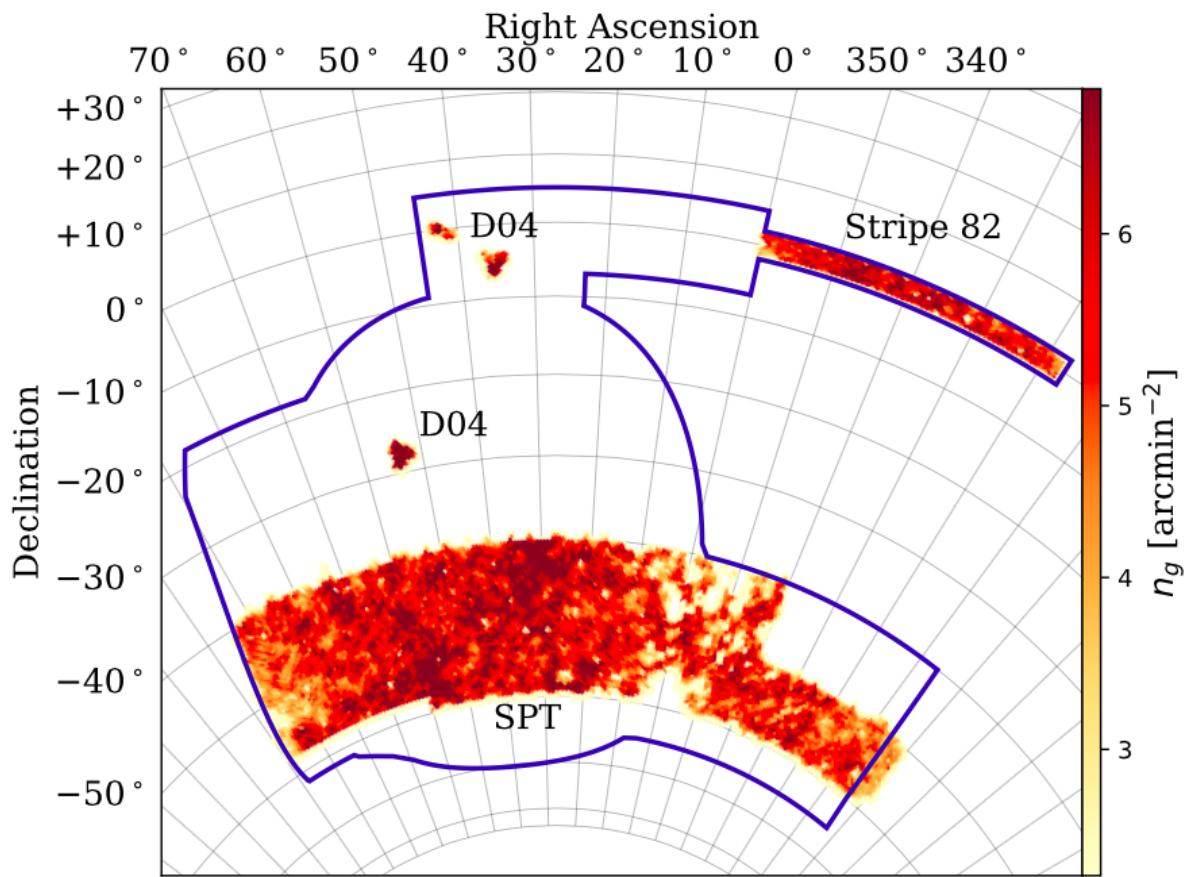


Figure 1.4: Footprint of the DES Y1 weak lensing shape catalogue, colour-coded in source density [arcmin<sup>-2</sup>]. The outline of the year 5 footprint of DES is shown in blue. Credits: Zuntz et al. (2018)

mean ellipticity measured from artificially sheared images with a selection based on un-sheared images.

The footprint of the DES Y1 weak lensing shape catalogue is shown in Fig. 1.4, overlapping with SPT and the stripe 82 from SDSS and colour-coded in source density [ $\text{arcmin}^{-2}$ ]. The outline of the year 5 footprint of DES is shown in blue.

### Photo-z catalogues

The photometric redshift catalogue of DES Y1 was produced and validated by Hoyle et al. (2018) using the template-based BPZ algorithm (Benítez, 2000; Coe et al., 2006). Two separate redshift estimates are derived: one based on the multi-epoch, multi-object fitting (MOF) based galaxy colours listed in the GOLD catalogue, and a second based on the photometric model obtained from METACALIBRATION. This second METACALIBRATION based redshift estimate is required to properly account for the selection response correction, however is found to have greater scatter compared to the MOF based redshift estimates. For this reason McClintock et al. (2019) opted to use the METACALIBRATION estimates only in selecting and weighting source-lens pairs. Hoyle et al. (2018) found these redshift estimates to be mildly biased in the mean redshift. Since the  $P(z)$  decomposition method is only sensitive to the relative shape of the  $P(z)$ -s, the impact of this bias is not expected to be significant.

#### 1.3.2 Sloan Digital Sky Survey

The Sloan Digital Sky Survey (SDSS, York et al., 2000) saw first light in 1998, with routine operations starting in the year 2000. Its goal was to obtain CCD imaging in five broad bands ( $u,g,r,i,z$ ) over 10,000 square degrees of the the high-latitude sky in the North Galactic Cap as well as spectroscopy of roughly  $10^6$  galaxies and  $10^5$  quasars over the same area. Additionally, a 275 square degree region on the Celestial Equator in the Southern Galactic Cap was imaged by SDSS, called “Stripe 82”. It was imaged multiple times, at times when the North Galactic Cap was not observable. In comparison to the SDSS single pass data, which reaches  $r \approx 22.4$  with a median seeing of  $1.4''$ , the co-added images of Stripe 82 can reach roughly 2 magnitudes deeper with a median seeing of  $1.1''$ . Stripe 82 overlaps with the DES footprint. The SDSS uses a wide-field 2.5m telescope (Gunn et al., 2006) located at the Apache Point Observatory (APO) near Sacramento Peak in Southern New Mexico. The images were predominantly taken under good seeing conditions and for stellar sources the 50% completeness limits of the images are  $u, g, r, i, z = 22.5, 23.2, 22.6, 21.9, 20.8$ , respectively (Abazajian et al., 2003). Figure 1.5 shows the footprint of SDSS photometric survey in the top panel in grey, which covers a contiguous area of the Northern Galactic Cap. Stripes indicated in blue are auxiliary imaging data, while stripes in red are part of the SEGUE survey (Yanny et al., 2009) which focuses on the astrophysics of the Milky Way Galaxy. In the bottom panel, the footprint of the spectroscopic survey is shown, again in grey.

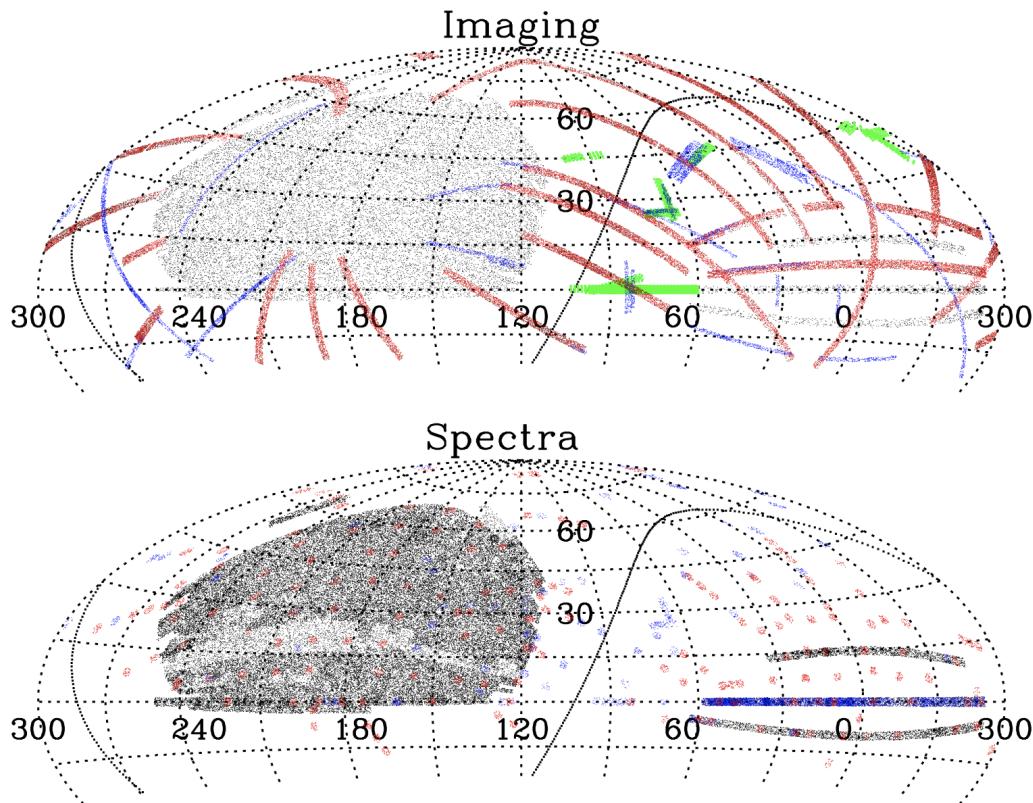


Figure 1.5: The top panel shows the footprint of SDSS photometric survey in grey, which covers a contiguous area of the Northern Galactic Cap. Stripes indicated in blue are auxiliary imaging data, while stripes in red are part of the SEGUE survey which focuses on the astrophysics of the Milky Way Galaxy. In the bottom panel, the footprint of the spectroscopic survey is shown, again in grey. Credits: Abazajian et al. (2009)

### 1.3.3 Kilo-Degree Survey

The Kilo-Degree Survey (KiDS) is an ongoing optical wide-field imaging survey carried out with the OmegaCAM camera (Kuijken, 2011) at the VLT Survey Telescope. It is currently in its fourth public data release (DR4, Kuijken et al., 2019), which has more than doubled the covered area since the previous data release DR3. KiDS was specifically designed with the purpose to measure weak gravitational lensing by galaxies and large-scale structure and eventually to constrain the equation-of-state of Dark Energy. Upon its completion it will have imaged 1350 square degrees in four filters (u,g,r,i). As of this data release, the aperture matched photometry from the companion VISTA Kilo-degree INfrared Galaxy survey (VIKING; Edge et al., 2013) observed by the VISTA telescope was added to the photometric catalogue. It covers the same area in five near-infrared bands: Z,Y,J,H,K. The combination of these two surveys has led to the creation of a nine-band matched aperture u-Ks catalogue, containing roughly 100 million galaxies. The depth of the survey is quantified by a signal-to-noise of  $5\sigma$  for point sources in 2" apertures. Its limiting AB magnitudes are  $u \approx 24.23$ ,  $g \approx 25.12$ ,  $r \approx 25.02$  and  $i \approx 23.68$ . This KiDS/VIKING combined data set is so far the largest-area optical+infra-red survey to this depth. Figure 1.6 shows the full KiDS+VIKING footprint in grey, tiles in green are those released for the first time in DR4 and tiles in blue were already included in previous data releases. DES Y1 and KiDS have been observed to similar depths, but KiDS exhibits superior seeing compared to DES which explains why both surveys are providing cosmic shear constraints of comparable power, despite the larger DES area.

### 1.3.4 Rubin

The Vera C. Rubin Observatory, previously known as the Large Synoptic Survey Telescope (LSST), is currently under construction in Chile with first light expected in 2021 and full operations to start in October 2022 (Ivezic et al., 2019). The goal of the Rubin project is to conduct the 10-year Legacy Survey of Space and Time (LSST), with the “simple” objective to conduct a deep survey over a vast sky area, with a frequency that enables images of every part of the visible sky to be obtained every few nights and continue in this mode for ten years to achieve astronomical catalogues thousands of times larger than have ever previously been compiled. Hence, Rubin will be able to address some of the most sought after questions about the evolution of the Universe, its structure and the objects within it. Some of its particular science goals include probing dark matter and dark energy via weak gravitational lensing, baryon acoustic oscillations, and photometry of type Ia supernovae; mapping objects in the solar system like near-Earth asteroids; detecting transient events such as novae, supernovae, and gamma-ray bursts; and finally mapping the Milky Way. The survey is going to cover approximately 18,000 deg<sup>2</sup> of the southern sky, imaging billions of objects in six colours in its main survey and also records the time evolution of these sources: the first motion picture of our Universe. The  $5\sigma$  limiting magnitudes are expected to be  $r < 24.5$  in single images and  $r < 27.8$  in the stacked data. About 90% of the observing time will be used by the main survey, whereas the remaining 10%

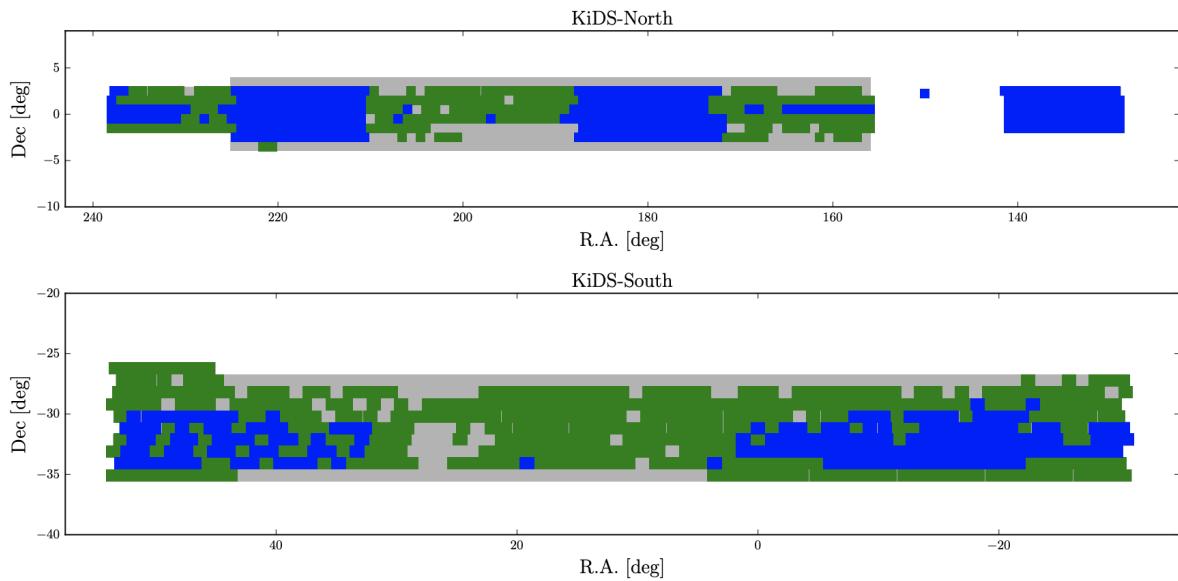


Figure 1.6: Sky distribution of survey tiles released in KiDS-DR4. Tiles in green were released for the first time in DR4; those in blue were included in the earlier data releases (DR1+2+3) but have been reprocessed for DR4. The full KiDS+VIKING area (1350 square degrees) is shown in grey. Top: KiDS-North. Bottom: KiDS-South. Credits: Kuijken et al. (2019)

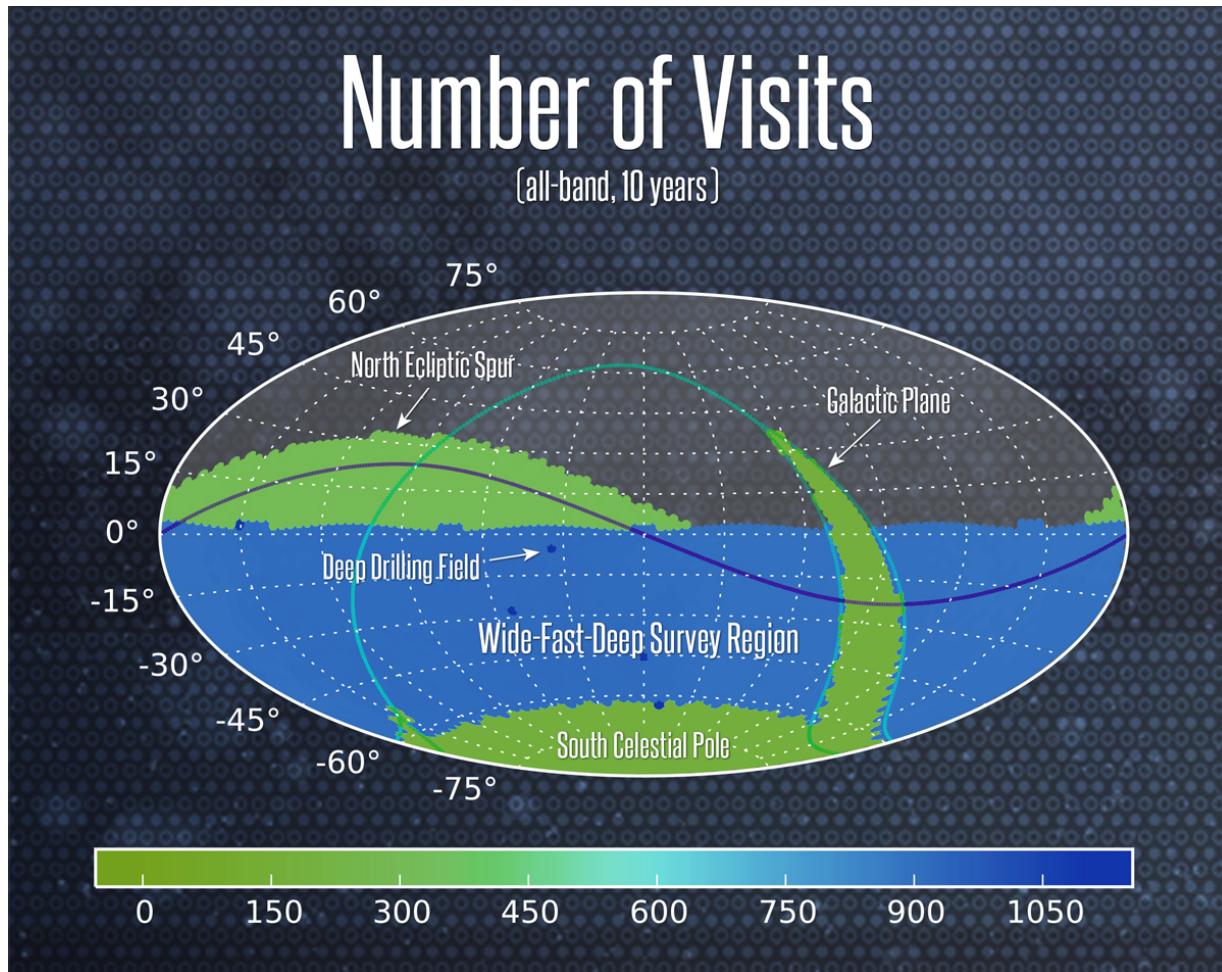


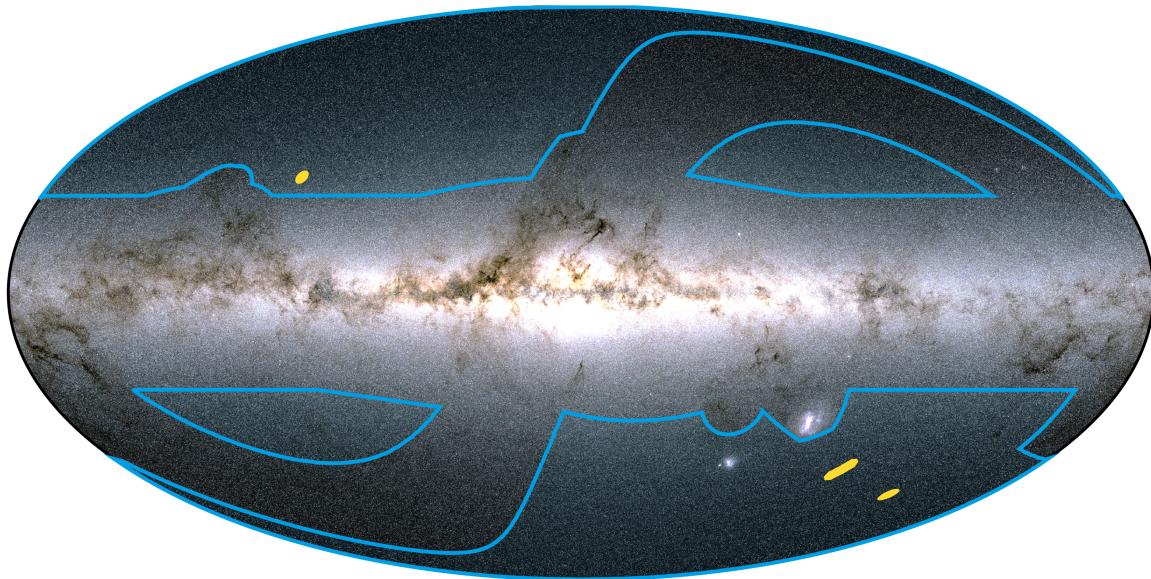
Figure 1.7: Simulation of the field selection and image acquisition process of the LSST over the 10-year life of the planned survey. The main survey will cover about  $18,000 \text{ deg}^2$ . Credit: Rubin Obs/NSF/AURA ([lsst.org](http://lsst.org))

will be used to acquire coverage for specific regions and goals. This includes very deep ( $r \sim 26$ ) observations, observations of “special” regions such as the Ecliptic, Galactic plane, and the Large and Small Magellanic Clouds, and areas covered in detail by multi-wavelength surveys such as COSMOS. Combined, this will increase the total area covered by Rubin to about 25,000 deg<sup>2</sup>. Figure 1.7 shows a simulation of the field selection and image acquisition process of the Rubin over the 10-year life of the planned survey. The simulation was created by the Operations Simulator, developed by the LSST Project, to verify that the Rubin science requirements could be met with the telescope design. The 8.4-meter Simonyi Survey Telescope, part of the Rubin Observatory, is located on the Cerro Pachón ridge in north-central Chile. It used a special three-mirror design allowing for an exceptionally wide field of view and combined with its large aperture will give it an extraordinarily big etendue. In fact, 3 times larger than the best existing telescopes like the Subaru Telescope with its Hyper Suprime Camera and more than an order of magnitude better than most large telescopes. The Rubin Observatory Camera is the largest digital camera ever constructed for the field of astronomy and contains over three billion pixels of solid state detectors. Even though the camera has six filters (ugrizy) covering 330 to 1080 nm wavelengths, its position between specific mirrors limits the size of its filter changer. It can only hold five filters at a time, so each day one of the six must be chosen to be omitted for the following night.

### 1.3.5 Euclid

With the current launch scheduled to be in 2022, the Euclid mission aims to survey more than 15,000 square degrees, called the wide survey (Laureijs et al., 2011). By observing with visible imaging and in near-infrared photometry, the survey will produce a large data set of about 1.5 billion resolved galaxies and of those roughly 50 million galaxies will be observed in near-infrared spectroscopy. Its main mission goals are to investigate cosmological phenomena like the evolution of galaxy clustering and gravitational weak lensing due to the presence of baryonic and dark matter between the observer and the background galaxies, helping us to characterise the expansion history of the Universe as well as understanding its acceleration, thought to be caused by dark energy. In addition to the wide survey, Euclid will also observe 3 deep surveys, two of which are spanning 10 square degrees and one spanning 20 square degrees. These deep surveys are two magnitudes deeper than the wide survey, thereby providing a useful data set to validate the main cosmological analysis of the wide survey, but also, due to returning to the same patch of sky multiple times, represents an important way to monitor stability of the mission and is essential for calibration purposes. The full footprint of the Euclid Surveys can be seen in Figure 1.8.

Euclid will produce shapes and masses of galaxies as well as star formation rates within those galaxies with a four times better resolution as well as a 15 times better sensitivity in the near-infrared than what would be possible from any ground based observation like DES. Especially, obtaining data comparable to Euclid’s deep fields from the ground would require several tens of years of continuous observing time from the best near-infrared facilities. For weak lensing, Euclid will observe at a depth of roughly 30 galaxies per arcmin<sup>2</sup> (roughly



The Euclid Wide Survey and the Euclid Deep Survey [Mollweide Galactic]

Euclid Wide Survey : 15,000 deg.<sup>2</sup>

Euclid Deep Fields : North=10 deg.<sup>2</sup>, Fornax=10 deg.<sup>2</sup>, South=20 deg.<sup>2</sup>



Background: ESA/Gaia/DPAC & Euclid Consortium

Figure 1.8: All-sky map showing the location of Euclid's wide survey in blue and the three complimentary deep field surveys in yellow. Two of the deep fields are spanning 10 square degrees, whereas the third deep field is spanning 20 square degrees. Credit: ESA/Euclid Consortium. Acknowledgment: Euclid Consortium Survey Group, Copyright: ESA/Gaia/DPAC; Euclid Consortium.

3 times more than DES) observed in one broad R+I+Z band, which corresponds to a detection limit of  $10\sigma$  at an AB magnitude of 24.5. The corresponding redshifts are derived from 3 additional near-infrared bands (Y,J,H) at a limiting AB magnitude of 24 ( $5\sigma$ ). This is 3 magnitudes deeper than what can be achieved from ground based observations over such a large area.

## 1.4 Outline and Motivation

In this introduction, the importance of an accurate cluster mass calibration for a solid cosmological analysis was highlighted, specifically by accounting for systematic uncertainties. Three different ways of detecting galaxy clusters were discussed and a common means of estimating clusters masses was introduced, namely gravitational weak lensing. It was argued why this method might be superior to others when it comes to measuring cluster masses. Finally, the wide-field imaging survey DES was introduced in detail, and comparisons to other imaging surveys like SDSS, KiDS, Rubin and Euclid were made.

In Chapter 2 we consider the so called cluster member contamination, an important systematic uncertainty affecting the weak lensing mass calibration of galaxy clusters. Cluster own galaxies can scatter into your background sample due to photometric uncertainties on their redshift estimates. This leads to a dilution of the overall shear measurement, as one is averaging over galaxies that have no shear signal at all. Therefore, in order to recover the true shear signal, the measured shear signal must be boosted. This can be done by applying a cluster member contamination correction, which involves estimating a contamination fraction based on the decomposition of the photometric redshift probability distribution function estimates,  $P(z)$ , of source galaxies. We apply this method to two cluster samples, the Sunyaev-Zel'dovich selected cluster sample SPT (Bleem et al., 2015b) and the X-ray selected cluster sample MARD-Y3 (Klein et al., 2019), over the DES Y1 footprint.

In Chapter 3 we present an analysis of the cluster contamination correction on the optically selected cluster sample redMaPPer (Rykoff et al., 2016a), which represents a 10 times larger cluster sample than the above mentioned cluster samples. The greatly extended sample size helps to reduce the uncertainty on the measurements and we compare these results to our findings for the SPT and MARD-Y3 cluster samples. Finally, we extract the cluster contamination with the same parameters as it was done in Varga et al. (2019) in order to meaningfully compare both works and estimate the impact on the cluster masses for any differences that were found.

In Chapter 4 we close with concluding remarks and a discussion on future works.

# Chapter 2

## Cluster Member Contamination of the Weak Lensing Shear Profiles from the Dark Energy Survey Year 1 Data: Application to SPT and MARD-Y3 clusters

M. Paulus<sup>1,2</sup>★, J. J. Mohr<sup>1,2</sup>, S. Grandis<sup>1</sup>, S. Bocquet<sup>1</sup>, M. Klein<sup>1</sup>,

2021, to be submitted to Monthly Notices of the Royal Astronomical Society (This draft is currently being reviewed by the collaborators)

### 2.1 Abstract

Cluster member galaxies can contaminate weak lensing source galaxy catalogues, thereby limiting the accuracy of cluster mass determinations. We employ an estimator for the fraction of cluster contaminants by decomposing the observed redshift distribution  $P(z)$  into cluster member and local field components. We apply our method to 250 South Pole Telescope SZE selected and to 208 MARD-Y3 X-ray selected cluster samples. The Dark Energy Survey Year 1 footprint overlaps these samples and allows for competitive weak lensing mass calibration. We model the cluster contamination as a Gaussian offset from the cluster redshift with amplitude depending on cluster richness and falling off as a projected NFW profile with distance from the cluster centre. We constrain the model parameters using the entire cluster sample. The data are adequately described by a single NFW concentration, with a power law variation of the amplitude with cluster richness. However, the redshift dependence of the amplitude is more complex and so we adopt discrete redshift

normalisations that are all fit simultaneously. We perform validation tests that allow us to justify our choice of model, and these include comparing the model prediction of the contamination fraction with the field subtracted  $P(z)$ . We find excellent agreement. This work is a prerequisite for the upcoming cosmological analyses of the two cluster samples.

## 2.2 Introduction

Galaxy clusters and their abundance and distribution in the Universe are regarded as powerful cosmological probes (Haiman et al., 2001; Allen et al., 2011; Dodelson et al., 2016). This presents two major tasks for a cosmological analysis: first, detecting galaxy clusters and then second, estimating their masses by calibrating their mass-observable relation (MOR), a way to compare observed and predicted cluster abundances. The main methods for identifying galaxy clusters include their observed galaxy populations in optical wavelengths (e.g. Koester et al., 2007; Oguri, 2014; Rykoff et al., 2016a; Bellagamba et al., 2017), their X-ray emission (Vikhlinin et al., 1998; Mantz et al., 2010; Klein et al., 2019) and their Sunyaev-Zel'dovich effect (SZE; Sunyaev & Zeldovich, 1970, 1972b; Staniszewski et al., 2009; Bleem et al., 2015b; Planck Collaboration et al., 2016b; Hilton et al., 2018). The SZE arises when Cosmic Microwave Background (CMB) photons inverse Compton scatter off hot electrons from the intra-cluster medium (ICM) and the X-ray emission is due to bremsstrahlung and line emission from the ICM. These methods for identifying clusters do not provide an accurate measure of the cluster masses, and thus the masses are estimated through their MOR that are calibrated using additional information.

One of the most promising ways to determine cluster masses today is via gravitational weak lensing. The induced tangential alignment of background galaxies around a foreground cluster is directly sensitive to the cluster mass. Hence, many studies have been carried out to specifically estimate cluster masses using gravitational weak lensing (Sheldon et al., 2001; Johnston et al., 2007; von der Linden et al., 2014a,b; Applegate et al., 2014; Hoekstra et al., 2015; Mantz et al., 2015; Okabe & Smith, 2016; Battaglia et al., 2016; Melchior et al., 2017; Simet et al., 2017; Murata et al., 2018; Dietrich et al., 2019; McClintock et al., 2019). Weak lensing cluster mass estimates come with a number of uncertainties, both systematic and statistical. As sky surveys are becoming larger, deeper and wield a larger statistical power it is becoming more important to properly characterise the systematic uncertainties. A recent analysis by McClintock et al. (2019) has shown that the stacked weak lensing mass estimates are for the first time dominated by systematic uncertainties. Therefore, in order to improve the cosmological constraining power from clusters, one has to first gain a better understanding of the systematic uncertainties affecting the cluster mass calibration.

An important systematic uncertainty impacting the weak lensing analysis results from cluster galaxies that scatter into the background sample of weak lensing galaxies. This is due to the photometric uncertainty on the redshift estimates. The contamination dilutes the shear measurement, because we are averaging over galaxies with no shear signal at all, which makes it necessary to boost the measured signal in order to recover the true signal.

Hence, many studies refer to this rescaling of the weak lensing signal as applying a *boost factor* (Sheldon et al., 2004; Applegate et al., 2014; Hoekstra et al., 2015; Gruen et al., 2014; Simet et al., 2017; Melchior et al., 2017; Medezinski et al., 2018a; Leauthaud et al., 2017; McClintock et al., 2019; Varga et al., 2019).

Multiple approaches in characterising cluster member contamination have been made use of in previous studies: Sheldon et al. (2004) and Simet et al. (2017) estimated the boost factor profiles from the transverse correlation of source galaxies around cluster centres, whereas others (Applegate et al., 2014; Medezinski et al., 2018a,b) applied a “colour-cut” method making use of colour information. Gruen et al. (2014) and Dietrich et al. (2019) estimated the cluster contamination by decomposing the source galaxies into a cluster member component and a background component. This method was expanded by Melchior et al. (2017) who estimated the contamination fraction based on a decomposition of the photometric redshift probability distribution function estimates,  $P(z)$ , of source galaxies. The same method has been used since (Chang et al., 2018; Stern et al., 2019; McClintock et al., 2019; Pereira et al., 2020) and has been validated by Varga et al. (2019) on simulations.

In this work, we employ the  $P(z)$  decomposition method to estimate the cluster member contamination for the SPT and MARD-Y3 cluster mass calibration over the Dark Energy Survey (DES) Year 1 (Y1) footprint. DES (DES Collaboration, 2005) is a wide-field imaging survey located in the southern sky that covers an area of  $5000 \text{ deg}^2$ , yielding approximately 300 million galaxies out to  $z = 1.4$ . At present, the available DES Y1 footprint covers about  $1800 \text{ deg}^2$  and shares a large overlap with the observed area of the South Pole Telescope (SPT, Carlstrom et al. (2011b)). Hence, the Sunyaev-Zel’dovich effect selected cluster sample from SPT (Bleem et al., 2015b) is an ideal candidate for this analysis. The  $P(z)$  decomposition method was already used by Stern et al. (2019) to aid the SPT weak lensing mass calibration over the DES Science Verification (SV) area. Here we are looking at a significantly larger cluster sample as well as the greatly extended footprint of DES Y1 compared to SV. In addition, we are studying the X-ray selected cluster sample MARD-Y3 (Klein et al., 2019) over the DES Y1 footprint. Furthermore, the SPT cluster sample has previously been used to constrain cosmology (Bocquet et al., 2019) and the results from this work will be part of the upcoming SPT cluster cosmology analysis (Bocquet et al, in prep.). In this analysis we are particularly interested in modelling the cluster contamination in a manner that allows us to fit for a variable normalisation in cluster redshift.

This paper is structured in the following way: In section 2.3 we present the DES Y1 data used in this work as well as the different cluster samples. In section 2.4 we introduce the framework as well as the formalism of the  $P(z)$  decomposition method and in section 2.5 we present our results. We assume a flat  $\Lambda\text{CDM}$  cosmology with  $\Omega_m = 0.3$  and  $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$  throughout.

## 2.3 Data

In the following subsections we describe the shear and photometric redshift (photo-z) catalogues from DES as well as the X-ray and SZE selected cluster catalogues from MARD-Y3 and SPT, respectively.

### 2.3.1 DES Y1 shear and photo-z catalogues

In the DES Y1 observations, roughly  $1800 \text{ deg}^2$  of the observed  $5000 \text{ deg}^2$  survey were covered in each of the four bands, g, r, i, z. Even though the imaging is shallower than the SV data, it covers a significantly larger area. These observations were processed via a range of photometric data reduction steps into the Y1 GOLD catalogue (Drlica-Wagner et al., 2017). This catalogue represents the main science catalogue of DES. It was decided that Y1 should target areas of the southern sky overlapping with the SPT footprint ( $1321 \text{ deg}^2$ ) and the "stripe 82" deep field region ( $116 \text{ deg}^2$ ) of the Sloan Digital Sky Survey (SDSS). After masking only about  $1500 \text{ deg}^2$  of the main survey is utilised. This area was then further processed by the METACALIBRATION algorithm (Huff & Mandelbaum, 2017; Sheldon & Huff, 2017) to define a shear galaxy catalogue (Zuntz et al., 2018). This shear catalogue contains ellipticity measurements  $e_i$  for each galaxy and corresponding shear response terms,  $R = R_\gamma + R_{sel}$ , that describe the change in ellipticity of each source due to a small applied shear. The photometric redshifts  $P(z)$ 's were calculated via the BPZ template based algorithm (Hoyle et al., 2018). Two different redshift estimates were derived. One is based on the galaxy colours, which are extracted using the multi-epoch, multi-object fitting algorithm (MOF), and are listed in the GOLD catalogue. The second is based on a model obtained from METACALIBRATION. These second redshift estimates are required to correctly account for the selection response correction, but they were found to show a larger scatter compared to the MOF estimated redshifts. Therefore, in previous DES cluster weak lensing analyses (McClintock et al., 2019; Pereira et al., 2020), the redshift estimates from METACALIBRATION were used only for selecting and weighting source lens pairs and the MOF based photo-z were used for determining the resulting source redshift distributions. In our analysis, we follow the same practice.

### 2.3.2 MARD-Y3 cluster catalogue

The second ROSAT All-Sky Survey source catalogue 2RXS (Boller et al., 2016) was used to produce the X-ray selected MARD-Y3 cluster catalogue over the full  $5000 \text{ deg}^2$  footprint of DES (Klein et al., 2019). The 2RXS is based on the RASS-3 processed photon event files and uses an improved methodology compared to the 1RXS catalogues (Voges et al., 1999, 2000). The full 2RXS catalogue contains 135,000 source of which  $\sim 30\%$  can be attributed to spurious sources (Boller et al., 2016). Around  $\sim 20,000$  of the total number of sources are within the DES footprint, and only 1,000-2,000 sources were expected to be clusters (Henry et al., 2006; Ebeling et al., 2013; Klein et al., 2018). Due to this low number an optical confirmation to identify a 2RXS source as a cluster is needed. Furthermore, due

to the clustering of a large number of optical systems as well as the high density of 2RXS sources, the likelihood of chance superpositions is significant. Hence, the probability that a 2RXS source with an optical counterpart is actually a cluster needs to be characterised. The Multi-Component Matched Filter Cluster Confirmation tool (Klein et al., 2018, MCMF) was developed to do exactly that. MCMF is designed for automated confirmation and redshift estimation using large scale imaging surveys such as DES. To be more precise, the internally available DES Y3A2 GOLD catalogue was used, which is a value-added version of the previously published DES DR1 dataset (Abbott et al., 2018).

To confirm a 2RXS source as an X-ray detected cluster, the so called red sequence (RS) technique is used, which takes advantage of the colour-magnitude-redshift dependency of passively evolving galaxies (Gladders & Yee, 2000). These RS galaxies dominate the galaxy population within the cluster virial region out to  $z \sim 1$  (Hennig et al., 2017), and even clusters at higher redshift contain significant numbers of RS galaxies (e.g., Strazzullo et al., 2019). Specifically, a scan of the RS population as a function of redshift is carried out at the location of each 2RXS source. The region scanned corresponds to the radius  $r_{500}$  that the cluster would have if all the X-ray flux could be attributed to the cluster. If one or more significant peaks are identified, then the richness and redshift of each peak is recorded. Only a few percent of the confirmed clusters have more than one significant RS peak along the line of sight (Klein et al., 2019). To eliminate contamination of the resulting cluster catalogue by chance superposition of physically independent X-ray point sources and RS galaxy clumps, Klein et al. (2019) compared the identified overdensities of each 2RXS source with those found along random lines of sight with similarly sized radial apertures and redshifts. For counterparts with richesses similar to those found at the same redshift along random lines of sight the probability is significant that they are simply random superpositions. By comparing the richness distributions of the 2RXS optical counterparts with those of random lines of sight it is then possible to quantify the incidence of random superpositions or contamination in the final cluster catalogue. The estimator  $f_{\text{cont}}$  contains this information. For a cluster candidate  $i$  with richness  $\lambda_i$ ,  $f_{\text{cont},i}$  is defined as

$$f_{\text{cont},i} = \frac{\int_{\lambda_i}^{\infty} f_{\text{rand}}(\lambda) d\lambda}{\int_{\lambda_i}^{\infty} f_{\text{obs}}(\lambda) d\lambda}, \quad (2.1)$$

where  $f_{\text{obs}}$  and  $f_{\text{rand}}$  are the smoothed distributions of richesses of the observed 2RXS sources (obs) and random lines of sight (rand) within  $\Delta z \approx 0.025$ . Cutting a candidate list at, for example, a particular value of  $f_{\text{cont}} < 0.05$  then produces a cluster catalogue with a fixed 5% contamination fraction, independent of redshift.

As noted previously, the line of sight toward each 2RXS source is searched for peaks in richness  $\lambda$ , and the 3 most significant peaks are fitted with  $\lambda(z)$  profiles derived using an RS model. The peak with the lowest  $f_{\text{cont}}$  value is then considered to be the best counterpart. The redshift corresponding to that peak is then the assigned cluster redshift,  $z_{\text{MCMF}}$ .

While typically X-ray positions are known to provide a good proxy for cluster centres, the large PSF of RASS and the low signal to noise of the 2RXS sources cause a large uncertainty on the position. Hence, the identification of cluster centres using optical data

is especially important and MCMF provides three different centres based on that data. The first estimate uses the centres that are identified in the density map of red sequence galaxies using SExtractor (Bertin & Arnouts, 1996). The richest peak identified within  $r_{500}$  of the X-ray position is chosen in the case of there being multiple peaks. The second estimate of the cluster centre is a by-product of the estimator of the dynamical state, which is based on a fit of a two dimensional King profile (King, 1962) to the RS galaxy density map. The fit is performed within the region of radius  $r_{500}$  centred on the X-ray source. The third approach relies on identifying the red brightest cluster galaxy (rBCG) within 1.5 Mpc of the X-ray source that has all colours within  $3\sigma$  of the RS at the cluster redshift. While the rBCG potentially provides the best proxy for the cluster centre available for this survey, its automated identification is not always successful.

By comparing the different cluster centre estimates one can test their reliability and identify possible failures in one of the estimates. The rBCG is taken to be the default centre as long as it is within 60 arcsec of the galaxy density peak, otherwise the galaxy density peak is used as the cluster centre.

In our analysis we use the optical centres of the optical counterparts rather than the X-ray centres of the 2RXS sources, due to the bad resolution of 2RXS in survey mode (compared to current telescopes like XMM or Chandra and also eRosita). As the contamination level of 2RXS is  $\sim 90\%$ , an  $f_{\text{cont}}$  cut of 0.05 was selected when the catalogue was created. This cut was chosen to ensure that the residual contamination is not larger than the poisson errors. The final cluster sample considered in our analysis contains 208 clusters ranging from  $0.15 < z_{\text{cl}} < 0.9$ . Clusters that were located too close to the edge of the DES Y1 footprint were removed, as well as those lying in the more sparsely populated regions of the footprint, meaning that a large portion of a radial bin is not inhabited by background galaxies (roughly 25%). This was done by visual inspection. This introduces a bias in the number of background sources detected for a certain cluster. The inspection of clusters was performed by eye. Figure 2.1 shows the mass and redshift distribution of the MARD-Y3 cluster sample (Klein et al., 2019).

### 2.3.3 SPT cluster catalogue

SPT (Carlstrom et al., 2011b) is a 10 m telescope located within  $\sim 1$  km of the geographical South Pole. Due to its resolution of  $\sim 1$  arcmin and 1 degree field of view, the telescope is well suited for a survey of high-mass clusters through the SZE signatures. From 2007 to 2011, the telescope observed three millimetre-wave bands centred at 95, 150, and 220 GHz with the SPT-SZ camera. Most of this time was spent on the SPT-SZ survey, a  $2500 \text{ deg}^2$  area located in the southern sky within the boundaries  $20h \leq \text{R.A.} \leq 7h$  and  $-65^\circ \leq \text{Dec.} \leq -40^\circ$ . The survey region was mapped to depths of approximately 40, 18, and  $70 \mu\text{K}\text{-arcmin}$  at 95, 150, and 220 GHz, respectively. Galaxy clusters are detected via their thermal SZE signature in the 95 and 150 GHz maps. For our sample of clusters detected in the SPT-SZ survey (Bleem et al., 2015b), we are primarily focusing on the optical cluster centres. The optical centres estimated for SPT are likely to be better than the MARD-Y3 optical centres as the input positions and mass proxy are better. SPT-SZ

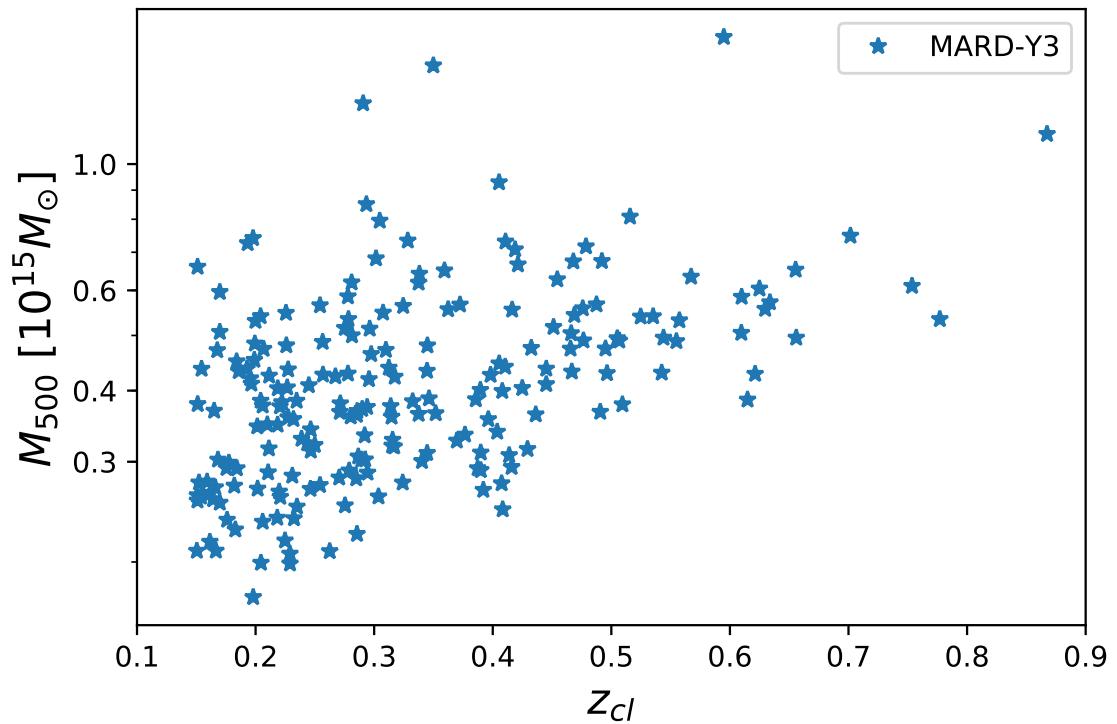


Figure 2.1: The distribution in mass and redshift of the 208 MARD-Y3 clusters ranging with  $0.15 < z_{cl} < 0.9$ . An  $f_{\text{cont}}$  cut of 0.05 was applied resulting in a 5% contamination fraction that is constant with redshift.

centres are found to be quite robust as well (Song et al., 2012; Saro et al., 2015; Bleem et al., 2020). We adopt the selection  $S/N \gtrsim 4$  to increase the sample size over which we can fit the contamination correction. This extended sample will be described in Klein et. al., in prep. The contamination level of the SPT catalogue is at  $\sim 50\%$  for the  $S/N \gtrsim 4$  sample, so we adopted an  $f_{\text{cont}}$  value of 0.1. This means that we have roughly the same amount of contamination by chance superposition as in the MARD-Y3 cluster sample due to the MARD-Y3 contamination level being roughly twice as large. Cluster redshifts and richnesses are estimated also via MCMF, adapted to fit an SZE selected cluster sample. The final cluster sample considered in our analysis contains 250 clusters ranging from  $0.15 < z_{\text{cl}} < 1.2$ . The lower redshift cut was adopted to allow for a larger sample size as well as a larger variety in cluster redshifts over which the contamination correction can be fit. The upper redshift cut was chosen to match the MARD-Y3 cluster sample. Clusters that were positioned too close to the edge of the DES Y1 footprint were removed, as well as those lying in the more sparsely populated regions of the footprint, again meaning that a large portion of a radial bin is not inhabited by background galaxies (roughly 25%). This was done by visual inspection. Figure 2.2 shows the mass and redshift distribution of the SPT clusters studied here.

### 2.3.4 SPT + MARD-Y3 cluster sample

We create a combined cluster sample with the aim to test our model on a larger sample size. The cluster sample was created by discarding duplicate clusters that occur in both samples. This is done by matching their positions as well as redshift and henceforth discarding the MARD-Y3 cluster. The resulting cluster sample contains 397 clusters ranging from  $0.15 < z_{\text{cl}} < 1.2$ .

There is an expected small scatter and bias between the MCMF richnesses on SPT and on MARD-Y3, caused by the scatter in position and redshift as well as in the mass proxy used to define  $r_{500}$  for the MCMF richnesses. Given that there is a bias in the mass proxy and different scatter between true cluster centre and SZE/X-ray position entering MCMF one gets a systematic difference in lambda. Furthermore, there is scatter in the mass proxies compared to true mass which drives the scatter between both lambdas as well. As we always choose the SPT cluster out of the overlapping clusters, and they have on average a lower richness compared to the MARD-Y3 counterpart, the median redshift of the combined sample is lower than the medians of the individual samples.

## 2.4 Cluster member contamination

The cluster contamination affects the weak lensing measurement in such a way that the shear signal in annuli around the cluster is diluted due to the presence of cluster galaxies in the background sample. With a measure of the contamination, the bias in the weak lensing signal can be corrected. The method described here estimates a cluster contamination fraction,  $f_{\text{cl}}$ , which is the fraction of cluster member galaxies that scatter into

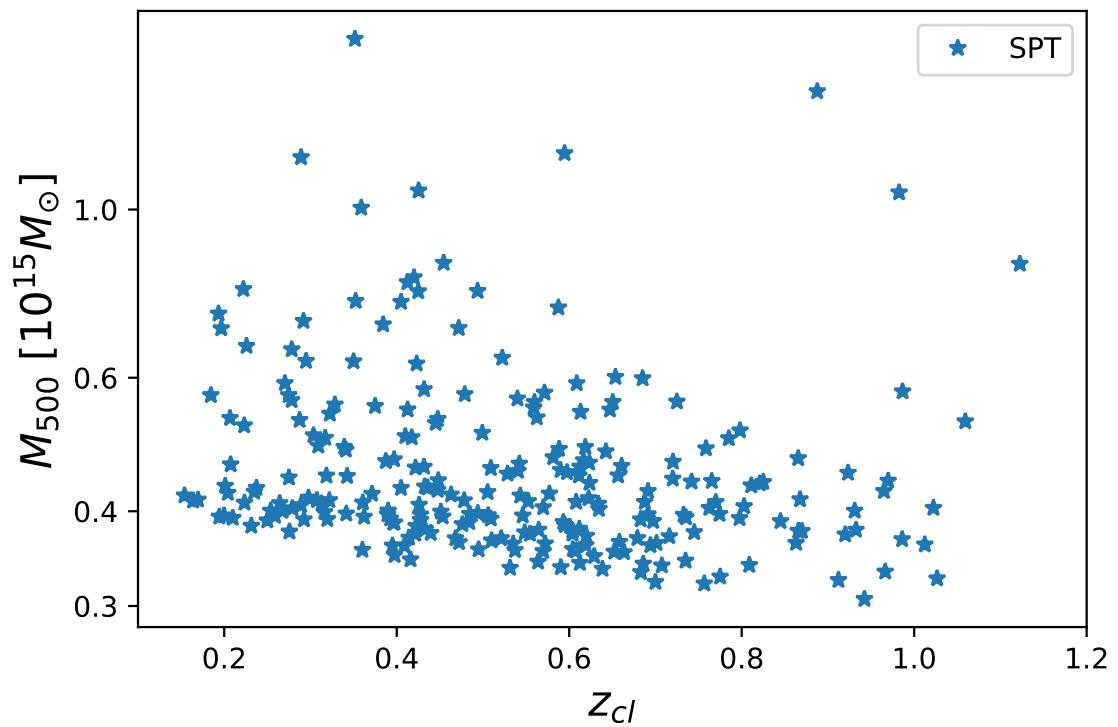


Figure 2.2: The distribution in mass and redshift of the 250 SPT clusters in the DES-Y1 footprint with  $0.15 < z_{\text{cl}} < 1.2$ . An  $f_{\text{cont}}$  cut of 0.1 was applied resulting in an overall 5% contamination fraction comparable to that within the MARD-Y3 sample.

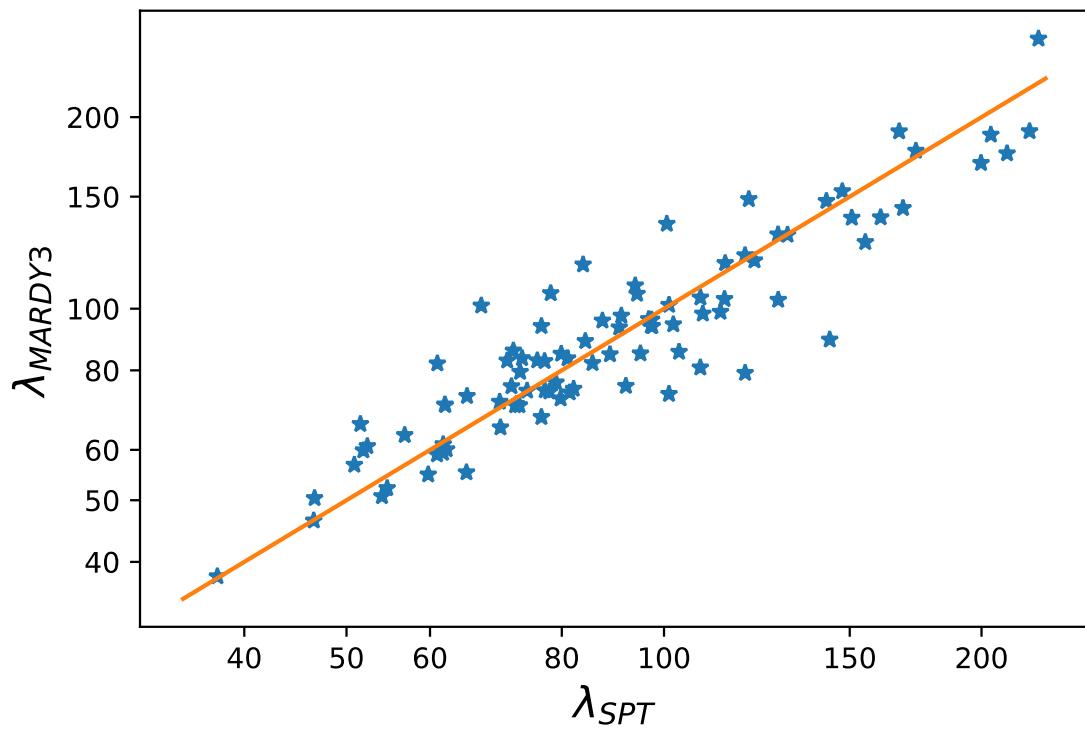


Figure 2.3: Comparison between the richness values of the overlapping SPT and MARD-Y3 clusters. A small scatter and bias between the MCMF richnesses on SPT and on MARD-Y3 is expected which is caused by the scatter in position and redshift as well as in the mass proxy used to define  $r_{500}$  for the MCMF richness.

the background shear sample used to constrain the cluster mass.  $f_{\text{cl}}$  is inferred from the photometric redshift probability distribution function  $P(z)$  of the appropriately selected and weighted source galaxies behind the cluster. We describe the method in the following sections.

### 2.4.1 Weak lensing

Analysing galaxy clusters using gravitational weak lensing requires a large sample of background source galaxies. Their images are distorted due to the gravitational potential of the cluster itself. This effect can be used to trace the underlying matter distribution of the lens (Bartelmann & Schneider, 2001).

To construct a catalogue of background source galaxies, photometric algorithms are used to extract estimates of their photo-z's. These redshifts come with large uncertainties (Hildebrandt et al., 2017; Hoyle et al., 2018), which consequently means that the source galaxy catalogue might include foreground galaxies or galaxies at the cluster redshift even after an attempt to exclude them using their photo-z's. In this work we use information from the probability distribution of the galaxy photo-z to constrain the contamination.

The fact that background galaxies at different redshifts contribute to the measured shear signal with different amplitudes is accounted for by the inverse of the *critical surface density*:

$$\Sigma_{\text{crit}}(z_l, z_s) = \frac{c^2}{4\pi G} \frac{D_s(z_s)}{D_l(z_l) D_{ls}(z_l, z_s)}, \quad (2.2)$$

where  $D_s$ ,  $D_l$  and  $D_{ls}$  denote *angular diameter* distances to the source galaxy, the lens, and between the lens and the source, respectively.

As source redshifts are generally estimated from photometry and are described by a probability distribution  $p_{\text{phot}}(z_s)$  for each source galaxy, one can estimate an *effective* critical surface density as

$$\langle \Sigma_{\text{crit}}^{-1} \rangle_{i,j} = \int dz_s p_{\text{phot}}(z_{s,i}) \Sigma_{\text{crit}}^{-1}(z_{l,j}, z_{s,i}), \quad (2.3)$$

where  $i$  and  $j$  are the indices of the source and the lens in a particular lens-source pair.

In the limit of weak gravitational lensing, we are interested in the reduced tangential shear component which is defined as  $g_T = \gamma_T / (1 - \kappa)$  where  $\gamma$  is the weak lensing shear and  $\kappa$  is the convergence. These are estimated from the shapes and alignments of the source galaxies through the ellipticity measurement  $\mathbf{e}$ , with the assumption that  $\langle \mathbf{e} \rangle \approx \langle \mathbf{g} \rangle$ . The shear signal can be related to the *excess surface mass density*  $\Delta\Sigma$ , by considering the mass distribution of the cluster:

$$\gamma_T(R) = \frac{\bar{\Sigma}(< R) - \bar{\Sigma}(R)}{\Sigma_{\text{crit}}} = \frac{\Delta\Sigma(R)}{\Sigma_{\text{crit}}}, \quad (2.4)$$

where  $\bar{\Sigma}(< R)$  represents the average surface mass density within a projected radius  $R$  and  $\bar{\Sigma}(R)$  refers to the average surface mass density at  $R$ , where  $R$  is the projected separation from the lens.

### Lensing estimator

Following Sheldon et al. (2004) one can define the maximum likelihood estimator for the shear as:

$$\widetilde{\gamma}_T(R) = \frac{\sum_i^s \sum_j^1 w_{i,j} \gamma_T(R)}{\sum_i^s \sum_j^1 w_{i,j}}, \quad (2.5)$$

with weights

$$w_{i,j} = \langle \Sigma_{\text{crit}}^{-1} \rangle_{i,j}^2 / \sigma_{\gamma, i}^2, \quad (2.6)$$

where  $\sigma_{\gamma, i}^2$  is the variance of the estimated shape for galaxy  $i$ .

The practical or more general form of the reduced tangential shear estimator for a single cluster,  $\langle g_T \rangle(R)$ , which is used in our analysis is defined as

$$\langle g_T \rangle(R) \equiv \frac{\sum_i \omega_i e_{T;i}}{\frac{N_{\text{rb}}}{N_{\text{tot}}} \left[ \sum_i \omega_i R_{\gamma,i}^T + \sum_i \omega_i \langle R_{\text{sel}}^T \rangle \right]}, \quad (2.7)$$

where the denominator is re-scaled according to the number of background objects in a specific radial bin,  $N_{\text{rb}}$ , relative to the total number of background objects for a single cluster,  $N_{\text{tot}}$ .  $R_{\gamma}^T$  is the shear response term measured from the distorted images of the galaxies and  $R_{\text{sel}}^T$  is the selection response term which examines the response of the selections to the shear. The reduced tangential shear is estimated over several radial bins,  $R$ . The weights are

$$\omega_{i,j} \equiv \Sigma_{\text{crit}}^{-1} (z_{l_j}, \langle z_{s_i}^{\text{MCAL}} \rangle) \text{ if } \langle z_{s_i}^{\text{MCAL}} \rangle > z_{l_j} + 0.1, \quad (2.8)$$

with  $\langle z_s^{\text{MCAL}} \rangle$  being the mean redshift of the source galaxy estimated from METACALIBRATION photometry. Rather than making use of the full redshift distribution of the source galaxy we opted for the mean redshift to speed up computation. Given the width of the estimated  $P(z)$ , this simplification is close to the optimal weight and is justified in (e.g. Malz et al., 2018).

#### 2.4.2 Source background distribution

The source background distribution of a specific cluster contains galaxies that have been selected and weighted in several logarithmically spaced radial bins from the cluster. We select the source galaxies such that they have  $z_{\text{mean}} > z_{\text{cl}} + 0.1$ , where  $z_{\text{mean}}$  is the mean redshift of each background galaxy and  $z_{\text{cl}}$  is the cluster redshift. We choose the  $\Delta z = 0.1$  offset to produce a good quality background sample with low cluster contamination, given

typical photo-z uncertainties. Given this sample, we construct the estimator of the true redshift distribution  $P(z)$  using the redshift estimator  $z_{\text{MC}}$ , which is a random sample of the redshift probability distribution of each source galaxy. The selected background galaxies are then weighted by  $\omega_i R_{\gamma;i}^T$  where  $\omega_i$  is described in Equation 2.8 and  $R_{\gamma;i}^T$  is the shear response term measured from the distorted images of the galaxies. Figure 2.4 shows an example of the source background distribution of a cluster at redshift  $z_{\text{cl}} = 0.47$ . In the bottom panel, the photo-z selected sample with  $z_{\text{mean}} > z_{\text{cl}} + 0.1$  is shown, and in the top panel the estimator of the true redshift distribution  $P(z)$  for the same galaxies is shown. One can see the excess of galaxies appearing near the redshift of the cluster. These are the contaminating galaxies in our background shear sample.

### 2.4.3 Cluster contamination and correction

We follow the method first described in Gruen et al. (2014), used in McClintock et al. (2019) and further analysed in Varga et al. (2019). A model is assumed that decomposes the measured source background distribution of each cluster into two parts: a cluster member component located near the cluster redshift and a field component that is constrained using a weighted distribution of local field galaxies:

$$\begin{aligned} \widetilde{\gamma}_T &= \frac{\sum_{j,i \in \text{cl}} w_{i,j} \gamma_T + \sum_{j,i \in \text{field}} w_{i,j} \gamma_T}{\sum w_{i,j}} \\ &= \left( \underbrace{\frac{\sum_{j,i \in \text{cl}} w_{i,j}}{\sum w_{i,j}}}_0 \right) \langle \gamma_T \rangle_{\text{cl}} + \left( \underbrace{\frac{\sum_{j,i \in \text{field}} w_{i,j}}{\sum w_{i,j}}}_{\langle \gamma_T \rangle_{\text{true}}} \right) \langle \gamma_T \rangle_{\text{field}} \end{aligned} \quad (2.9)$$

$\langle \gamma_T \rangle_{\text{cl}}$  carries no signal, while  $\langle \gamma_T \rangle_{\text{field}}$  describes the “true” signal we would measure if there were no contamination.  $\sum_{j,i \in \text{cl}}$  and  $\sum_{j,i \in \text{field}}$  denotes a sum over source-lens pairs for cluster members and background galaxies, respectively. The contamination fraction  $f_{\text{cl}}$  is then defined as follows:

$$f_{\text{cl}}(R) = \frac{\sum_{j,i \in \text{cl}} w_{i,j}}{\sum_{j,i} w_{i,j}} , \quad (2.10)$$

where the summation is over all sources within a radial bin. The contamination fraction used in our analysis can be derived from the general form of the shear estimator. The selection response term will be ignored here due to it being sub-dominant with respect to the shear response  $R_{\gamma,i}^T$  (McClintock et al., 2019). Hence the shear estimator for a single cluster takes the following form:

$$\langle g_T \rangle(R) \equiv \frac{\sum_i \omega_i e_{T;i}}{\frac{N_{\text{rb}}}{N_{\text{tot}}} \left[ \sum_i \omega_i R_{\gamma,i}^T \right]} , \quad (2.11)$$

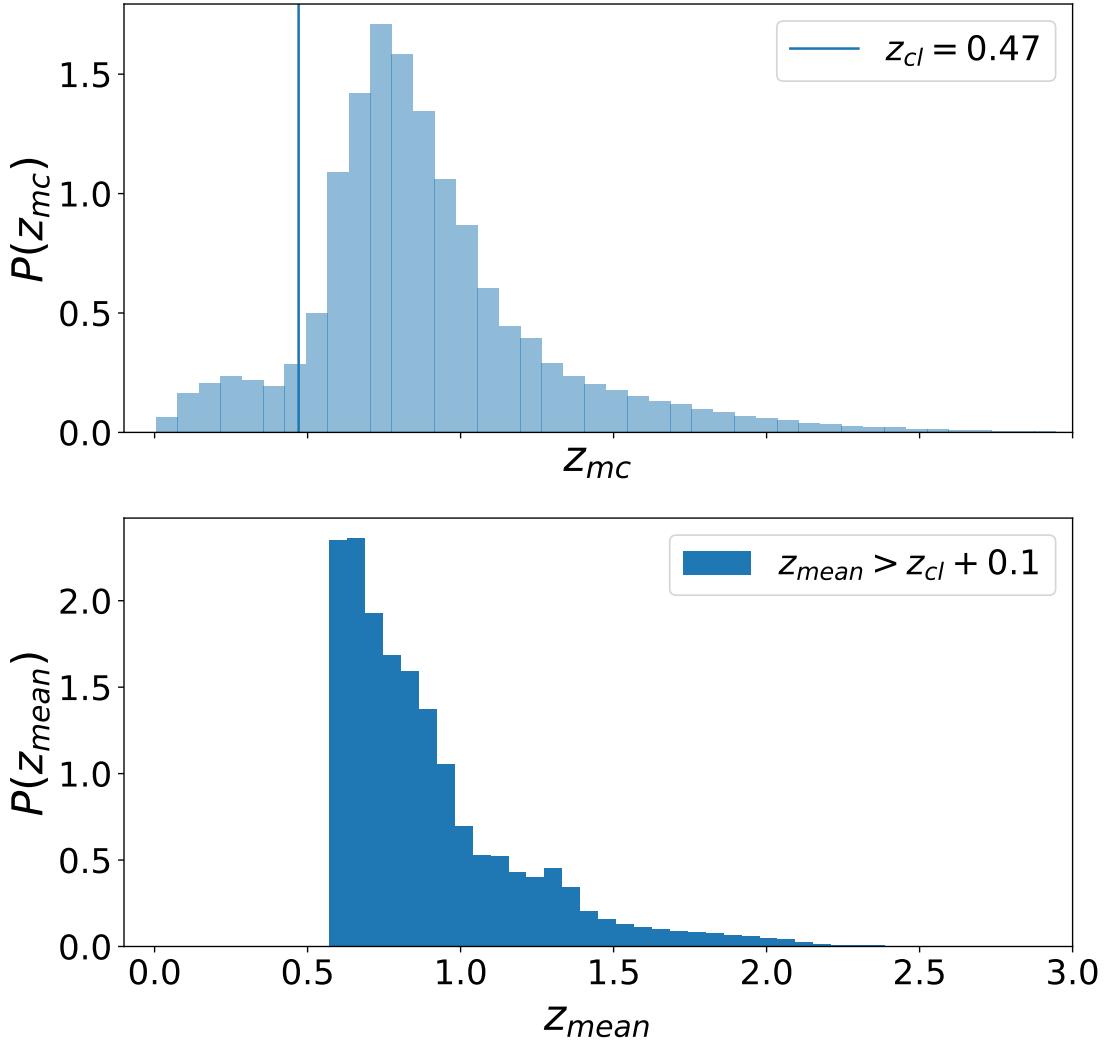


Figure 2.4: The selection criterion of the source galaxies is such that  $z_{\text{mean}} > z_{\text{cl}} + 0.1$ , where  $z_{\text{mean}}$  is the mean redshift of each background galaxy and  $z_{\text{cl}}$  is the cluster redshift. The  $P(z)$  is then created using  $z_{\text{MC}}$ , a random sample of the redshift distribution of each source galaxy. The selected background galaxies are weighted by  $\omega_i R_{\gamma; i}^T$  where  $\omega_i$  is described in Equation 2.8 and  $R_{\gamma; i}^T$  is the shear response term. Here we show this with the example of a cluster at redshift  $z_{\text{cl}} = 0.47$

Following Equations 2.9 and 2.10 the contamination fraction is

$$f_{\text{cl}}(R) = \frac{\sum_{\text{cl}} \omega_i R_{\gamma;i}^T}{\sum \omega_i R_{\gamma;i}^T}. \quad (2.12)$$

which can be used to correct the predicted shear profile to recover the true signal

$$g_{\text{true}}(R) = \frac{g(R)}{1 - f_{\text{cl}}(R)}. \quad (2.13)$$

#### 2.4.4 P(z) decomposition method

The background sample in our analysis is studied within 9 logarithmically spaced radial bins from the cluster centre ranging from 0.25 Mpc to 10.4 Mpc. The lower limit was set to avoid the dense cluster core where blending is more prominent (Medezinski et al., 2018a) and where there are few background galaxies available in the shear catalogues (especially in higher redshift clusters,  $z_{\text{cl}} > 0.6$ ). The upper limit of 10.4 Mpc is larger than the upper limit we are going to use for the WL mass calibration, and smaller than the radius for which you would be worried that the field distribution starts to be severely affected by sample variance. The last radial bin was chosen to be far enough from the cluster to be useful to constrain the local field distribution of each cluster. Due to field depth variations in the DES Y1 survey dataset, one cannot use the survey average. Otherwise, field depth variations in the METACALIBRATION catalogue from DES Y1 will introduce additional noise and a bias in the estimate of the cluster contamination.

A visual representation of source galaxies (in orange) behind a particular cluster is shown in Figure 2.5. Small masked areas are apparent, and the edges of the radial bins are marked with circles of alternating colour. As mentioned in Sections 2.3.2 and 2.3.3, clusters with very large masked areas or clusters that are positioned near the edge of the DES Y1 footprint are taken out of the sample. This was performed via visual inspection. The cluster presented here has a redshift of  $z_{\text{cl}} = 0.17$ .

As described in Section 2.4.3 we model the background galaxy redshift distribution as a combination of two terms; a cluster member component and a local field component of appropriately selected and weighted galaxies. We further assume that the redshift distribution of cluster members can be modelled as a Gaussian as it was shown in Varga et al. (2019) (section 3.2.5) that a Gaussian distribution is a reasonably good representation of the contaminating cluster galaxies. Hence, the redshift distribution takes the following form:

$$\begin{aligned} P_{\text{pred}}(z, R) &= f_{\text{cl}}(R)P_{\text{mem}}(z) + (1 - f_{\text{cl}}(R))P_{\text{field}}(z) \\ &= f_{\text{cl}}(R)\mathcal{N}(z, \mu, \sigma) + (1 - f_{\text{cl}}(R))P_{\text{field}}(z), \end{aligned} \quad (2.14)$$

where  $z$  is the redshift of the background object,  $R$  is the area-weighted average radius of the radial bin in Mpc,  $f_{\text{cl}}(R)$  is the contamination fraction in the bin of radius  $R$ ,

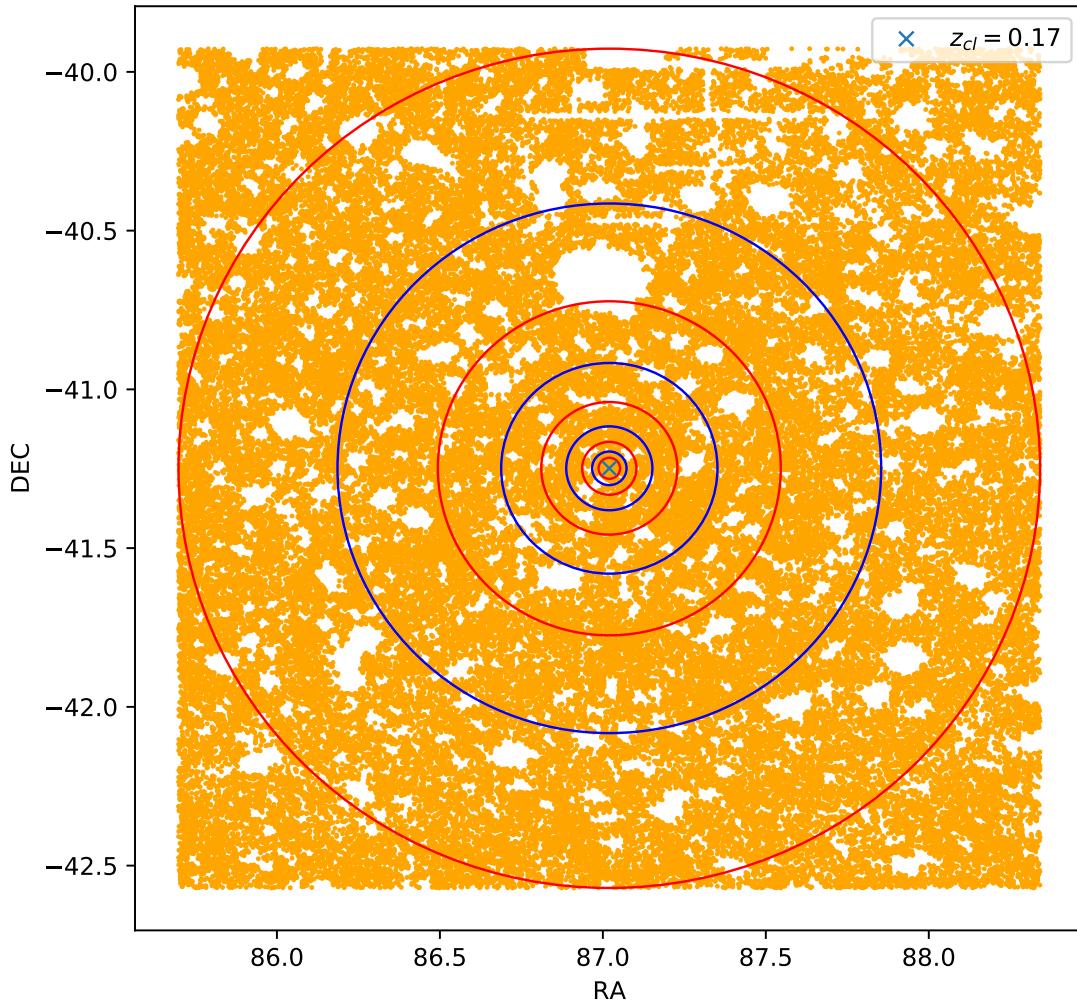


Figure 2.5: The contamination in the background galaxy sample of each cluster is studied in 9 logarithmic annuli with metric radii ranging from 0.25 Mpc to 10.4 Mpc. The last two radial bins are used to represent the uncontaminated background galaxy sample around the cluster. The source galaxies are shown in orange. Small masked areas are apparent. As mentioned in Sections 2.3.2 and 2.3.3, clusters with very large masked areas or that are positioned at the edge of the DES Y1 footprint are taken out of the sample. The cluster presented here has a redshift of  $z_{cl} = 0.17$ .

$\mathcal{N}(z, \mu, \sigma)$  describes our model of the distribution of cluster members to be Gaussian and finally,  $P_{\text{field}}(z)$  is the uncontaminated background galaxy redshift distribution.

The Gaussian distribution is described in the following way:

$$\begin{aligned} \mathcal{N}(z; \mu, \sigma^2) &= \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2}\frac{(z-\mu)^2}{\sigma^2}\right) \\ \mu &= z_{\text{cl}} + \mu_0 \quad , \quad \sigma = \sigma_0 \end{aligned} \quad (2.15)$$

Here,  $\mu_0$  describes the offset from the cluster redshift,  $z_{\text{cl}}$ . Both,  $\mu_0$  and  $\sigma_0$  are free parameters in our analysis.

Our *ansatz* for the contamination fraction,  $f_{\text{cl}}(R, \lambda, z_{\text{cl}})$ , is defined in the following way:

$$(1 - f_{\text{cl}}(R, \lambda, z_{\text{cl}}))^{-1} = 1 + A_i(z_{\text{cl}}) \left(\frac{\lambda}{\lambda_{\text{piv}}}\right)^B \Sigma(R, r_s), \quad (2.16)$$

where  $A_i(z_{\text{cl}})$  describe the set of discrete cluster redshift dependent normalisation parameters for specific cluster redshift bins and  $i$  refers to the assumed redshift bins.  $B$  describes the power-law dependence of the contamination fraction on cluster richness and  $\Sigma(R, r_s)$  characterises the projected Navarro-Frenk-White (NFW) profile (Navarro et al., 1996) with scale radius  $r_s$ . The scale radius is given by  $r_s = r_{200}/c$ , where  $r_{200}$  is defined as the radius within which the mass density of the halo is 200 times the critical density of the Universe and  $c$  is the concentration. The values of  $r_{200}$  for each cluster are inferred from their masses,  $M_{500}$ .  $M_{500}$  was determined using an observable-mass relation. For the SPT clusters the used scaling relation is described in Bocquet et al. (2019) and for the MARD-Y3 clusters the applied scaling relation is described in Klein et al. (2019). The concentration,  $c$ , was set to be a free parameter. The projected NFW profile was normalised to unity at  $\Sigma(R/r_s = 1.01)$ . For the cluster richness pivots,  $\lambda_{\text{piv}}$ , the median values of the cluster samples were assumed. For the SPT cluster sample  $\lambda_{\text{piv}} = 74$ , for MARD-Y3  $\lambda_{\text{piv}} = 66$  and for the combined SPT + MARD-Y3 sample  $\lambda_{\text{piv}} = 65$ . As mentioned in subsection 2.3.4, due to the small scatter between the MCMF richesses on SPT and MARD-Y3 and the on average lower assigned richness of SPT clusters compared to their MARD-Y3 counterpart, the median richness of the combined sample is lower than the medians of the individual samples.

As mentioned above, this method of estimating the contamination fraction involves its estimation for different redshift bins for a discrete set of cluster redshift dependent normalisation parameters, but this means that the correction we apply depends on the redshift bin limits. However, the true cluster contamination is almost certainly a smooth function and henceforth should be modelled as such. Our final model allows for the discrete amplitude fit parameters but with a Gaussian smoothing kernel applied. This method is inspired by radial basis function interpolation with the distinction that we centre the basis functions at a discrete set of redshifts spread over the full range of interest. This approach represents a convenient way to obtain a smooth prediction over full range of cluster redshifts

we consider. It takes the following form:

$$A(z_{\text{cl}}) = \exp [A_{\text{asymptote}} + \sum_i e^{-\left(\frac{z_{\text{cl}} - z_i}{\sigma_{\text{corr}}}\right)^2} A_i] \quad (2.17)$$

where  $A(z_{\text{cl}})$  describes the amplitude at an arbitrary cluster redshift. The sum runs over all amplitude fit parameters  $A_i$  and all redshifts  $z_i$ . The Gaussian weight factor leads to “nearby” (in redshift) amplitudes contributing more strongly than the  $A_i$  that are far away. The variance  $\sigma_{\text{corr}}^2$  in the Gaussian weight is an additional parameter of the model that acts as a correlation length.

There is not enough information in our samples to constrain the correlation length well, and so through comparison of the interpolation function with the originally binned results we select the form  $\sigma_{\text{corr}} = 0.1(1 + z_{\text{cl}})^2$ . A fixed correlation length of 0.1 works well at low cluster redshifts but allows for too much variation at high cluster redshifts (above  $z_{\text{cl}} \approx 0.7$ ). Increasing the correlation length with redshift, either as  $0.1(1 + z_{\text{cl}})$  or  $0.1(1 + z_{\text{cl}})^2$ , improves the consistency of the interpolation with the binned results, and the latter form with steeper redshift increases provides the better description of the binned results.

The  $A_{\text{asymptote}}$  parameter sets the value  $A(z_{\text{cl}})$  for redshifts that are well away from the set of redshifts  $z_i$ . We set this parameter to  $-2.21$ . The exponential form for the function is chosen to ensure that the contamination will always be positive. Finally, with the contamination amplitude for the multi-Gaussian method determined as above, we estimate the contamination fraction at a given cluster redshift  $z_{\text{cl}}$ , richness  $\lambda$  and projected radius from the centre  $R$  in the following way:

$$(1 - f_{\text{cl}}(R, \lambda, z_{\text{cl}}))^{-1} = 1 + A(z_{\text{cl}}) \left(\frac{\lambda}{\lambda_{\text{piv}}}\right)^B \Sigma(R, r_s), \quad (2.18)$$

The adopted form for  $f_{\text{cl}}$  with  $A(z_{\text{cl}}) \geq 0$  and arbitrary  $B$  ensures the range  $0 < f_{\text{cl}} < 1$  (see, e.g., Varga et al., 2019). All parameters, except for the amplitude are as described above. With nine (eight) amplitude parameters  $A_i$  at different redshifts, our multi-Gaussian model then has 13 (12) free parameters for SPT (MARD-Y3).

#### 2.4.5 Likelihood

The likelihood we adopt in solving for the best fit parameters is

$$\ln L_{\text{Poisson}} = \sum_{z,R} [n_{\text{meas}}(z, R) \times \ln(n_{\text{pred}}(z, R)) - n_{\text{pred}}(z, R)] \quad (2.19)$$

where

$$n_{\text{meas}}(z, R) = N(R) \Delta z P_{\text{meas}}(z, R) \quad (2.20)$$

$$n_{\text{pred}}(z, R) = N(R) \Delta z P_{\text{pred}}(z, R). \quad (2.21)$$

Here,  $N(R)$  is the total number of background galaxies in each radial bin and  $\Delta z$  is the width of the redshift bins.  $P_{\text{meas}}(z, R)$  and  $P_{\text{pred}}(z, R)$  are normalised to unity and are the measured and predicted redshift distributions, respectively.

The cluster log-likelihoods are summed to obtain the total log-likelihood, and the best fit parameters are all determined in a simultaneous fit

$$\ln L = \sum_i \ln L_{\text{Poisson}}[i]. \quad (2.22)$$

As discussed later, we examine additional redshift freedom in the other model parameters and find no compelling evidence that an extension of our model is needed.

#### 2.4.6 Validation with mock catalogues

To validate our fitting code, we create a mock catalogue by first choosing certain input parameters similar to those recovered from the data and then drawing Poisson realisations of the contaminated redshift distributions for each cluster in our sample. The mock has a 10 times larger sample size than our cluster sample so that we can probe for any biases at a level corresponding to a  $\sim \sqrt{10}$  smaller than the statistical uncertainties that result from our real data. We then analyse those mock catalogues, recovering the input parameters with good accuracy. A figure showing the results of the combined SPT + MARD-Y3 mock cluster sample is presented in the appendix (see Figure A.1).

## 2.5 Cluster contamination measurements

By allowing multiple free normalisation parameters distributed over the redshift range of interest, one can model the redshift trends of the cluster contamination. We present results for an analysis using the multi-Gaussian interpolation scheme (equations 2.17 and 2.18) but we also carry out a redshift bin based analysis (equation 2.16) for comparison.

For the binned analysis, we adopt the following redshift bins:  $z_{\text{cl}} \in \{0.15, 0.2, 0.25, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.2\}$  for SPT and SPT + MARD-Y3 and  $z_{\text{cl}} \in \{0.15, 0.2, 0.25, 0.3, 0.4, 0.5, 0.6, 0.7, 0.9\}$  for MARD-Y3, where always the lower limit is included and the upper limit is not. An even finer binning, especially at higher redshifts, was found to result in a too few clusters in a cluster redshift bin to still successfully constrain the model parameters.

For the multi-Gaussian method we adopt the following redshift arrays for the discrete amplitude fit parameters:  $z_i \in \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.9, 1.1\}$  for SPT and SPT+MARDY3 and  $z_i \in \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.9\}$  for MARDY3. The redshift spacing was doubled at high redshifts because the steep increase in correlation length with redshift requires less dense redshift sampling. Shifting the assumed redshift arrays by  $\delta z = 0.05$  did not lead to a statistically significant change in the estimated  $A(z_{\text{cl}})$  values. If not stated otherwise, all results presented in this paper are shown for the multi-Gaussian method.

Figure 2.6 shows the contour plots of the best-fit parameters of the SPT cluster sample in blue and of the MARD-Y3 cluster sample in green where we assume different normalisations for the different cluster redshift bins. Because the SPT and MARD-Y3 results are comparable for most parameters (there is some tension in  $\mu$  that we discuss below), we also present the combined results as black curves. The combined SPT + MARD-Y3 sample contains 397 clusters where multiples between the two samples have been removed. We note that there is a small degeneracy between most of the various normalisation parameters,  $A_i(z_{\text{cl}})$  and the concentration parameter  $c$ . A similar degeneracy was also found by McClintock et al. (2019). They conclude that the impact of the degeneracy was minimal on the accuracy of the estimated contamination profile. We do not find any degeneracy between the individual normalisation parameters,  $A_i(z_{\text{cl}})$ . A negligible difference was found comparing the resulting best fit parameters when using the optical versus the SZ centres for the SPT cluster sample.

For comparison, Figure 2.7 shows the contour plots of the best-fit parameters of the SPT cluster sample in blue and of the MARD-Y3 cluster sample in green for the multi-Gaussian method. Again we present the combined results as black curves.

The parameter posterior distributions are sampled using the package emcee (Foreman-Mackey et al., 2013), which enables a parallelised exploration of the parameter space. We use 156 (144) walkers for SPT (MARD-Y3) with 1000 steps each, discarding the first 40,000 samples.

In Table 2.1 we compare the results of the best fit parameters of our different data sets – SPT, MARD-Y3 and SPT + MARD-Y3 – when using the binning method, whereas Table 2.2 compares the results of the best-fit parameters of our different data sets when using the multi-Gaussian interpolation method. As is clear from the table, there is tension between the MARD-Y3 and SPT datasets in the parameters  $\mu_0$  and  $B$ , but the other parameters are in reasonably good agreement. We quantify these tensions by doing a “probability of consistency check” and find a tension of  $2.8\sigma$  for  $\mu_0$  and a tension of  $1.3\sigma$  for  $B$  (see discussion in Section 2.5.2). These numbers are fully consistent with the best fit values and uncertainties presented in Table 2.2.

The redshift normalisation,  $A(z_{\text{cl}})$ , for SPT, MARD-Y3 as well as SPT + MARD-Y3 is plotted in Figure 2.8. The resulting normalisation parameters from the binned method (shown as points with error bars) are in good agreement with the multi-Gaussian method (shown as line with  $1\sigma$  uncertainty band). It is clear that a simple power law would not have been sufficient to describe the redshift variation of the cluster contamination for these cluster samples. The redshift behaviour of the cluster contamination normalisation is reasonably consistent in the two samples and therefore is reflective of intrinsic properties of the DES Y1 dataset such as, e.g., the photo-z behaviour.

### 2.5.1 Validation of the cluster contamination results

To validate these results we adopt a way of estimating contamination fraction within stacks of clusters at similar redshift or richness. This allows us to, for example, test whether the model we fit above is adequate to describe the data. We estimate the contamination

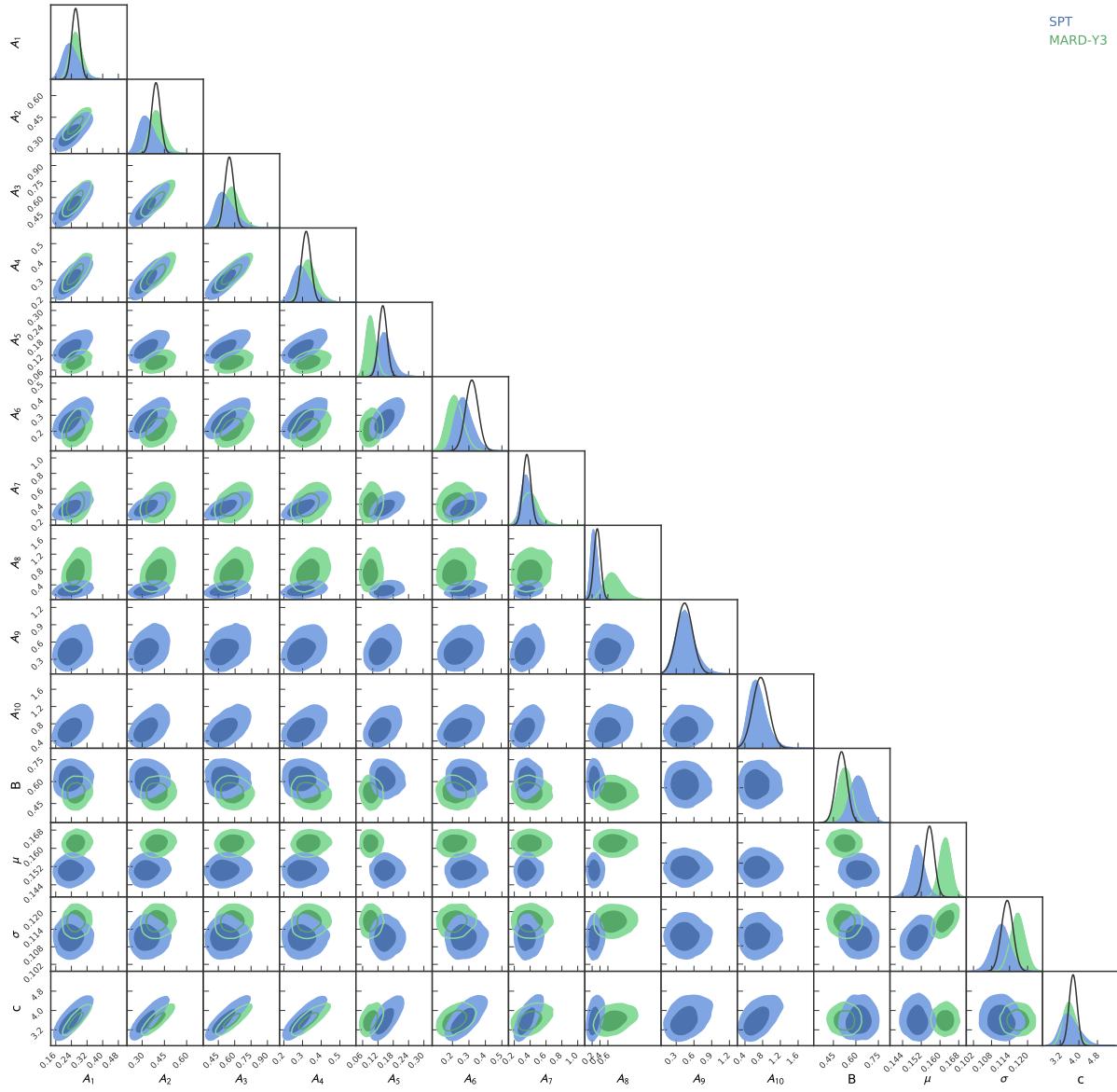


Figure 2.6: Constraints on the contamination fraction model  $f_{\text{cl}}$  for the SPT (blue), MARD-Y3 (green) and combined SPT + MARD-Y3 samples (black line) where we assume discrete amplitude fit parameters  $A_i$  for different redshift bins. We find a modest degeneracy between the individual normalisation parameters  $A_i(z_{\text{cl}})$  and concentration  $c$ , which is expected given the concentration dependence of the normalisation radius  $1.01r_s$ .

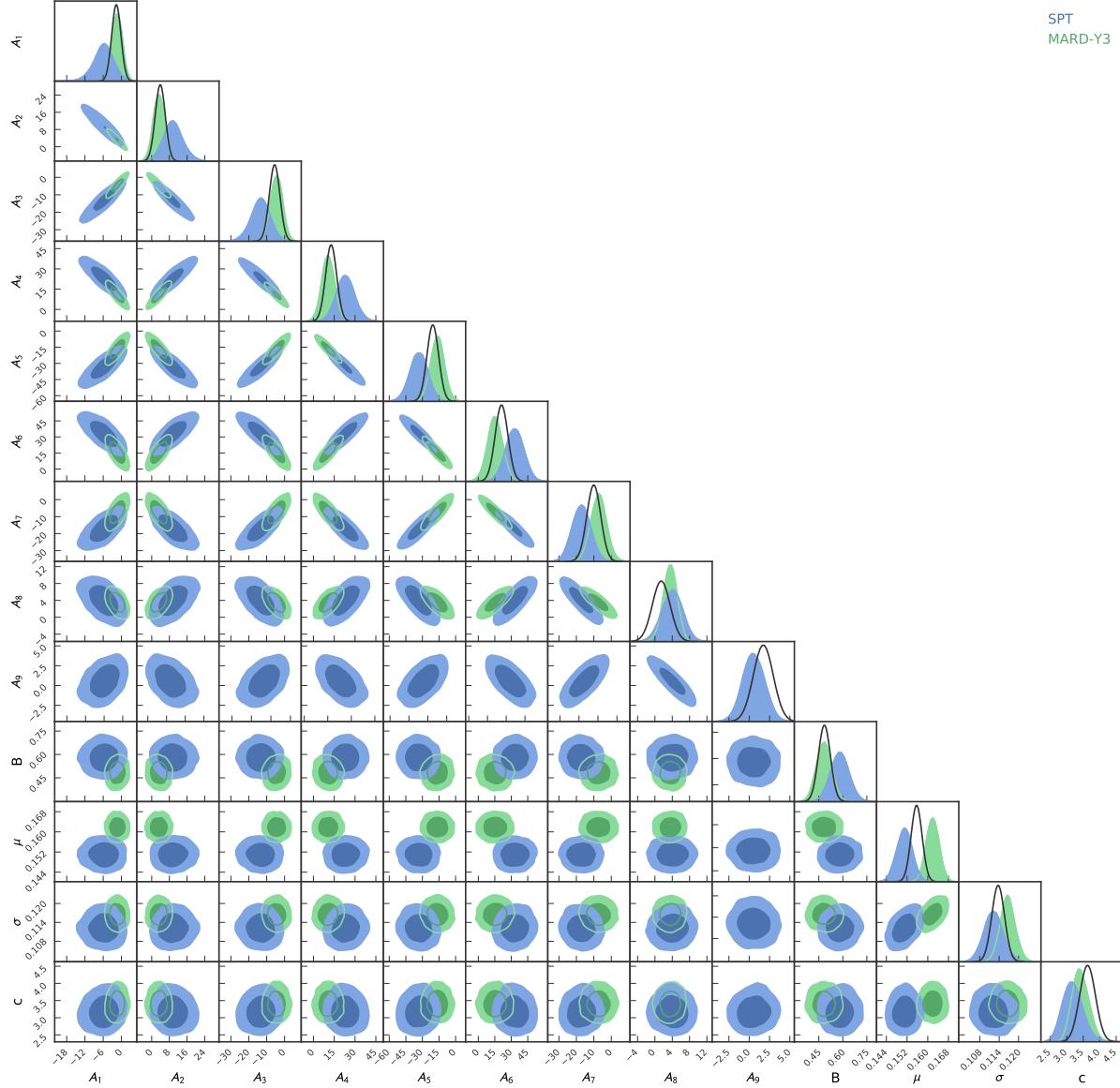


Figure 2.7: Constraints on the contamination fraction model  $f_{\text{cl}}$  for the SPT (blue), MARD-Y3 (green) and combined SPT + MARD-Y3 samples (black line) where we assume discrete amplitude fit parameters  $A_i$  for a given redshift array and use a multi-Gaussian interpolation method. Note the strong anti-correlation among neighbouring normalisation parameters of the multi-Gaussian function. This could be reduced by adopting fewer amplitude parameters, but we use this function only for interpolation purposes (see Fig. 2.8) where the parameter degeneracies are accounted for.

Table 2.1: Binned results: Best fit parameters for SPT, MARD-Y3 and SPT+MARD-Y3 of Equations 2.15 & 2.16. For the SPT and SPT + MARD-Y3 cluster samples the normalisation parameters  $A_i(z_{\text{cl}})$  are fitted in the following cluster redshift ranges:  $A_1 - A_{10}$ :  $z_{\text{cl}} \in \{0.15, 0.2, 0.25, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.2\}$ . For the MARD-Y3 cluster sample we have  $A_i(z_{\text{cl}})$  being fitted in the ranges:  $A_1 - A_8$ :  $z_{\text{cl}} \in \{0.15, 0.2, 0.25, 0.3, 0.4, 0.5, 0.6, 0.7, 0.9\}$ . In both cases the lower limit is included whereas the upper limit is not.

Parameters	SPT	MARD-Y3	SPT+MARD-Y3
$A_1$	$0.241 \pm 0.041$	$0.265 \pm 0.030$	$0.262 \pm 0.020$
$A_2$	$0.337 \pm 0.055$	$0.402 \pm 0.046$	$0.393 \pm 0.029$
$A_3$	$0.515 \pm 0.085$	$0.583 \pm 0.070$	$0.551 \pm 0.041$
$A_4$	$0.301 \pm 0.049$	$0.337 \pm 0.042$	$0.320 \pm 0.025$
$A_5$	$0.151 \pm 0.027$	$0.093 \pm 0.018$	$0.140 \pm 0.016$
$A_6$	$0.273 \pm 0.051$	$0.217 \pm 0.048$	$0.318 \pm 0.038$
$A_7$	$0.365 \pm 0.072$	$0.403 \pm 0.100$	$0.358 \pm 0.050$
$A_8$	$0.262 \pm 0.089$	$0.746 \pm 0.222$	$0.330 \pm 0.085$
$A_9$	$0.463 \pm 0.157$		$0.442 \pm 0.134$
$A_{10}$	$0.692 \pm 0.195$		$0.756 \pm 0.181$
$B$	$0.619 \pm 0.053$	$0.525 \pm 0.045$	$0.504 \pm 0.037$
$\mu_0$	$0.150 \pm 0.0027$	$0.162 \pm 0.0024$	$0.155 \pm 0.0021$
$\sigma_0$	$0.111 \pm 0.0029$	$0.117 \pm 0.0023$	$0.113 \pm 0.0019$
$c$	$3.631 \pm 0.385$	$3.608 \pm 0.272$	$3.751 \pm 0.171$

Table 2.2: Gaussianized interpolation results: We present the best-fit parameters for SPT, MARD-Y3 and SPT+MARD-Y3 of Equations 2.15, 2.17 & 2.18 when using the multi-Gaussian method. For the SPT and SPT + MARD-Y3 cluster samples the normalisation parameters  $A_i$  are being fitted for the following cluster redshift array:  $A_1 - A_9: z_{\text{cl}} \in \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.9, 1.1\}$ . For the MARD-Y3 cluster sample we have  $A_i$  being fitted in the ranges:  $A_1 - A_8: z_{\text{cl}} \in \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.9\}$ . The  $A_{\text{asymptote}}$  parameter was set to  $-2.21$ .

Parameters	SPT	MARD-Y3	SPT+MARD-Y3
$A_1$	$-5.992 \pm 3.227$	$-1.377 \pm 1.683$	$-1.687 \pm 1.516$
$A_2$	$9.776 \pm 4.438$	$3.266 \pm 2.518$	$3.876 \pm 2.257$
$A_3$	$-13.421 \pm 5.340$	$-4.506 \pm 3.323$	$-5.575 \pm 2.951$
$A_4$	$22.969 \pm 6.593$	$10.610 \pm 4.535$	$12.755 \pm 3.983$
$A_5$	$-32.945 \pm 7.969$	$-16.414 \pm 6.058$	$-20.617 \pm 5.263$
$A_6$	$32.731 \pm 8.070$	$15.180 \pm 6.702$	$21.142 \pm 5.811$
$A_7$	$-16.793 \pm 5.087$	$-7.245 \pm 4.332$	$-10.030 \pm 3.956$
$A_8$	$3.845 \pm 2.278$	$3.477 \pm 1.467$	$1.448 \pm 2.048$
$A_9$	$0.623 \pm 1.270$		$1.729 \pm 1.226$
$B$	$0.581 \pm 0.053$	$0.482 \pm 0.046$	$0.485 \pm 0.037$
$\mu_0$	$0.151 \pm 0.0029$	$0.162 \pm 0.0024$	$0.156 \pm 0.0021$
$\sigma_0$	$0.112 \pm 0.0029$	$0.117 \pm 0.0022$	$0.114 \pm 0.0020$
$c$	$3.186 \pm 0.300$	$3.400 \pm 0.233$	$2.631 \pm 0.240$

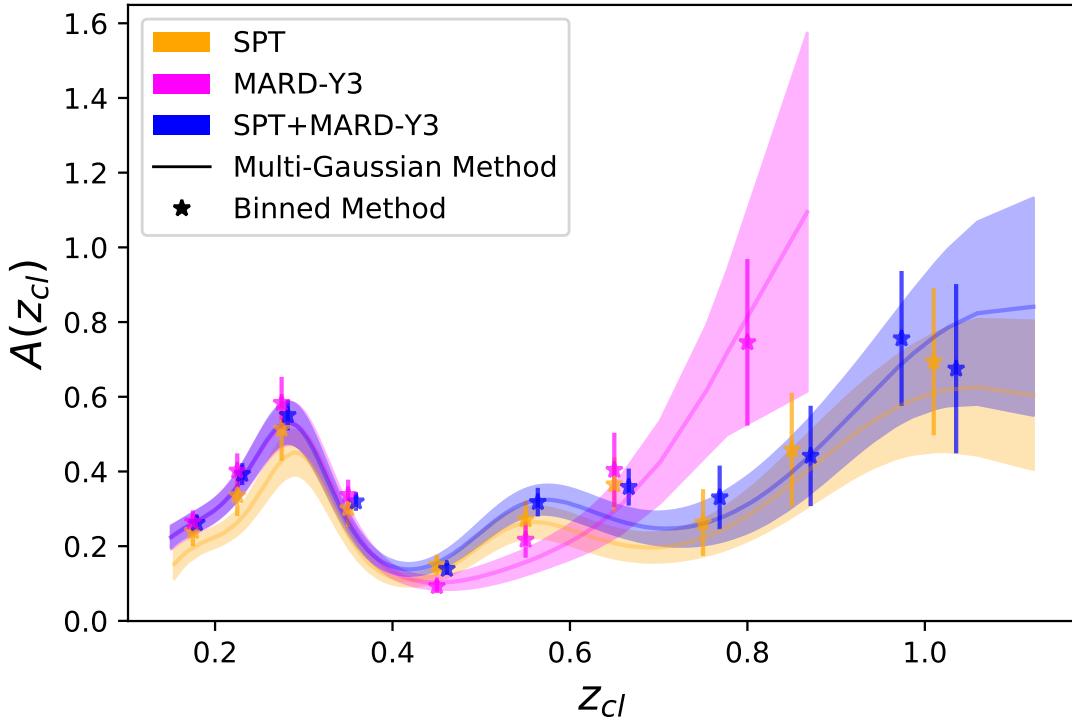


Figure 2.8: The plot shows the different normalisation parameters of the cluster contamination versus cluster redshift for SPT (orange), MARD-Y3 (magenta) and SPT + MARD-Y3 (blue). We compare the best fit normalisation parameters resulting from the binned method,  $A_i(z_{cl})$ , indicated by the stars, to the resulting normalisation parameters for each cluster,  $A(z_{cl})$ , from the multi-Gaussian method. We find them to be in good agreement. The redshift variation shows a rather complex behaviour, and the MARD-Y3 and SPT samples are reasonably consistent.

fraction by subtracting a re-normalised local field distribution,  $P_{\text{field}}(z)$ , from the observed background distribution of the cluster,  $P_{\text{meas}}(z, R)$ . The renormalisation uses the ratio between the respective high redshift tails (e.g.,  $z > z_{\text{cl}} + 0.5$ ) of the predicted background distributions and the local field distributions, hence providing us with a model independent measurement of the cluster contamination:

$$f_{\text{cl}}(R, \lambda) P_{\text{mem}}(z) = P_{\text{meas}}(z, R) - \frac{P_{\text{meas}}(z > z_{\text{cl}} + 0.5)}{P_{\text{field}}(z > z_{\text{cl}} + 0.5)} P_{\text{field}}(z) \quad (2.23)$$

In other words, we use the high redshift portion of the background galaxy redshift distribution where the cluster contamination is expected to be minimal to determine the correct normalisation of the field distribution so that it can be subtracted away. For a single cluster the signal to noise is too low to see much, but by combining these field-subtracted distributions in the space of  $z - z_{\text{cl}}$  one can clearly see the excess of galaxies near the cluster redshift. Moreover, uncertainties on this stacked redshift distribution can be determined using bootstrap resampling.

These distributions produced within bins of redshift or richness are used both for validation of the model we fit and to extract an average cluster contamination for all the clusters in a particular redshift or richness bin by simply integrating over the excess associated with the cluster contamination.

In Figure 2.9 we show an example redshift distribution stacked in the redshift range  $0.2 < z_{\text{cl}} < 0.25$  in blue along with the best fit model in orange, averaged in the same cluster redshift range. This demonstrates that the adopted model provides a good description of the data. We fit our model for the first 8 radial bins, out to 6.87 Mpc. The last radial bin is used to describe the local field distribution around the cluster.

A complementary set of figures showing the contamination distributions as a function of redshift within eight different redshift bins (Figures A.2 to A.5) and as a function of richness or four different richness bins (Figures A.6 to A.9) appear in section A.2. In each of these cases we present the results for the innermost four radial bins. The complex redshift behaviour is clearly identifiable and a modest increase of the contamination with richness is visible.

The next pair of figures show the cluster contamination fraction  $f_{\text{cl}}$  as a function of radius within different cluster redshift and richness bins, respectively. In Figure 2.10 we plot the model predicted cluster contamination fraction (solid lines with  $1\sigma$  shaded regions), colour coded by the cluster redshift bins over which the predicted  $f_{\text{cl}}$  values of each cluster in these redshift bins have been averaged. Shown as points with error bars are the contaminations  $f_{\text{cl}}$  estimated by subtracting a renormalised field distribution from the full  $P(z)$  as described above. The points are the integral over the excess of galaxies within the  $\pm 3\sigma$  region around the cluster member contamination peak. The results are shown for a subsample of the SPT + MARD-Y3 cluster sample, and show reasonably good agreement between the model and the directly measured contamination within the bin. This shows that the adopted radial model for the cluster contamination is a good description of the data, and it confirms the complex variation of the cluster contamination with redshift.

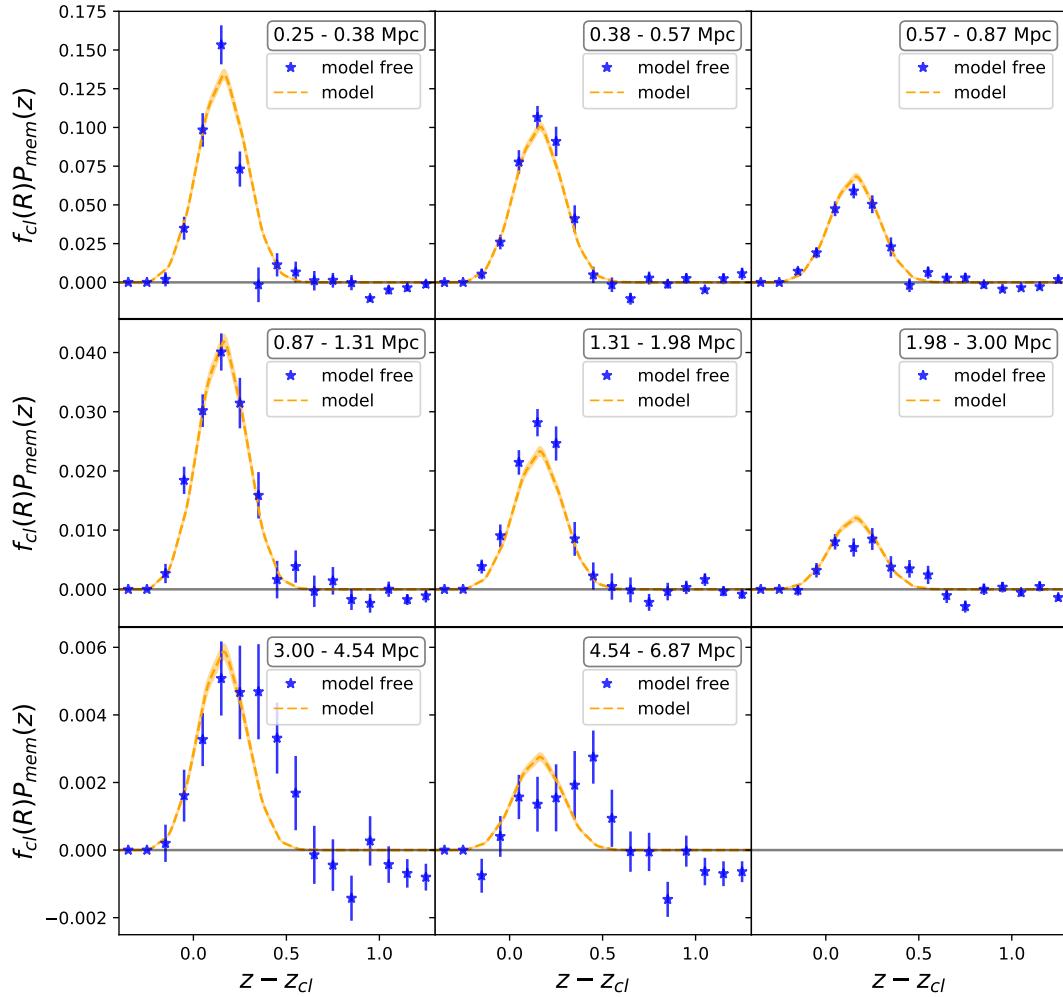


Figure 2.9: An example field-subtracted redshift distribution stacked in  $0.2 < z_{cl} < 0.25$  in blue along with the best fit model in orange, stacked in the same cluster redshift range. The average cluster contamination for all the clusters in this particular redshift bin can be extracted by simply integrating over the excess associated with the cluster contamination, which corresponds to the  $\pm 3\sigma$  region around the cluster member contamination peak. The plot shown here is for the SPT + MARD-Y3 cluster sample.

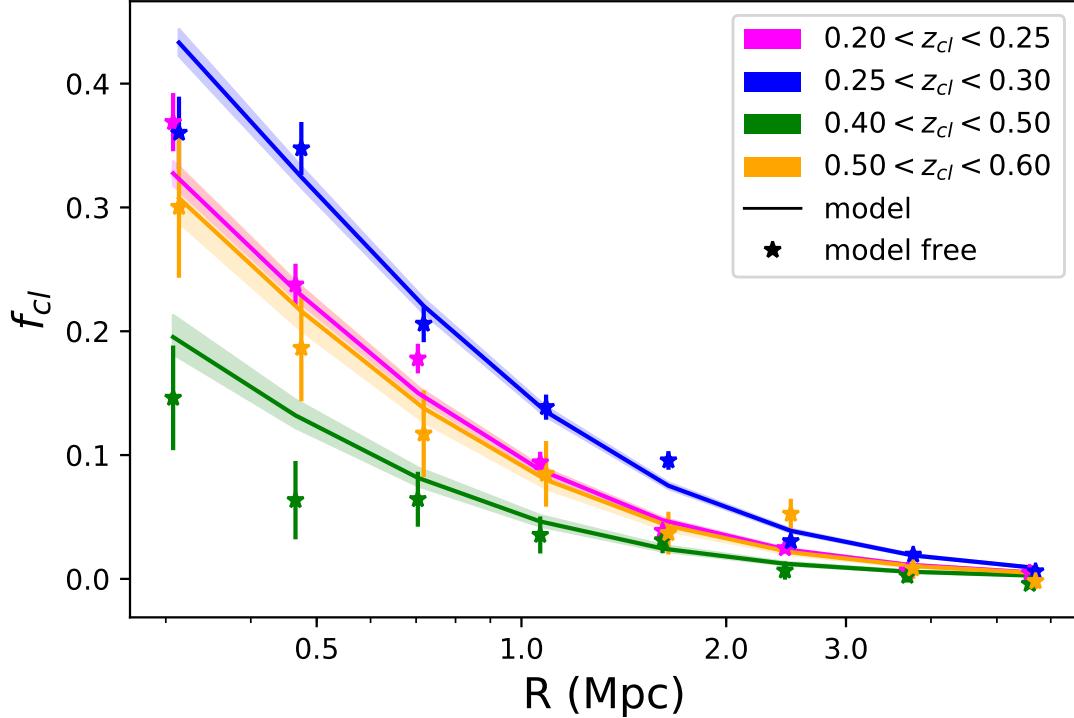


Figure 2.10: The best-fit model for the cluster contamination  $f_{\text{cl}}$  is plotted (solid line with  $1\sigma$  shaded regions) as a function of radius within four different colour-coded redshift ranges over all richnesses. Points with error bars and similar colours represent the cluster contamination extracted from the field subtracted redshift distributions stacked for all clusters within the same redshift ranges (see discussion in Section 2.5.1). The results shown are for the SPT + MARD-Y3 cluster sample. The measurements from the stacked clusters show similar radial and redshift behaviour as the models.

In Figure 2.11 we show a similar plot but within different richness bins. In this case the model predicted  $f_{\text{cl}}$  values of each cluster have been averaged over the chosen richness bins. The agreement between the model and the data in different richness bins indicates that the best fit power law variation of the cluster contamination with richness reproduces the behaviour of the cluster stacks. A modest increase in contamination fraction with richness is identifiable, suggesting that the cluster contamination correction is not heavily dependent on richness.

### 2.5.2 Evaluating the adequacy of the model

This test is for the binned method. We gain constraining power on  $\mu$ ,  $\sigma$ ,  $B$  and  $c$  by adopting a common, fixed value for each parameter at all redshifts. Here we explore

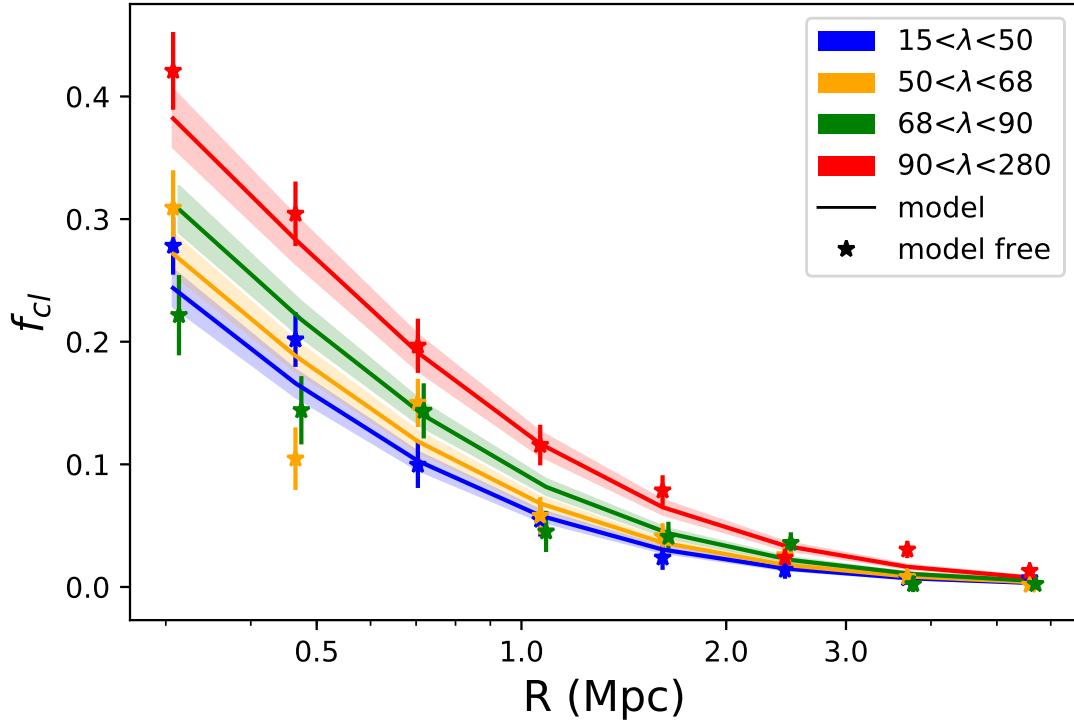


Figure 2.11: The best fit model for the cluster contamination  $f_{cl}$  is plotted (solid line with  $1\sigma$  shaded regions) as a function of radius within four different colour-coded richness ranges. Points with error bars and similar colours represent the cluster contamination extracted from the field subtracted redshift distributions stacked for all clusters within the same richness ranges (see discussion in Section 2.5.1). The results shown are for the SPT + MARD-Y3 cluster sample. The measurements from the stacked clusters show similar radial and richness behaviour as the models.

whether the data are in tension with this approach. Namely, we evaluate the likelihood for each assumed cluster redshift bin individually, allowing not only the normalisation,  $A$ , to be free, but allowing also the parameters  $B$ ,  $\mu$ ,  $\sigma$  and  $c$  to vary within priors that are taken from the resultant best fit values over the full cluster sample. We quantify the tension between the priors and the individual best fit parameters by doing a "probability of consistency check" of the distributions of the priors and the best fit values (Liu et al., 2015; Bocquet et al., 2015).

We begin by calculating the probability density distribution of the difference  $\Delta\theta$  for any pair of the distributions  $P_i(\theta)$ :

$$P(\Delta\theta) = \int d\theta P_1(\theta)P_2(\theta - \Delta\theta). \quad (2.24)$$

Then, we calculate the likelihood  $p$  that the origin ( $\Delta\theta = 0$ ) lies within the distribution:

$$p = \int_S d\Delta\theta P(\Delta\theta), \quad (2.25)$$

where  $S$  is the area where  $P(\Delta\theta) < P(\Delta\theta = 0)$ . The resulting  $p$ -value can then be converted to an equivalent number of sigmas, assuming Gaussian statistics.<sup>1</sup>

Figure 2.12 shows the results of the best fit values for  $B$ ,  $\mu$ ,  $\sigma$  and  $c$  fitted for individual SPT + MARD-Y3 cluster redshift bins rather than over the full sample for set priors with free  $A$ . The black curves mark the Gaussian priors for the individual parameters, which were set to be the best fit values of these parameters over the full cluster sample. Here we show the results for a sub-sample of the chosen cluster redshift bins. None of the parameters aside from the normalisation  $A$  show a statistically significant tension between redshift bins, and this justifies our adoption of single parameters over all redshifts.

### 2.5.3 Redshift variation of cluster contamination

The redshift trend of the cluster contamination amplitude is complex, as shown in Figure 2.8. We have considered the underlying causes of this behaviour, but the situation is quite intricate, with redshift dependent changes in both the cluster and background galaxy samples. For example, as one pushes to higher redshift the cluster galaxies become fainter, and a smaller fraction of these galaxies are sufficiently bright to be included in the shear galaxy sample. The same is valid for the true source galaxies. Moreover, the fraction of cluster-associated galaxies that are star forming and therefore have larger photo-z uncertainties (and therefore presumably contribute more to the cluster contamination sample) increases with redshift (Hennig et al., 2017).

At the same time, due to our cut in photo-z of the shear galaxy sample at  $z_{\text{cl}} + 0.1$  and due to the increasing angular diameter distance, there is an ever smaller number of shear galaxies available within bins of fixed metric radius as one pushes to higher redshift. In addition, there are redshift trends in the performance of the photo-z's. The shifting of

---

<sup>1</sup>The code is available at <https://github.com/SebastianBocquet/PosteriorAgreement>

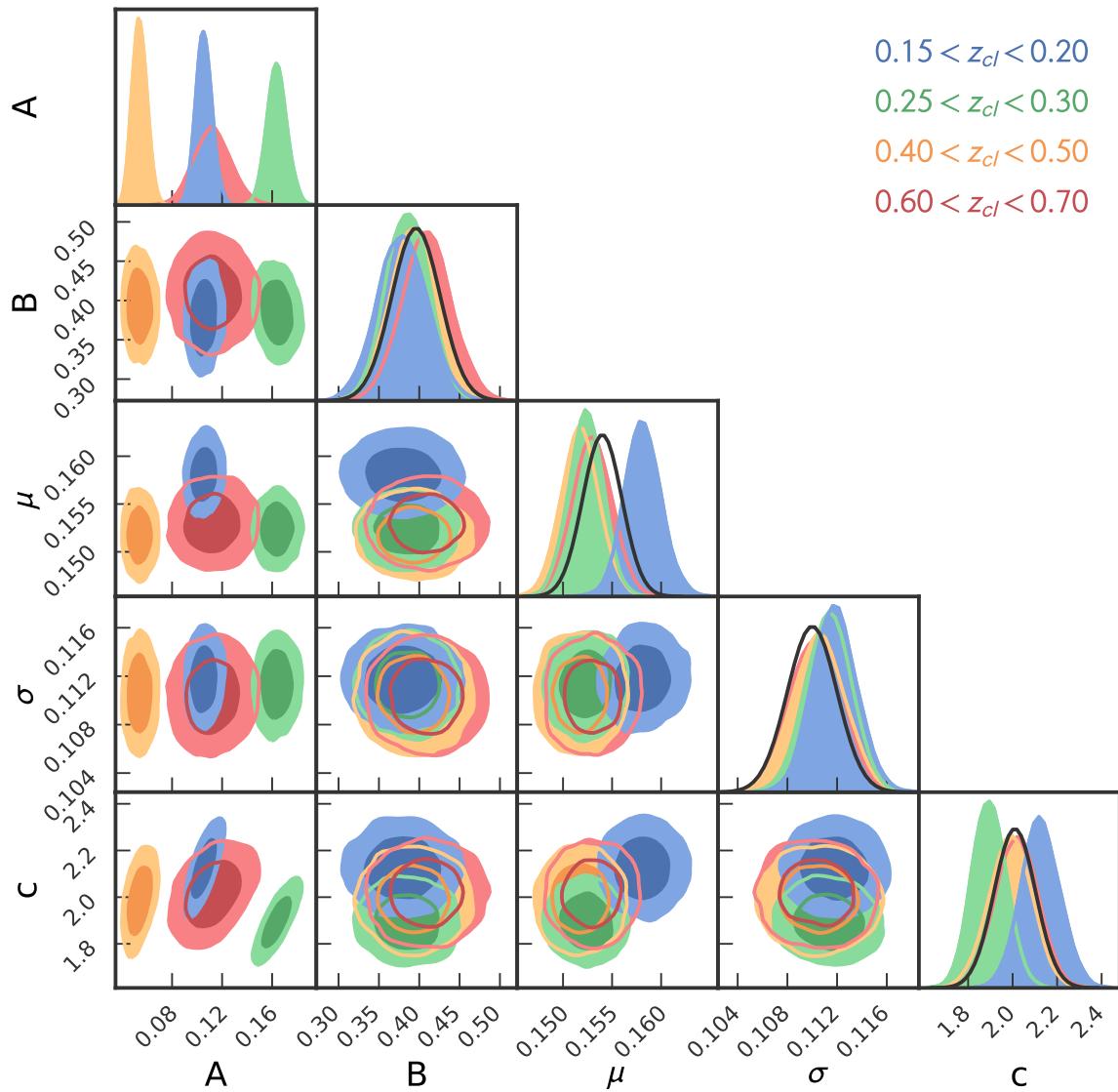


Figure 2.12: Results of the constraints for  $B$ ,  $\mu$ ,  $\sigma$  and  $c$  fitted independently within SPT + MARD-Y3 cluster redshift bins rather than over the full sample. The black curves indicate the assumed Gaussian priors for the individual parameters, which were taken from the best fit values of these parameters over the full cluster sample. Here we show the results for a sub-sample of the chosen cluster redshift bins, but the results are consistent with all bins. No parameter aside from  $A$  shows a statistically significant tension from redshift bin to bin, which allows us to adopt single parameters over all redshifts.

Table 2.3: The weak-lensing biases  $b_{z_i}$ , the natural logarithms of the weak-lensing variances  $s_{z_i}$ , and their mass trends ( $b_M$ ,  $\sigma_M$ ) with their uncertainty induced by the cluster member contamination (see Section 2.5.4). These systematic uncertainties are derived at redshifts at 0.25, 0.47, 0.78, and the mass trend parameters are constant for all redshift. The analysis was repeated twice for two different inner fitting radii  $R_{\min}$ . The cluster member contamination uncertainties for the SPT sample corresponds to, depending on cluster mass, a 1% to 2% systematic floor in mass accuracy and a 2% to 4% systematic floor in mass scatter.

$R_{\min}$	0.29 Mpc	0.71 Mpc
$b_{z_1}$	-0.016 <sup>+0.008</sup> <sub>-0.008</sub>	0.009 <sup>+0.007</sup> <sub>-0.007</sub>
$b_{z_2}$	-0.047 <sup>+0.007</sup> <sub>-0.007</sub>	-0.031 <sup>+0.006</sup> <sub>-0.006</sub>
$b_{z_3}$	-0.087 <sup>+0.015</sup> <sub>-0.013</sub>	-0.090 <sup>+0.010</sup> <sub>-0.010</sub>
$b_M$	0.968 <sup>+0.009</sup> <sub>-0.009</sub>	0.964 <sup>+0.008</sup> <sub>-0.009</sub>
$s_{z_1}$	-3.063 <sup>+0.036</sup> <sub>-0.040</sub>	-3.029 <sup>+0.038</sup> <sub>-0.040</sub>
$s_{z_2}$	-3.049 <sup>+0.032</sup> <sub>-0.035</sub>	-3.058 <sup>+0.031</sup> <sub>-0.034</sub>
$s_{z_3}$	-3.000 <sup>+0.047</sup> <sub>-0.046</sub>	-2.834 <sup>+0.048</sup> <sub>-0.048</sub>
$s_M$	-0.237 <sup>+0.051</sup> <sub>-0.051</sub>	-0.533 <sup>+0.047</sup> <sub>-0.049</sub>

the 4000 Å break or Balmer break of cluster member galaxies from the  $g$  to the  $r$  band at  $z \sim 0.35$  and from  $r$  to  $i$  at  $z \sim 0.75$  impacts the photo-z uncertainties. At these transitions, one might expect the photo-z scatter and outlier fraction to change, leading to rapid changes in the fraction of cluster galaxies that scatter into the background source galaxy population. In Hoyle et al. (2018) (see their figure 2) there is indeed a modest uptick in the characteristic photo-z scatter around  $z \sim 0.3$  where our cluster contamination peaks. On the other hand, there is no indication in our data that the Gaussian width or redshift offset of the contaminant population is changing rapidly over this redshift range (see discussion in Section 2.5.2 above).

At even higher cluster redshifts one would expect again an increase in contamination as the photometric redshifts of the fainter galaxies generally become noisier and a larger fraction would be expected to scatter beyond  $z_{\text{cl}} + 0.1$ . This is broadly consistent with the trend in amplitude of the cluster contamination at  $z > 0.4$ . So while our considerations have identified many contributing factors for explaining the complex redshift trend, we have not yet developed a clear physical explanation for the behaviour we see in both the models and the validation plots. Nonetheless, we are capable of modelling the contamination in a way that accurately represents the data.

#### 2.5.4 Quantifying the impact on cluster mass

In the following we explore the systematic mass uncertainty implied by the parameter uncertainties on the cluster member contamination constraints presented above. We follow the framework presented in Grandis et al. (2021). In that work  $10^3$  realisations of cluster

catalogues with shear profiles extracted from hydrodynamical simulations are produced and weak-lensing masses are measured with the goal of quantifying weak-lensing mass systematics coming from a broad range of effects. In contrast to that work, we focus here on the cluster member contamination as the sole source of systematic uncertainty.

The systematic uncertainty on the weak-lensing mass measurement is captured in the constraints on the parameters of the weak-lensing – halo mass relation. In each realisation of the cluster shear profiles, we draw a different set of cluster member contamination parameters from the posterior presented in Section 2.5 (here we use results from the SPT column in Table 2.2). To define the weak-lensing mass, we choose an NFW model for the reduced shear profile for the DES Y1 source redshift distribution, corrected by the mean cluster member contamination profile. The weak-lensing masses are then measured for each simulated cluster by fitting the model to the simulated shear profile. This gives us the opportunity to vary the inner fitting radius between  $R_{\min} = 0.29$  and  $0.71$  Mpc, and as in Grandis et al. (2021) the outer radius is chosen to avoid the onset of the two halo term  $R_{\max} \sim 4.5$  Mpc.

We create the shear profiles for simulation snapshots at  $z_i \in (0.25, 0.47, 0.78)$ . The weak-lensing – halo mass relation is modeled as a log-normal with mean

$$\left\langle \ln \left( \frac{M_{\text{WL}}}{M_0} \right) \right\rangle = b_{\text{WL}}(z) + b_M \ln \left( \frac{M}{M_0} \right), \quad (2.26)$$

and variance in weak-lensing mass around the relation of

$$\ln \sigma_{\text{WL}}^2 = s(z) + s_M \ln \left( \frac{M}{M_0} \right), \quad (2.27)$$

with a pivot mass of  $M_0 = 2.9 \times 10^{14}$  M<sub>⊙</sub>. For the redshift evolution of the bias  $b_{\text{WL}}(z)$ , we fit an independent bias parameter  $b_i$  for each simulation snap-shot at  $z_i \in (0.25, 0.47, 0.78)$ . For the redshift evolution of the bias  $s(z)$ , we fit an independent bias parameter  $s_i$  for each simulation snap-shot at  $z_i \in (0.25, 0.47, 0.78)$ . The mass trends of the bias  $b_M$  and  $s(M)$  are extracted from the full ensemble at all three redshifts. Each of the  $10^3$  simulated cluster ensembles leads to an estimate of these parameters.

The resulting constraints on the WL – halo mass relation parameters are reported in Table 2.3. What are shown are the median values along with the effective uncertainties on those median values that are defined to be the 16% and 84% confidence regions for the parameters. The median values of the parameters are not of particular interest except to note that the approach we have taken leads to a relative low bias relation between WL and halo masses. Of special interest to this work are the uncertainties on the parameters. The uncertainties on the mass bias  $b_i$  are of order of  $\sim 1\%$ , slightly larger at higher redshift and slightly larger for smaller inner fitting radii. The induced uncertainty on the mass trend of the bias  $b_M$  is also  $\sim 1\%$ , independent of inner fitting radius. The uncertainty on the natural logarithm of the variance  $s_i$  is  $\sim 0.04$ , corresponding to  $\sim 2\%$  to the uncertainty on the weak-lensing scatter, and the uncertainty on the mass trend parameter  $s_M$  is  $\sim 5\%$  for both inner fitting radii.

These results imply, then, that given the uncertainties on the cluster contamination presented here for the SPT sample, there is a corresponding systematic uncertainty in our ability to use weak lensing measurements to constrain the cluster halo masses at the 1%, 1.2% and 1.9% level for clusters of mass 3, 6 and  $15 \times 10^{14} M_\odot$ , respectively. Similarly, there is a corresponding systematic uncertainty in our ability to estimate the scatter of the WL mass estimates about the best fit relation of 2%, 2.5% and 4% level for the same cluster masses. This is smaller than other sources of uncertainty found in, e.g., current DES Y1 weak lensing mass calibration analyses, but these systematic floors in our ability to measure cluster masses from mis-centering alone are not negligible. Given that the mis-centering uncertainties presented here are derived from a sample of 250 SPT selected clusters, one can expect that when applied to much larger sample of clusters the uncertainties will fall and then correspond to significantly lower systematic floors in our weak lensing mass calibration analyses.

## 2.6 Summary and Conclusions

In this work we characterise the cluster member contamination of the background shear galaxy sample in two cluster samples as necessary steps toward cluster weak-lensing mass calibrations of both. We model the contamination as a Gaussian in redshift space that is offset from the cluster redshift and whose amplitude varies with redshift, richness and distance from the cluster centre. We model the contamination of each cluster individually, combining the likelihoods across the full cluster sample to extract the best-fit parameters and uncertainties of this model.

Because cluster galaxies can scatter into the background shear galaxy sample due to photo-z uncertainties, this approach relies on decomposing the redshift distribution  $P(z)$  of the shear galaxies into two parts: cluster member contaminants and a true background component. The latter is extracted from an uncontaminated sky region near each cluster. Fitting a model—as opposed to examining the behaviour of large stacks of clusters in richness and redshift bins—allows for the dependences of the contamination on the physical parameters of the cluster to be extracted for the full ensemble rather than estimated on average within binned subsamples. Our method is motivated by the need to quantify the cluster contamination in rather small cluster samples of  $10^2$  to  $10^3$  systems where dividing into binned subsamples is undesirable. Moreover, our method offers advantages over the study of binned subsamples in the presence of strong trends in redshift or richness, as demonstrated in our study.

We present cluster contamination results for the X-ray selected MARD-Y3 cluster sample (208 clusters) (Klein et al., 2019) and the SZE selected SPT cluster sample (250 clusters) (Bleem et al., 2015b) together with the shear galaxy catalogue extracted in the DES Y1 analysis (Zuntz et al., 2018; Hoyle et al., 2018). Our measured cluster contamination will be employed to correct for the associated biases from cluster member contamination in the mass calibration and associated cosmological analyses of these cluster samples.

To evaluate the validity of our results, we carry out a series of tests:

- We demonstrate our method is unbiased using a mock cluster sample that is an order of magnitude larger than the real dataset (subsection 2.4.6). All input parameters are recovered.
- We validate our best-fit model for cluster contamination through comparison to the cluster contamination component after subtracting a re-normalised local field distribution from the observed background distribution of a single cluster. We achieve high signal to noise by stacking these distributions using samples of many clusters within bins of redshift or richness (subsection 2.5.1). By averaging our best fit model over similar bins and comparing, we find that the cluster contamination distribution is well described by a Gaussian, and that the amplitude of this contamination varies with redshift, richness and distance from the cluster centre in a manner that is described by our model.
- We test whether our dataset prefers additional free parameters to follow the potential redshift variation in the redshift offset  $\mu$  or dispersion  $\sigma$  of the Gaussian contamination distribution, its richness trend parameter  $B$  and radial concentration  $c$  (subsection 2.5.2). We find that a single value of each parameter is adequate to describe the current dataset over the full redshift range of interest.
- We quantify the impact that the cluster contamination correction uncertainties from the SPT sample of 250 clusters have on the systematic mass and scatter uncertainty arising from a weak lensing mass calibration analysis (subsection 2.5.4). The systematic uncertainties are derived for three redshift snap shots as well as two different inner fitting radii, and the resulting systematic uncertainty floor on the weak lensing masses lie between 1% and 2% and the scatter between 2% and 4%, depending on cluster mass. We expect larger cluster samples and improved methods to select background galaxies will lead to reductions of these systematic floors.

An intriguing finding is that the redshift dependence of the cluster contamination amplitude exhibits a rather complex behaviour (see Figure 2.8 and discussion in Section 2.5.3). Similar behaviour is exhibited by both cluster samples. Moreover, preliminary results indicate that a similar behaviour is exhibited by the optically selected redMaPPer sample used in the DES Y1 cluster cosmology analysis (DES Collaboration, 2020). A more concrete understanding awaits further work on cluster samples with weak lensing information from DES Y1 and Y3.

## 2.7 Acknowledgments

We acknowledge financial support from the MPG Faculty Fellowship program, the new ORIGINS cluster funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany’s Excellence Strategy – EXC-2094 – 390783311, and the Ludwig-Maximilians-Universität Munich.

Funding for the DES Projects has been provided by the U.S. Department of Energy, the U.S. National Science Foundation, the Ministry of Science and Education of Spain, the Science and Technology Facilities Council of the United Kingdom, the Higher Education Funding Council for England, the National Center for Supercomputing Applications at the University of Illinois at Urbana-Champaign, the Kavli Institute of Cosmological Physics at the University of Chicago, the Center for Cosmology and Astro-Particle Physics at the Ohio State University, the Mitchell Institute for Fundamental Physics and Astronomy at Texas A&M University, Financiadora de Estudos e Projetos, Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro, Conselho Nacional de Desenvolvimento Científico e Tecnológico and the Ministério da Ciência, Tecnologia e Inovação, the Deutsche Forschungsgemeinschaft and the Collaborating Institutions in the Dark Energy Survey.

The Collaborating Institutions are Argonne National Laboratory, the University of California at Santa Cruz, the University of Cambridge, Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas-Madrid, the University of Chicago, University College London, the DES-Brazil Consortium, the University of Edinburgh, the Eidgenössische Technische Hochschule (ETH) Zürich, Fermi National Accelerator Laboratory, the University of Illinois at Urbana-Champaign, the Institut de Ciències de l'Espai (IEEC/CSIC), the Institut de Física d'Altes Energies, Lawrence Berkeley National Laboratory, the Ludwig-Maximilians Universität München and the associated Excellence Cluster Universe, the University of Michigan, the National Optical Astronomy Observatory, the University of Nottingham, The Ohio State University, the University of Pennsylvania, the University of Portsmouth, SLAC National Accelerator Laboratory, Stanford University, the University of Sussex, Texas A&M University, and the OzDES Membership Consortium.

Based in part on observations at Cerro Tololo Inter-American Observatory, National Optical Astronomy Observatory, which is operated by the Association of Universities for Research in Astronomy (AURA) under a cooperative agreement with the National Science Foundation.

The DES data management system is supported by the National Science Foundation under Grant Numbers AST-1138766 and AST-1536171. The DES participants from Spanish institutions are partially supported by MINECO under grants AYA2015-71825, ESP2015-66861, FPA2015-68048, SEV-2016-0588, SEV-2016-0597, and MDM-2015-0509, some of which include ERDF funds from the European Union. IFAE is partially funded by the CERCA program of the Generalitat de Catalunya. Research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Program (FP7/2007-2013) including ERC grant agreements 240672, 291329, and 306478. We acknowledge support from the Brazilian Instituto Nacional de Ciéncia e Tecnologia (INCT) e-Universe (CNPq grant 465376/2014-2).

This manuscript has been authored by Fermi Research Alliance, LLC under Contract No. DE-AC02-07CH11359 with the U.S. Department of Energy, Office of Science, Office of High Energy Physics.

Figures 2.7, 2.12, and A.1 were created with PYGTC (Bocquet & Carter, 2016).

## Affiliations

<sup>1</sup> Faculty of Physics, Ludwig-Maximilians Universität München, Scheinerstr. 1, 81679 Munich, Germany

<sup>2</sup> Max Planck Institute for Extraterrestrial Physics, Giessenbachstrasse, 85748 Garching, Germany



# Chapter 3

## Cluster Member Contamination of the Weak Lensing Shear Profiles from the Dark Energy Survey Year 1 Data: Application to redMaPPer clusters

### 3.1 Introduction

Over the years, galaxy clusters have proven to be powerful cosmological tools. Their abundance and spatial distribution are sensitive to the growth of structures in the Universe as well as cosmic expansion (Haiman et al., 2001; Allen et al., 2011; Dodelson et al., 2016). Furthermore, cluster abundance constrains the  $\sigma_8$  parameter in combination with the mean matter density of the Universe,  $\Omega_M$ , where  $\sigma_8$  measures the amplitude of the linear power spectrum on the scale of  $8h^{-1}$  Mpc.

Currently, studies involving cluster abundances are restricted by their ability to accurately measure cluster masses through calibration of their mass - observable relation (MOR) where the observable is used as a mass proxy. Among the various techniques to calibrate the MOR, gravitational weak lensing is regarded as one of the most robust methods currently available to determine cluster masses. The induced tangential alignment of background galaxies around a foreground cluster is directly sensitive to all of the cluster mass, not just its baryonic component. Furthermore, the weak lensing signal is insensitive to the dynamical state of the cluster. For all of these reasons, many studies have been carried out to specifically estimate cluster masses via gravitational weak lensing (Sheldon et al., 2001; Johnston et al., 2007; von der Linden et al., 2014a,b; Applegate et al., 2014; Hoekstra et al., 2015; Mantz et al., 2015; Okabe & Smith, 2016; Battaglia et al., 2016; Melchior et al., 2017; Simet et al., 2017; Murata et al., 2018; Dietrich et al., 2019; McClintock et al., 2019). But, weak lensing cluster mass estimates are affected by a variety of systematic uncertainties, including shear biases, photometric redshift biases, halo triaxiality, miscentering, cluster

member contamination and projection effects, each of which are significant contributions to the total error budget (Melchior et al., 2017; Medezinski et al., 2018a; Miyatake et al., 2019; McClintock et al., 2019) and hence are limiting the cosmological constraining power.

A recent study combined cluster abundances and weak lensing mass estimates to simultaneously constrain cosmology and the observable – mass relation over the Year 1 (Y1) footprint of the Dark Energy Survey (DES) (DES Collaboration, 2020). DES (DES Collaboration, 2005) is a wide-field imaging survey, covering an area of 5000 deg<sup>2</sup> located in the southern sky that yields approximately 300 million galaxies out to  $z = 1.4$ . The DES Y1 footprint covers about 1800 deg<sup>2</sup> and shares a large overlap with the observed area of the South Pole Telescope (SPT, Carlstrom et al. (2011b)). DES Collaboration (2020) uses the optically - selected redMaPPer cluster catalogue (Rykoff et al., 2014a) to perform their cosmological analysis and relies on the results from McClintock et al. (2019) for the mass estimates using a stacked weak lensing analysis of the same cluster sample. Surprisingly, they found that the cosmological posteriors of their unblinded analysis are in  $5.6\sigma$  tension with the Planck CMB results (Planck, 2020) and in  $2.4\sigma$  tension with the results from the DES 3x2pt cosmological analysis (DES Collaboration, 2018) in the  $\sigma_8 - \Omega_M$  plane. These tensions seem to be driven by the low  $\Omega_M$  value preferred by the DES Y1 cluster sample. After further tests, it was found that the low richness systems with  $\lambda < 30$  are responsible for the tension as these systems push the slope and amplitude posteriors of the richness–mass relation towards lower values compared to the ones preferred by the combination of number counts and weak lensing data at  $\lambda > 30$ , as well as by the analysis of Bleem et al. (2020) using SPT clusters ( $\lambda \geq 40$ ). In conclusion, these tensions might be a consequence of either an overestimation of the selection effect correction at low richness, or due to an unmodelled systematic uncertainty affecting the stacked weak lensing signal of optically selected redMaPPer clusters that is expected to be richness-dependent.

Clearly, there are still aspects of the weak lensing analysis that are not fully understood. We want to take this as a motivation to examine the richness - dependent cluster member contamination for the redMaPPer sample with our own method. We have seen very good results on the analysis of the cluster member contamination with the SPT and MARD-Y3 cluster samples (see Chapter 2), and so we want to apply our technique to the redMaPPer sample as a way for trying to understand whether cluster contamination is the cause for the strange cosmological results.

Briefly, the cluster member contamination is caused by cluster own galaxies that can scatter into the background sample of weak lensing source galaxies, due to the photometric uncertainty on their redshift estimates. This resulting contamination dilutes the weak lensing shear measurement, as we are averaging over galaxies that do not exhibit a shear signal at all. This makes it necessary to boost the measured shear signal in order to recover the true signal. Therefore, multiple studies have referred to this rescaling of the weak lensing signal as applying a *boost factor* (Sheldon et al., 2004; Applegate et al., 2014; Hoekstra et al., 2015; Gruen et al., 2014; Simet et al., 2017; Melchior et al., 2017; Medezinski et al., 2018a; Leauthaud et al., 2017; McClintock et al., 2019; Varga et al., 2019).

The most recent method to characterise the cluster member contamination involves the

decomposition of the source galaxies into a cluster member component and a background component. A contamination fraction is estimated based on a decomposition of the photometric redshift probability distribution function estimates,  $P(z)$ , of source galaxies. This method has been applied in e.g. Chang et al. (2018); Stern et al. (2019); McClintock et al. (2019); Pereira et al. (2020) and most recently in Paulus et al, in prep. and has been validated on simulations by Varga et al. (2019).

Here, we make use of the  $P(z)$  decomposition method to estimate the cluster member contamination for the redMaPPer sample over DES Y1. In this work we are interested in modelling the cluster contamination of the redMaPPer sample on an individual cluster basis and in a manner that allows us to fit for a variable normalisation in cluster redshift. In the end, we are comparing our results to Varga et al. (2019) who used the  $P(z)$  decomposition method to aid the stacked weak lensing mass calibration of redMaPPer clusters over the DES Y1 area. To this end, we are adopting the same redshift - richness binning as well as a similar radial range.

This chapter is structured in the following way: In section 3.2 we present the DES Y1 data used in this work as well as the redMaPPer cluster sample. In section 3.3 we introduce the framework as well as the formalism of the  $P(z)$  decomposition method and in section 3.4 we present our results for the redMaPPer analysis. Finally, in section 3.5 we compare our results to Varga et al. (2019). We assume a flat  $\Lambda$ CDM cosmology with  $\Omega_m = 0.3$  and  $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$  throughout.

## 3.2 Data

In the following subsections we describe the shear and photo-z catalogues from DES as well as the optically selected redMaPPer cluster catalogue.

### 3.2.1 DES Y1

In the DES Y1 observations, approximately  $1800 \text{ deg}^2$  of the observed  $5000 \text{ deg}^2$  survey were covered in the four bands, g, r, i, z. Compared to the SV data, the imaging is shallower but it covers a significantly larger area. These observations were then processed via a range of photometric data reduction steps into the Y1 GOLD catalogue (Drlica-Wagner et al., 2017), which represents the main science catalogue of DES. In Y1 areas of the southern sky overlapping with the SPT footprint ( $1321 \text{ deg}^2$ ) and the "stripe 82" deep field region ( $116 \text{ deg}^2$ ) of the Sloan Digital Sky Survey (SDSS) were targeted. Due to masking only about  $1500 \text{ deg}^2$  of the main survey is utilised. This area was then further processed by the METACALIBRATION algorithm (Huff & Mandelbaum, 2017; Sheldon & Huff, 2017) to define a shear galaxy catalogue (Zuntz et al., 2018). This shear catalogue contains ellipticity measurements  $e_i$  for each galaxy and corresponding shear response terms,  $R = R_\gamma + R_{sel}$ . The shear response terms describe the change in ellipticity of each source due to a small applied shear. The photometric redshifts  $P(z)$ 's were calculated via the BPZ template based algorithm (Hoyle et al., 2018) and two different redshift estimates were derived.

One is based on galaxy colours, which are extracted using the multi-epoch, multi-object fitting algorithm (MOF), and are listed in the GOLD catalogue. The second is based on a model obtained from METACALIBRATION. These second redshift estimates are required to correctly account for the selection response correction, but they were found to show a larger scatter compared to the MOF estimated redshifts. Therefore, in previous DES cluster weak lensing analyses (McClintock et al., 2019; Pereira et al., 2020, Paulus et al, in prep) only the redshift estimates from METACALIBRATION were used for selecting and weighting source lens pairs. In our analysis, we are doing the same.

### 3.2.2 redMaPPer Cluster Catalogue

We are using the same cluster sample used in the weak lensing analysis by McClintock et al. (2019) and hence also in the cluster contamination analysis by (Varga et al., 2019, hereafter V19). This volume limited galaxy cluster sample was detected in the DES Y1 photometric data by using the **red-sequence Matched-filter Probabilistic Percolation** (redMaPPer) cluster finding algorithm v6.4.17 (Rykoff et al., 2014b, 2016b). The redMaPPer cluster catalogue used in this analysis is a “locally volume-limited” sample, meaning that a galaxy cluster is only included in the sample if all the cluster galaxies that are brighter than the set luminosity threshold used to define the cluster richness in redMaPPer lie above  $10\sigma$  in  $z$ ,  $5\sigma$  in  $i$  and  $r$ , and  $3\sigma$  in  $g$  according to the survey MOF depth maps (Drlica-Wagner et al., 2017). At this threshold the sample is  $> 90 - 95\%$  complete.

Galaxy clusters are identified as overdensities of red-sequence galaxies by the redMaPPer algorithm. Beginning with a preliminary set of spectroscopic seed galaxies, a model for the local red-sequence is fit, finding cluster candidates while allocating a membership probability to each potential member. The clusters are centred on bright galaxies that have been selected using an iteratively self-trained matched-filter method. This way a probability is assigned to each galaxy of the cluster of being the central galaxy. The final membership probabilities of all galaxies in the field are assigned based on spatial, colour, and magnitude filters.

The richness  $\lambda$  is estimated via the sum of the membership probabilities over all galaxies that lie within a pre-defined, richness-dependent projected radius  $R_\lambda$ . This radius is defined in relation to the cluster richness via  $R_\lambda = 1.0(\lambda/100)^{0.2} h^{-1}\text{Mpc}$ . In Rykoff et al. (2012) it was found that this relation minimises the scatter between richness and X-ray luminosity. A redshift was estimated for each cluster by maximising the probability that the observed colour-distribution of likely members matches the self-calibrated red-sequence model of redMaPPer.

The final cluster sample used in this analysis contains 5750 clusters ranging from  $0.2 < z_{\text{cl}} < 0.65$  with  $\lambda > 20$ .

### 3.3 Cluster member contamination

For a more in depth description and derivation of the cluster member formalism we refer the interested reader to a variety of literature that deals with this topic in detail, e.g. (Melchior et al., 2017; McClintock et al., 2019; Varga et al., 2019, Paulus et al, in prep). In general, a weak lensing measurement is affected by cluster contamination in such a way that in annuli around the cluster the shear signal is diluted. This is due to the existence of cluster own galaxies in the background sample. By estimating the contamination, the introduced bias in the weak lensing signal can be corrected. The method used in this work measures a cluster contamination fraction,  $f_{\text{cl}}$ , which is defined as the fraction of cluster member galaxies that scatter into the background shear sample. We infer  $f_{\text{cl}}$  from the photometric redshift probability distribution function  $P(z)$  of appropriately selected and weighted source galaxies behind the cluster.

#### 3.3.1 Source background distribution

Photometric algorithms are used to extract estimates of photometric redshifts (photo-z's) of clusters, which are needed to construct a catalogue of background source galaxies. These redshifts are associated with large uncertainties (Hoyle et al., 2018; Hildebrandt et al., 2017), meaning that the final galaxy source catalogue might include foreground galaxies or galaxies at the cluster redshift even after an attempt to exclude them using their photo-z's. The information gained from the probability distribution of the galaxy photo-z is used in this work to constrain the contamination.

Background galaxies at different redshifts contribute to the measured shear signal with different amplitudes. This is accounted for by the inverse of the *critical surface density*:

$$\Sigma_{\text{crit}}(z_l, z_s) = \frac{c^2}{4\pi G} \frac{D_s(z_s)}{D_l(z_l) D_{ls}(z_l, z_s)}, \quad (3.1)$$

where  $D_s$ ,  $D_l$  and  $D_{ls}$  denote *angular diameter* distances to the source galaxy, the lens, and between the lens and the source, respectively.

As source redshifts are generally estimated from photometry and are described by a probability distribution  $P_{\text{phot}}(z_s)$  for each source galaxy, one can estimate an *effective* critical surface density as

$$\langle \Sigma_{\text{crit}}^{-1} \rangle_{i,j} = \int dz_s p_{\text{phot}}(z_{s,i}) \Sigma_{\text{crit}}^{-1}(z_{l,j}, z_{s,i}), \quad (3.2)$$

where  $i$  and  $j$  are the indices of the source and the lens in a particular lens-source pair.

The source background distribution,  $P(z)$ , of a specific cluster contains galaxies that have been selected and weighted in several logarithmically spaced radial bins from the cluster. The source galaxies have been selected in such a way that  $z_{\text{mean}} > z_{\text{cl}} + 0.1$ , where  $z_{\text{mean}}$  is the mean redshift of each background galaxy and  $z_{\text{cl}}$  is the cluster redshift. The

offset of  $\Delta z = 0.1$  was chosen to produce a good quality background sample with low cluster contamination, given typical photo-z uncertainties. Given this sample, we construct the estimator of the true redshift distribution  $P(z)$  using the redshift estimator  $z_{\text{MC}}$ , which is a random sample of the redshift probability distribution of each source galaxy. The selected background galaxies are then weighted by  $\omega_i R_{\gamma; i}^T$ .  $\omega_i$  is defined by

$$\omega_{i,j} \equiv \Sigma_{\text{crit}}^{-1} \left( z_{l_j}, \langle z_{s_i}^{\text{MCAL}} \rangle \right) \text{ if } \langle z_{s_i}^{\text{MCAL}} \rangle > z_{l_j} + 0.1, \quad (3.3)$$

with  $\langle z_s^{\text{MCAL}} \rangle$  being the mean redshift of the source galaxy estimated from METACALIBRATION photometry.  $R_{\gamma; i}^T$  is the shear response term measured from the distorted images of the galaxies.

### 3.3.2 P(z) decomposition method

We study the background sample in this analysis within the same parameters as for our previous work (Paulus et al, in prep.), namely within 9 logarithmically spaced radial bins from the cluster centre ranging from 0.25 Mpc to 10.4 Mpc. This allows for a more direct comparison between the samples. The last radial bin was chosen to represent the local field distribution of each cluster. It is far enough from the cluster to be useful to constrain the local field distribution of a cluster, but also not so far that one would start to worry about the field distribution starting to be severely affected by sample variance. Field depth variations in the DES Y1 survey data set do not allow for the usage of the survey average. This would lead to the introduction of additional noise and a bias in the estimate of the cluster member contamination.

The background galaxy redshift distribution is modelled as a product of two terms; a cluster member component and a local field component of appropriately selected and weighted galaxies. We model the redshift distribution of the cluster members as a Gaussian. This was shown by Varga et al. (2019) (section 3.2.5) to be a reasonably good representation of the contaminating cluster galaxies. Hence, the redshift distribution takes the following form:

$$\begin{aligned} P_{\text{pred}}(z, R) &= f_{\text{cl}}(R) P_{\text{mem}}(z) + (1 - f_{\text{cl}}(R)) P_{\text{field}}(z) \\ &= f_{\text{cl}}(R) \mathcal{N}(z, \mu, \sigma) + (1 - f_{\text{cl}}(R)) P_{\text{field}}(z), \end{aligned} \quad (3.4)$$

where  $z$  is the redshift of the background source,  $R$  refers to the area-weighted average radius of the radial bin in Mpc,  $f_{\text{cl}}(R)$  is the contamination fraction estimated in bins of radius  $R$ ,  $\mathcal{N}(z, \mu, \sigma^2)$  describes our assumption of the distribution of cluster members to be Gaussian and finally,  $P_{\text{field}}(z)$  is the uncontaminated background galaxy redshift distribution.

At the moment our correction depends very sensitively on the exact bin limits that were chosen. However, the true function is almost certainly a smooth function and henceforth should be modelled as such. Therefore, our final model for estimating the contamination fraction allows for discrete amplitude fit parameters in redshift bins but with a Gaussian

smoothing kernel applied. This ‘‘Multi-Gaussian method’’ presents a convenient way to interpolate over the full redshift range between the binned results. It takes the following form:

$$A(z_{\text{cl}}) = \exp [A_{\text{asymptote}} + \sum_i e^{-\left(\frac{z_{\text{cl}} - z_i}{\sigma_{\text{corr}}}\right)^2} A_i] \quad (3.5)$$

Here  $A(z_{\text{cl}})$  describes the amplitude at an arbitrary redshift with the sum running over all amplitude fit parameters  $A_i$ , where  $i$  here refers to a specific value in an assumed redshift array. The Gaussian weight factor leads to ‘‘nearby’’ (in redshift) amplitudes contributing more strongly than the  $A_i$  that are far away. The variance,  $\sigma_{\text{corr}}^2$ , in the Gaussian weight is an additional parameter of the model that acts as a correlation length. There is not enough information in our samples to constrain the correlation length well, and so through comparison of the interpolation function with the originally binned results we decided to fix  $\sigma_{\text{corr}}$  at 0.1. The  $A_{\text{asymptote}}$  parameter sets the value  $A(z_{\text{cl}})$  when one is well away from the specific amplitudes  $A_i$  and so can be adjusted to be the typical observed contamination amplitude. We set this to -2.21. Applying an exponential at the very end guarantees that the function, together with all error intervals, will always be positive. Finally, with the contamination amplitude for the multi-Gaussian method determined as above we estimate the contamination fraction at a given cluster redshift  $z_{\text{cl}}$ , richness  $\lambda$  and projected radius from the centre  $R$  in the following way:

$$(1 - f_{\text{cl}}(R, \lambda, z_{\text{cl}}))^{-1} = 1 + A(z_{\text{cl}}) \left(\frac{\lambda}{\lambda_{\text{piv}}}\right)^B \Sigma(R, r_s), \quad (3.6)$$

Due to the natural parameterisation with  $A(z_{\text{cl}}) > 0$  and an arbitrary  $B$ , it is ensured that  $0 < f_{\text{cl}} < 1$ , where  $B$  describes the power-law dependence of the contamination fraction on cluster richness. For the cluster richness pivot,  $\lambda_{\text{piv}}$ , the median value of the cluster sample was assumed. For our redMaPPer sample,  $\lambda_{\text{piv}} = 27$ .  $\Sigma(R, r_s)$  characterises the projected NFW profile (Navarro et al., 1996), where  $r_s$  is the scale radius. It is given by  $r_s = r_{200}/c$ , where  $r_{200}$  is defined as the radius within which the mass density of the halo is 200 times the critical density of the Universe.  $c$  is the so called concentration. The values of  $r_{200}$  for each cluster are inferred from the mass - richness relation found in McClintock et al. (2019). A normalisation to unity at  $\Sigma(R/r_s = 1.01)$  was implemented. For the binned method, instead of iterating over a redshift array for the individual discrete amplitude parameters, we adopt individual amplitude parameters for assumed redshift bins. Therefore, the contamination fraction,  $f_{\text{cl}}(R, \lambda, z_{\text{cl}})$ , is defined in the following way:

$$(1 - f_{\text{cl}}(R, \lambda, z_{\text{cl}}))^{-1} = 1 + A_i(z_{\text{cl}}) \left(\frac{\lambda}{\lambda_{\text{piv}}}\right)^B \Sigma(R, r_s), \quad (3.7)$$

where  $A_i(z_{\text{cl}})$  describe the set of discrete cluster redshift dependent normalisation parameters for specific cluster redshift bins and  $i$  refers to the assumed redshift bin. Finally, with six amplitude parameters  $A_i$  at different redshifts, our multi-Gaussian model then has 10 free parameters including the mass - trend  $B$ ,  $\mu$  and  $\sigma$  defined in the Gaussian distribution and the concentration,  $c$ .

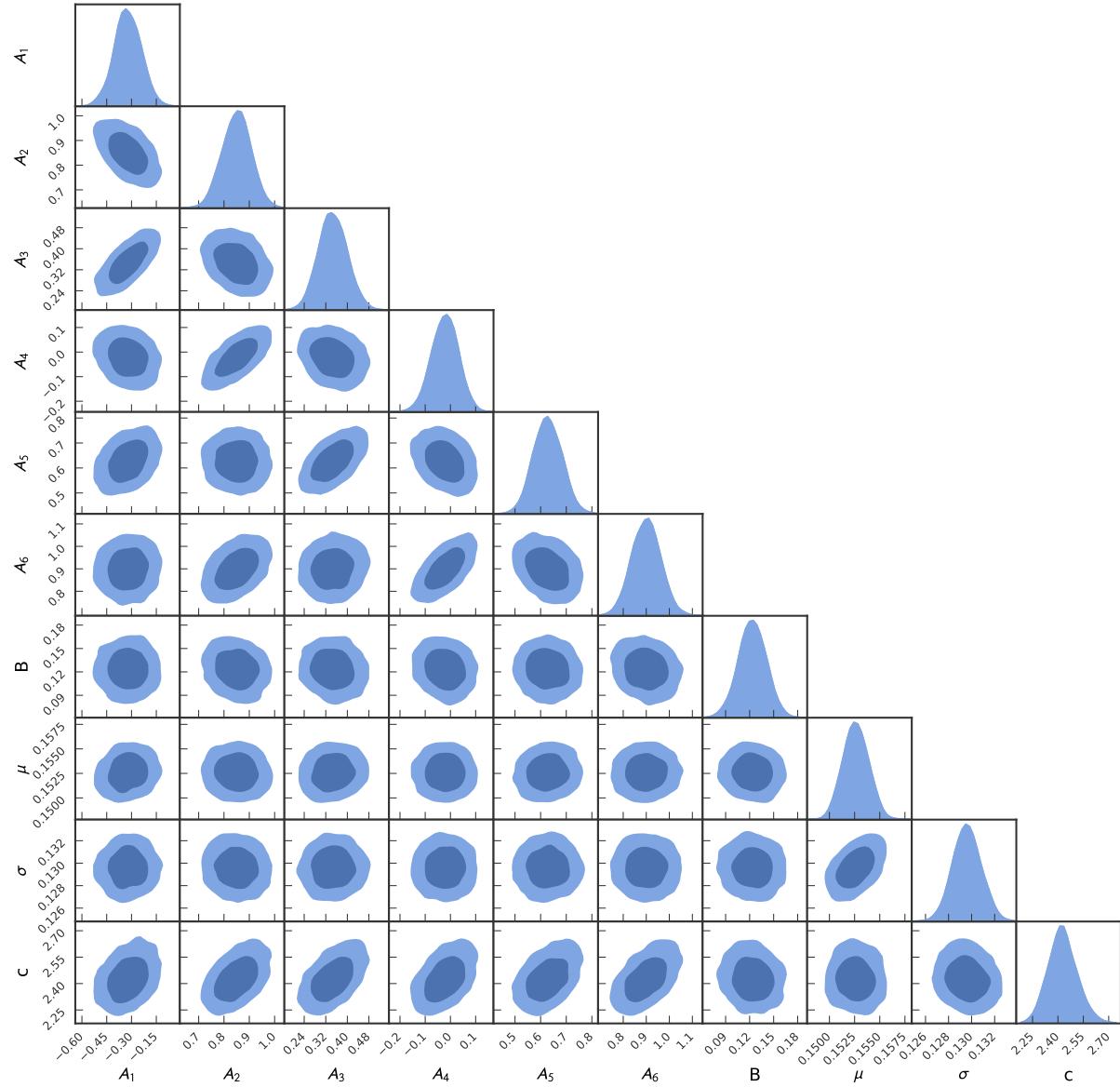


Figure 3.1: Constraints on the assumed free parameters of the contamination fraction model for the redMaPPer sample, where discrete amplitude fit parameters  $A_i$  were assumed for a given redshift array.

## 3.4 Cluster contamination measurements

The adopted ‘‘Multi-Gaussian method’’ allows for the actual redshift evolution of the cluster contamination in the redMaPPer sample to be traced. The assumed redshift array for the discrete amplitude fit parameters is  $z_{\text{cl}} \in \{0.15, 0.25, 0.35, 0.45, 0.55, 0.65\}$ . For the binned method we assumed the following redshift bins:  $z_{\text{cl}} \in \{0.20, 0.25, 0.3, 0.35, 0.4, 0.45, 0.5, 0.55, 0.6, 0.65\}$  where always the lower limit is included and the upper limit is not. If not stated otherwise, all results presented in this chapter are shown for the multi-Gaussian method.

Figure 3.1 shows the constraints on the best fit parameters of the redMaPPer cluster sample. Discrete normalisation fit parameters for a given cluster redshift array were assumed. The posteriors of the free parameters assumed in the model were sampled using the package emcee (Foreman-Mackey et al., 2013). It enables a parallelised exploration of the parameter space. 120 walkers with 1000 steps each were used, discarding the first 30,000 samples. In Figure 3.2 we show the redshift normalisation versus cluster redshift for the redMaPPer sample, where we compare the resulting normalisation parameters from the Multi-Gaussian method (solid line),  $A(z_{\text{cl}})$ , with the results from the binned method (stars),  $A_i(z_{\text{cl}})$ . The band around the normalisation parameters from the Multi-Gaussian method signifies the  $1\sigma$  region of uncertainty. As previously shown in Paulus et al, in prep., also here it is clear that a simple power law would not have been sufficient to describe the redshift variation of the cluster contamination for this cluster sample, as it exhibits a rather complex behaviour. Figure 3.3 shows a comparison between the results of the amplitude parameters of this work with those from Paulus et al, in prep., where we looked at the Sunyaev-Zel’dovich effect selected South Pole Telescope SPT cluster sample (Bleem et al., 2015b) and the X-ray selected MARD-Y3 clusters sample (Klein et al., 2019). In general, all three samples follow to some degree the same redshift behaviour. As one can see, the amplitude of the first feature up to a redshift of about  $z_{\text{cl}} = 0.35$  is significantly smaller for the redMaPPer sample when compared to the SPT or MARD-Y3 sample. When looking at possible reasons for that, we looked at the best fit parameters for each of the samples and found that the richness scaling parameter,  $B$ , is significantly smaller for the redMaPPer sample than for the other two. We find  $B_{\text{REDMAPPER}} = 0.12$ , whereas  $B_{\text{SPT}} = 0.58$  and  $B_{\text{MARD-Y3}} = 0.48$  (Paulus et al, in prep.). It is not intuitively clear why these parameters should alter by such a large amount. One can look at the resulting best fit parameters for especially  $B$ , but also  $\mu$ ,  $\sigma$  and  $c$  for different richness cuts of the redMaPPer sample. The idea is that by every cut we are closer to a cluster sample that represents the richness distribution of clusters within the SPT or MARD-Y3 samples and one might expect the  $B$  parameter to shift towards a value close to that of either SPT or MARD-Y3.

Figure 3.4 shows the resulting best fit parameters for  $B$ ,  $\mu$ ,  $\sigma$  and  $c$  for richness cuts at  $\lambda > 20$  (full sample),  $\lambda > 30$ ,  $\lambda > 40$  and  $\lambda > 50$ . A clear trend is detectable for all the parameters.  $\mu$  and  $\sigma$  drop with an increasing cut in richness, whereas  $B$  and  $c$  are increasing the larger the cut in richness. The  $B$  parameters never quite reach the same scaling as found for SPT or MARD-Y3. The low concentration for the full sample could be explained by the fact that we are using the  $M - \lambda$  relation found in (McClintock et al., 2019) to estimate  $r_{200}$ , needed for our NFW profile. As discussed in e.g. DES Collaboration (2020),

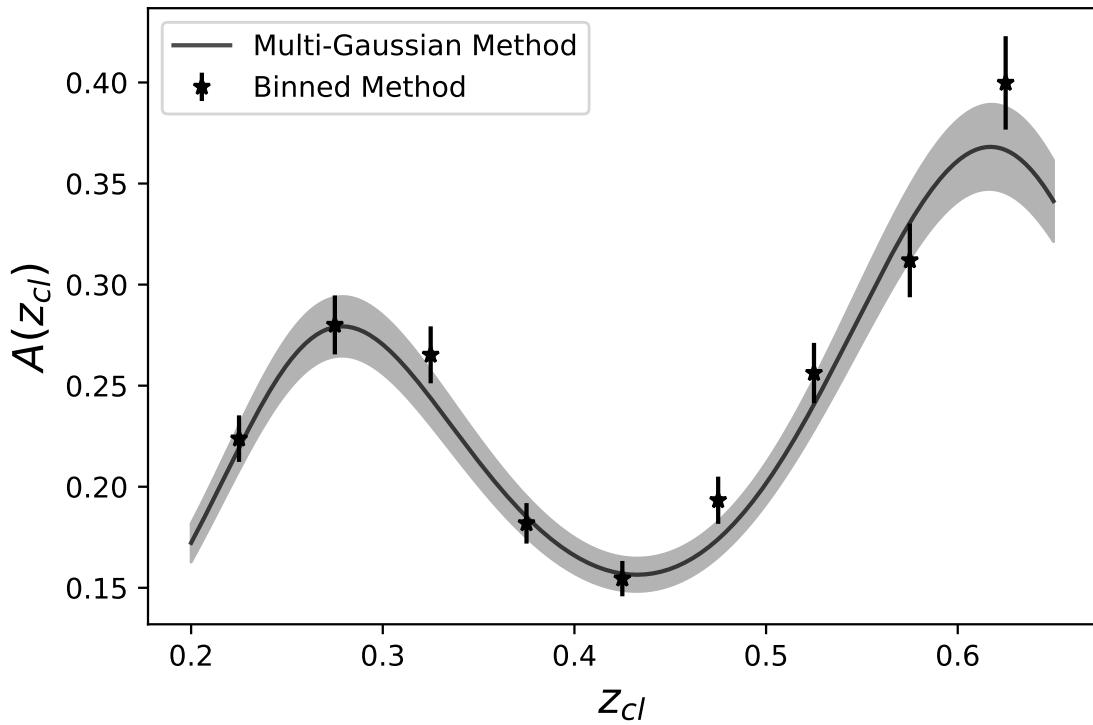


Figure 3.2: The plot shows the different normalisation parameters of the cluster contamination versus cluster redshift for the redMaPPer sample. We compare the best fit normalisation parameters resulting from the binned method,  $A_i(z_{cl})$ , indicated by the stars, to the resulting normalisation parameters for each cluster,  $A(z_{cl})$ , from the Multi-Gaussian method, indicated by the solid line. Both methods appear to be in good agreement and portray a rather complex behaviour in redshift.

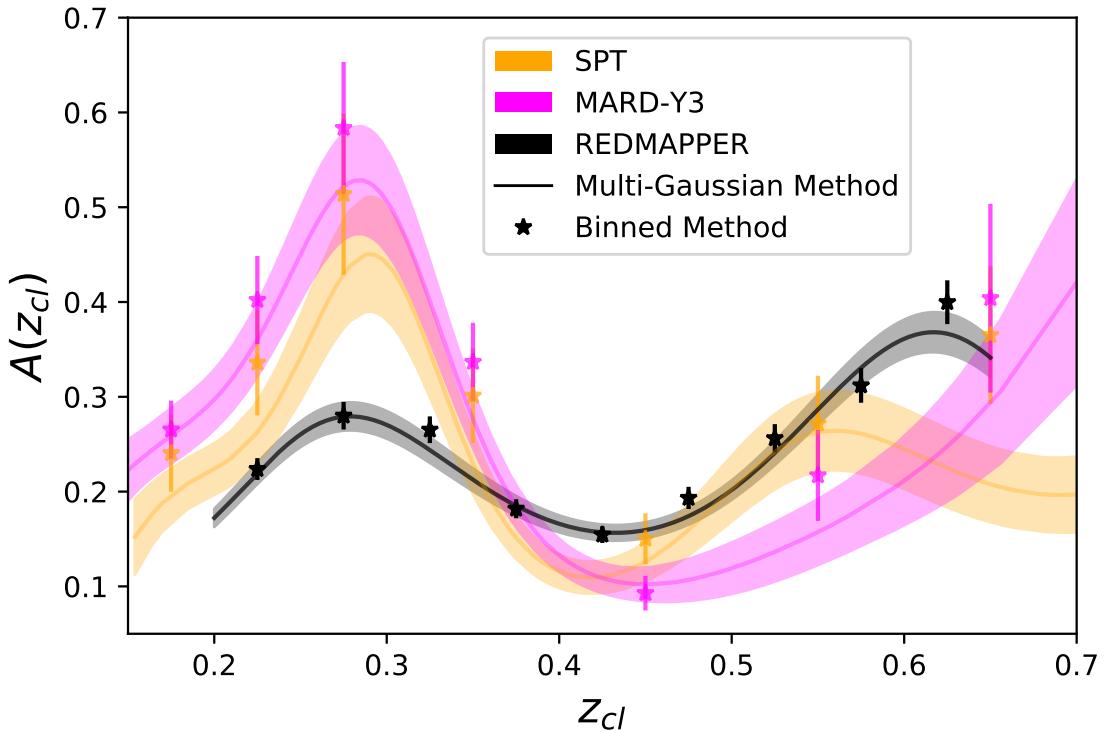


Figure 3.3: The plot shows the different normalisation parameters of the cluster contamination versus cluster redshift for the redMaPPer sample (black), the SPT sample (orange) and the MARD-Y3 sample (magenta). We always show the best fit normalisation parameters resulting from the binned method,  $A_i(z_{\text{cl}})$ , indicated by the stars, and the resulting normalisation parameters for each cluster,  $A(z_{\text{cl}})$ , from the Multi-Gaussian method, indicated by the solid line for each of the samples. The general shape of the redshift behaviour is similar in all 3 samples but the amplitude of the first feature for the redMaPPer case is significantly smaller than for the other two samples.

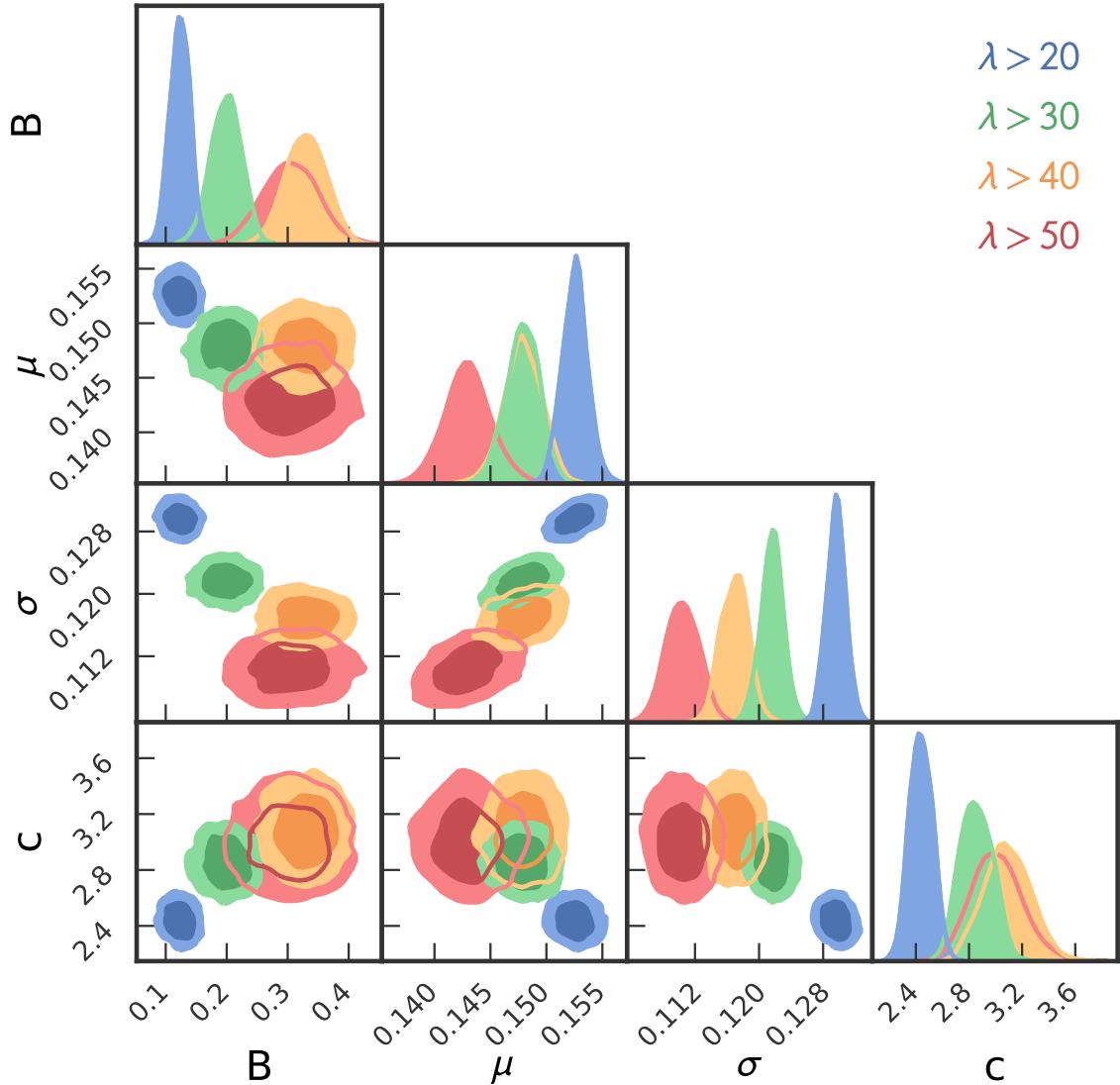


Figure 3.4: The contour plot show the resulting best fit parameters for  $B$ ,  $\mu$ ,  $\sigma$  and  $c$  for different richness cuts of the redMaPPer sample. These are  $\lambda > 20$  (full sample) in blue,  $\lambda > 30$  in green,  $\lambda > 40$  in orange and  $\lambda > 50$  in red. We find that the  $B$  parameters are indeed increasing with an increasing cut in richness but never reach the same amplitudes in  $B$  as for SPT or MARD-Y3.

the slope and amplitude posteriors of this richness - mass relation are pushed towards lower values by the systems with  $\lambda < 30$ . The masses for low richness systems seem to be underestimated therefore. It is also exactly for that sample that includes clusters below a richness of 30 that the  $B$  parameter is surprisingly small and statistically further away from the other sample cuts. In general, these samples have very different richness distributions, but also projection effects might play a role for the redMaPPer sample. Correlated and uncorrelated structures along the line of sight can impact photometric cluster richness estimates, as the width of the red sequence as well as photometric uncertainties limit the resolution a photometric cluster-finding algorithm can achieve along the line of sight (Cohn et al., 2007). Therefore, one might expect richness estimates to be contaminated by the galaxy content of nearby structures. Multiple sources of observational evidence have been found for projection effects in the SDSS redMaPPer cluster catalogue (e.g., Farahi et al., 2016; Zu et al., 2017). Recently, Costanzi et al. (2018) set out to quantify the effect projection effects have on the DES redMaPPer cluster catalogue and they found that projection effects can substantially alter the observed richness. The effect is especially strong in rich galaxy clusters due to the abundance of correlated structures around these systems. As we are mostly experiencing a small scaling in richness for the full sample, which is dominated by low richness systems this might, projection effects are not likely causing this effect but might contribute to it. In the end it seems there is some internal variance in the data, with the pattern not being as simple and richness dependent as we had originally thought.

### 3.4.1 Model-free estimation of the cluster contamination

We thought of a way to determine the cluster contamination without the need for a model prediction that would help us understand if our model was describing the data in an accurate way. This method estimates the contamination fraction by subtracting a re-normalised local field distribution,  $P_{\text{field}}(z)$ , from the observed background distribution of the cluster,  $P_{\text{meas}}(z, R)$ . The re-normalisation was achieved by taking the ratio between the respective high redshift tails ( $z > z_{\text{cl}} + 0.8$ ) of the predicted background distributions and the local field distributions, as we assume the cluster contamination to be minimal there and can be subtracted away. This results in a model independent measurement of the cluster contamination:

$$f_{\text{cl}}(R, \lambda) P_{\text{mem}}(z) = P_{\text{meas}}(z, R) - \frac{P_{\text{meas}}(z > z_{\text{cl}} + 0.8)}{P_{\text{field}}(z > z_{\text{cl}} + 0.8)} P_{\text{field}}(z) \quad (3.8)$$

Applying this method over an ensemble of clusters, combined in a space of  $z - z_{\text{cl}}$ , ensures enough signal to noise to identify features that can be associated with the excess of galaxies near the cluster redshift. An average cluster contamination for clusters within a certain redshift or richness bin can be extracted by integrating over the excess associated with the cluster contamination.

Figure 3.5 shows a stacked redshift distribution in the redshift range  $0.2 < z_{\text{cl}} < 0.3$  in blue together with the best fit model in orange, averaged over the same cluster redshift

range. This demonstrates that the adopted model provides a good description of the data.

In Figure 3.6 we show the cluster contamination fraction  $f_{\text{cl}}$  as a function of radius for different cluster redshift bins. We compare the model predicted cluster contamination fraction (solid lines with  $1\sigma$  shaded regions), colour coded by the cluster redshift bins over which the predicted  $f_{\text{cl}}$  values of each cluster in these redshift bins have been averaged, to the contamination fractions  $f_{\text{cl}}$  estimated by subtracting a re-normalised field distribution from the full  $P(z)$  as described above, shown as points with error bars. The points are the integral over the excess of galaxies within the  $\pm 3\sigma$  region around the cluster member contamination peak, like the peaks shown in Figure 3.5. We find a fairly good agreement between the model prediction and the model independent estimation of the cluster contamination within each redshift bin. This confirms that our adopted radial model for the cluster contamination is a good enough description of the data.

### 3.4.2 Redshift variation of cluster contamination

The redshift trend of the cluster contamination amplitude is complex, as shown in Figure 3.2 and has previously been shown in Paulus et al, in prep. There we have already considered the underlying causes of this behaviour, so we refer the interested reader to this paper and will instead just give a brief summary here. Redshift trends in the performance of the photo-z's might help us understand the features that we see. The shifting of the 4000 Å break of cluster member galaxies from the  $g$  to the  $r$  band at  $z \sim 0.35$  impacts the photo-z uncertainties and at this transition one might expect the photo-z scatter and outlier fraction to change, leading to rapid changes in the fraction of cluster galaxies that scatter into the background source galaxy population. As one pushes to higher redshift the cluster galaxies become fainter, and a smaller fraction of these galaxies are sufficiently bright to be included in the shear galaxy sample. At the same time, due to the cut in photo-z of the shear galaxy sample at  $z_{\text{cl}} + 0.1$  and due to the increasing angular diameter distance, there is an ever smaller number of shear galaxies available within bins of fixed metric radius as one pushes to higher redshift.

## 3.5 Comparison to the literature

In this section we compare our findings to those of V19, who has carried out a similar analysis to extract the cluster contamination fraction of clusters, but in a stacked analysis for different redshift-richness bins, rather than for individual clusters. The assumed cluster redshift bins and richness bins are the following:  $z_{\text{cl}} \in \{0.20, 0.35, 0.5, 0.65\}$  and  $\lambda \in \{20, 30, 45, 60, \infty\}$ , where the lower limit is included and the upper limit is not. In total that gives 12 redshift - richness bins. The background sample in their analysis is studied in 11 logarithmically spaced radial bins from the cluster centre ranging from 0.2 Mpc to 30 Mpc and the  $P(z)$  of sources in the outermost radial bin was chosen to represent the field component. Their so called “boost factor estimates” of the contamination fraction,

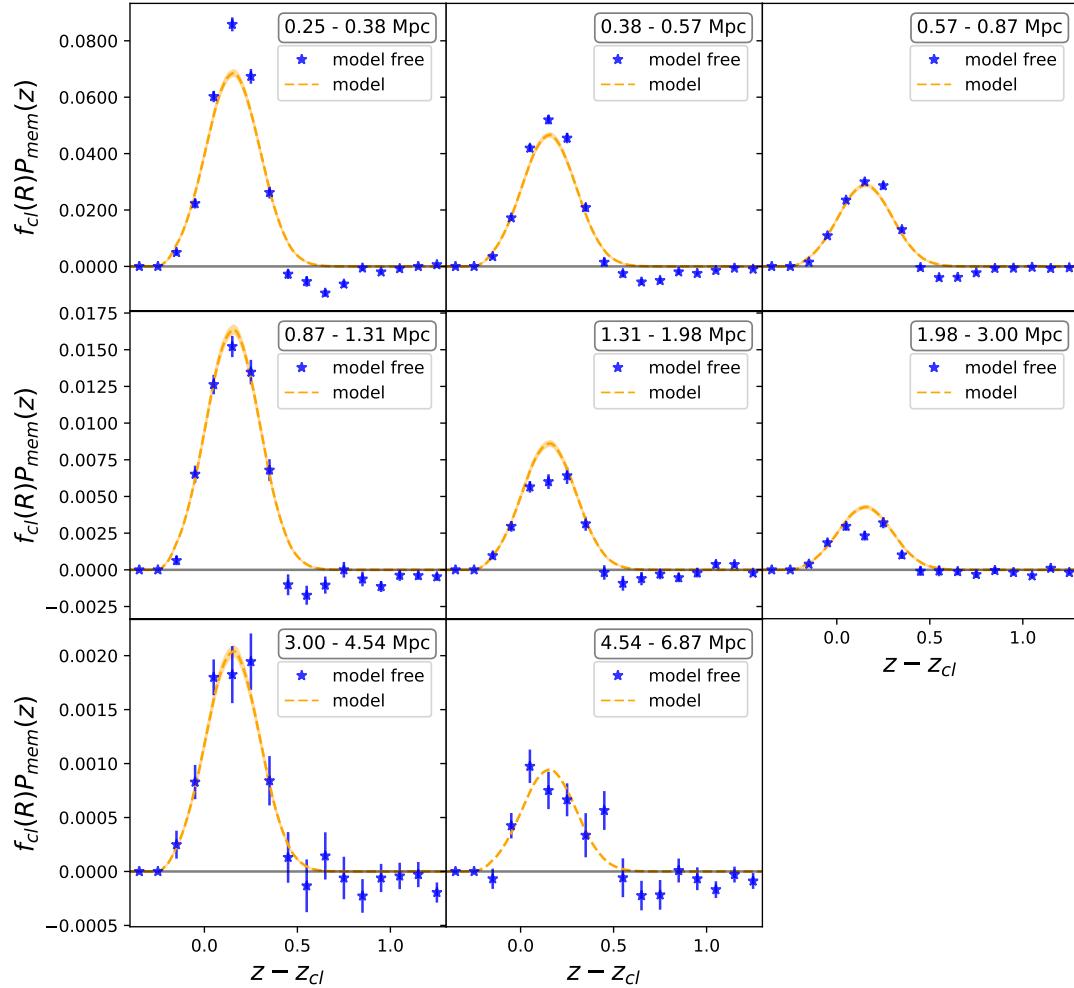


Figure 3.5: An example field-subtracted redshift distribution stacked in  $0.2 < z_{cl} < 0.3$  in blue along with the best fit model in orange, stacked in the same cluster redshift range.

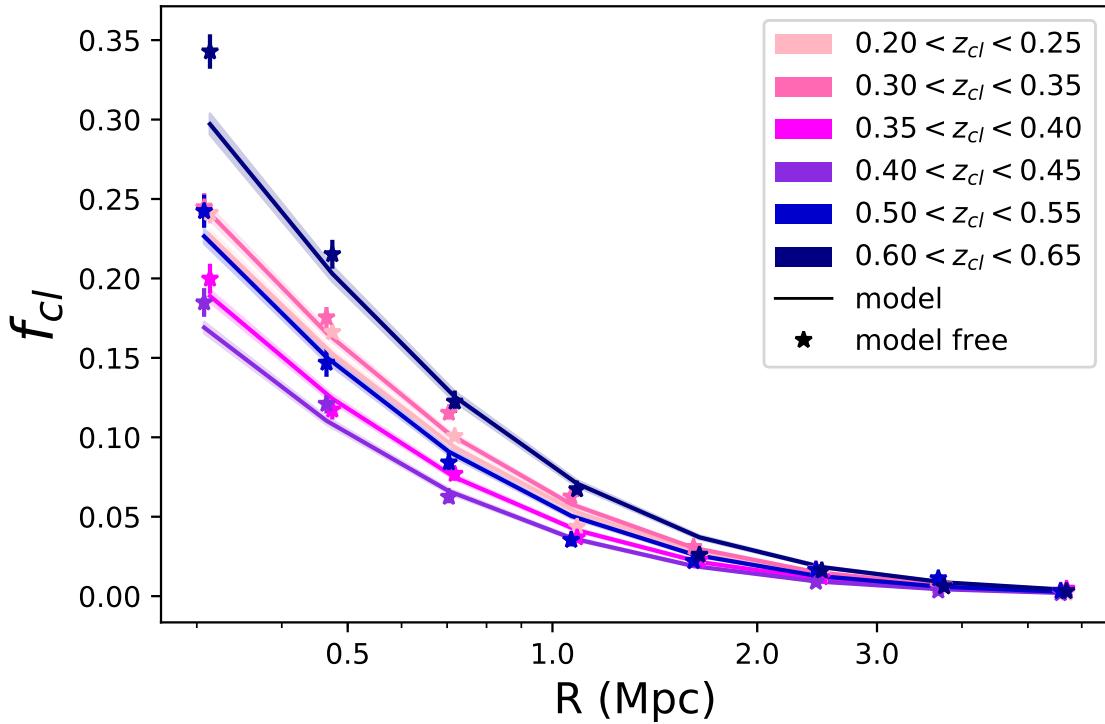


Figure 3.6: The solid line with  $1\sigma$  shaded regions show the best fit model for the cluster contamination  $f_{cl}$  as a function of radius for different colour coded redshift bins. Points with error bars and similar colours represent the cluster contamination extracted from the field subtracted redshift distributions stacked for all clusters within the same redshift ranges. The measurements from the stacked clusters show similar radial and redshift behaviour as the models.

$f_{cl}(R)$ , are modelled using an NFW profile:

$$(1 - f_{cl}(R))^{-1} = 1 + A \Sigma(R, R_s), \quad (3.9)$$

where the normalisation  $A$  and the scale radius  $R_s$  are free parameters. No scaling factor in richness is assumed. They also employ the decomposition method as described above with the assumption that the cluster member galaxies can be modelled as a Gaussian.

In order to compare our results, we strive to analyse our sample in the same manner as it was done in V19. Meaning that we adopt the same redshift-richness binning and fit each bin individually as well as adopting the same radial range. Furthermore, we adopt the last radial bin to represent the field component and exclude the richness dependence from our model of the cluster contamination due to the small assumed richness bins. A notable difference is that we set the concentration,  $c$ , as a free parameter, whereas in V19 the scale radius,  $R_s$ , is allowed to be free. This means that we introduce an extra step in our analysis, as we need to estimate  $r_{200}$  from the mass-richness scaling relation found in McClintock et al. (2019).

As described in Section 3.4.1 we can estimate a model independent measure of the cluster contamination, this time for the individual redshift - richness bins assumed for this analysis and compare it to our model prediction for the same binning. In Figure 3.7 we show the resulting field-subtracted redshift distributions in blue along with the best fit models in orange stacked in the various cluster redshift - richness bins and compare it to the the model prediction when using the cluster contamination results from V19 and multiply them with our results for the cluster member distribution,  $P_{\text{mem}}$  in magenta. The plot shows all results for the first radial bin. We find relatively good agreement between our model independent estimate of the cluster contamination and our model prediction for most bins, but find that the results including the cluster contamination estimates from V19 are significantly lower in amplitude for most bins.

Figure 3.8 shows the cluster contamination fraction  $f_{cl}$  as a function of radius for the 12 different cluster redshift and richness bins. We compare the model predicted cluster contamination fraction (orange solid lines with  $1\sigma$  shaded regions) for the different redshift - richness bins over which the predicted  $f_{cl}$  values of each cluster in these bins have been averaged, with the model free estimation of the contamination fraction, shown as points with error bars in blue. Furthermore, we show the cluster contamination estimates from V19 as magenta dashed lines with  $1\sigma$  shaded regions. There is good agreement between the model and the directly measured contamination within the bins and even the more complex redshift behaviour is identifiable within the three redshift bins for each lambda bin. But the cluster contamination estimates from V19 is much lower for the last two redshift bins across all associated richness bins and is clearly inconsistent with the data. Estimating our model free cluster contamination in  $z$  - space rather than  $z - z_{\text{cl}}$  - space also did not lead to similar results as were found in V19, in an attempt to explain the contrast we see to the data.

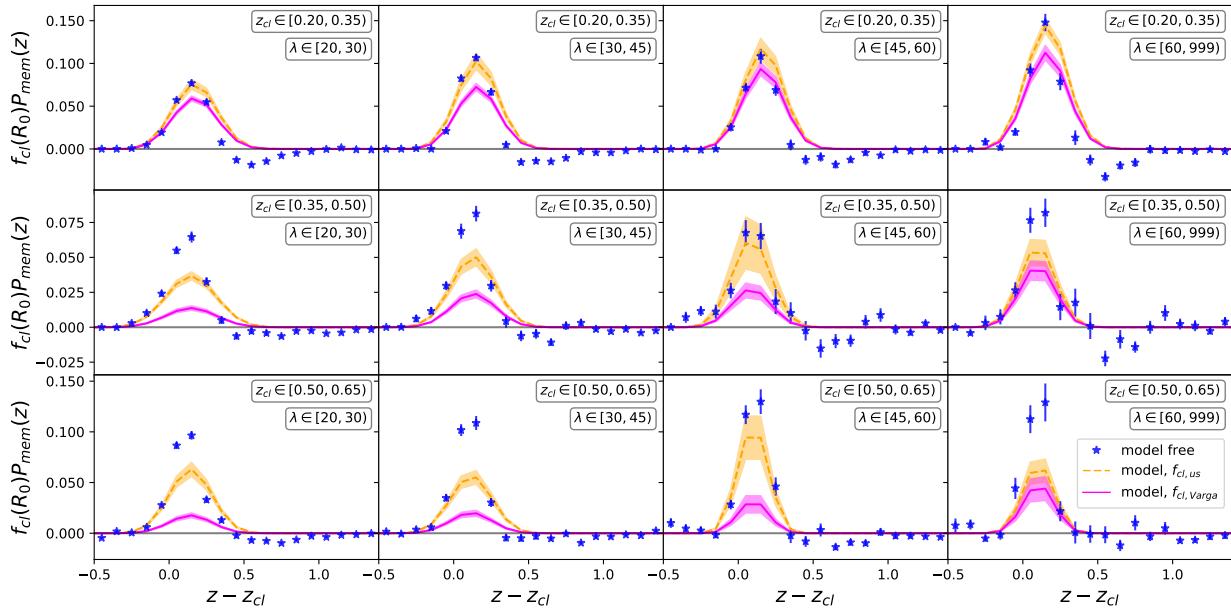


Figure 3.7: We show the field-subtracted redshift distributions (blue) along with the best fit models (orange) stacked in various cluster redshift - richness bins and compare it to the the model prediction when using the cluster contamination results from V19 and multiply them with our results for the cluster member distribution,  $P_{\text{mem}}$  in magenta. Here the results are presented for the first radial bin.

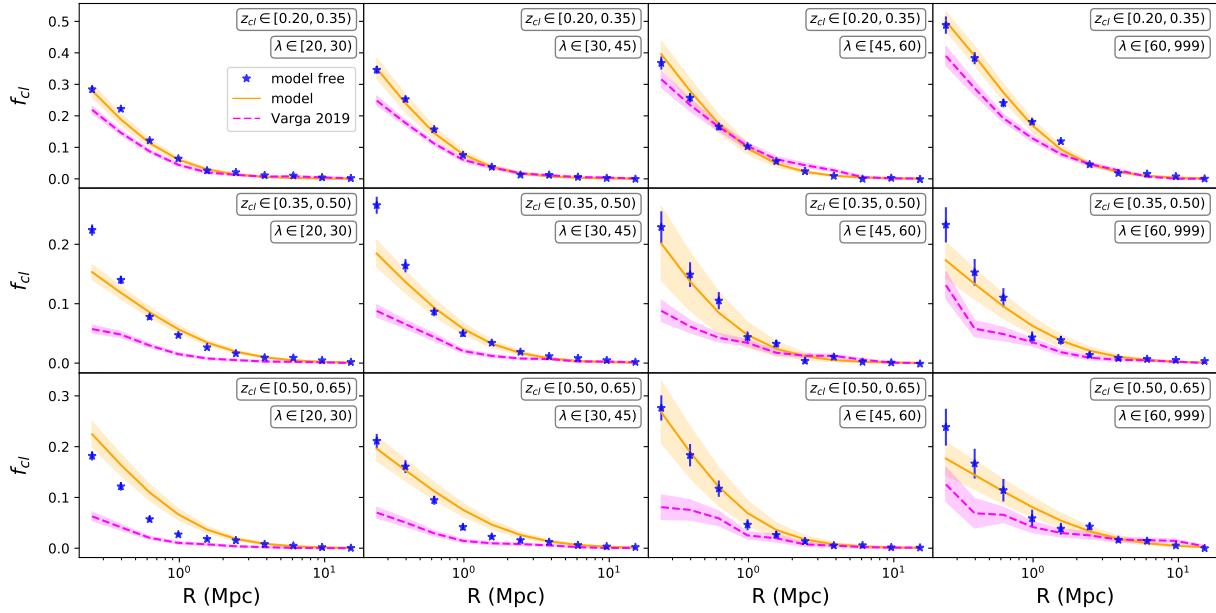


Figure 3.8: The best fit model for the cluster contamination  $f_{\text{cl}}$  is plotted in orange solid lines with  $1\sigma$  shaded regions as a function of radius for all assumed redshift - richness bins. Points with error bars represent the cluster contamination extracted from the field subtracted redshift distributions stacked for all clusters within the same binning. We compare these results to the cluster contamination estimates from V19, here shown as magenta dashed lines with  $1\sigma$  shaded regions. There is good agreement between the model and the directly measured contamination within the bins, but the cluster contamination estimates from V19 appear to be much lower for the last two redshift bins across all associated richness bins.

## 3.6 Impact on Cluster Mass

We are interested in the possible impact on the mean cluster mass estimates for each redshift-richness bin due to the different results of the cluster contamination fraction for most of these bins between V19 and this work. The idea is to create mock shear ( $\Delta\Sigma$ ) profiles for different masses, correct them for miscentering effects and by applying the two different cluster contamination estimates, find the best fit mass for each bin for each of the two cases and compare them.

### 3.6.1 Lensing signal

In the limit of weak gravitational lensing, we are interested in the reduced tangential shear component which is defined as  $g_T = \gamma_T / (1 - \kappa)$  where  $\gamma$  is the weak lensing shear and  $\kappa$  is the convergence. These are estimated from the shapes and alignments of the source galaxies through the ellipticity measurement  $\mathbf{e}$ , with the assumption that  $\langle \mathbf{e} \rangle \approx \langle \mathbf{g} \rangle$ . The shear signal can be related to the *excess surface mass density*  $\Delta\Sigma$ , by considering the mass distribution of the cluster:

$$\gamma_T(R) = \frac{\bar{\Sigma}(< R) - \bar{\Sigma}(R)}{\Sigma_{\text{crit}}} = \frac{\Delta\Sigma(R)}{\Sigma_{\text{crit}}}, \quad (3.10)$$

and the convergence is defined as

$$\kappa(R) = \bar{\Sigma}(R) / \Sigma_{\text{crit}}, \quad (3.11)$$

where  $\bar{\Sigma}(R)$  refers to the average surface mass density at  $R$

$$\bar{\Sigma}(R) = \Omega_m \rho_{\text{crit}} \int_{-\infty}^{+\infty} dz \xi_{\text{hm}}(r|M) \sqrt{R^2 + z^2}, \quad (3.12)$$

with  $R$  being the projected separation from the lens and  $z$  the separation along the line of sight.  $\rho_{\text{crit}}$  is the critical density of the Universe at a redshift of zero and  $\Omega_m$  is the mass density of the Universe.  $\bar{\Sigma}(< R)$  represents the average surface mass density within a projected radius  $R$ , given by

$$\bar{\Sigma}(< R) = \frac{2}{R^2} \int_0^R dR' R' \bar{\Sigma}(R'). \quad (3.13)$$

$\xi_{\text{hm}}(r|M)$  is the halo-matter correlation function at the halo redshift. While  $\xi_{\text{hm}}$  is dominated by the so called "1-halo" term at small scales, it is dominated by the "2-halo" term at large scales. To model  $\xi_{\text{hm}}$  we follow Zu et al. (2014) which takes the maximum between those two terms:

$$\xi_{\text{hm}}(r|M) = \max(\xi_{1h}(r|M), \xi_{2h}(r|M)). \quad (3.14)$$

For the 1-halo term we use an NFW profile  $\rho_{\text{NFW}}(r|M)$

$$\xi_{1h}(r|M, c) = \frac{\rho_{\text{NFW}}(r|M, c)}{\Omega_m \rho_{\text{crit}}} - 1, \quad (3.15)$$

where

$$\rho_{\text{NFW}}(r|M, c) = \frac{\Omega_m \rho_{\text{crit}} \delta_c}{\frac{r}{r_s} (1 + \frac{r}{r_s})^2}. \quad (3.16)$$

We assume a fixed concentration value of  $c_{200m} = 5$ . The 2-halo term  $\xi_{2h}(r|M)$  is the non-linear matter correlation function  $\xi_{\text{nl}}(r)$  scaled by the halo bias  $b(M)$  of Tinker et al. (2008)

$$\xi_{2h}(r|M) = b(M) \xi_{\text{nl}}(r), \quad (3.17)$$

where  $\xi_{\text{nl}}(r)$  is the 3D Fourier transform of the non linear power spectrum  $P_{\text{nl}}$  (Smith et al., 2003; Takahashi et al., 2012), given by

$$\xi_{\text{nl}}(r) = \frac{1}{2\pi^2} \int_0^\infty dk k^2 P_{\text{nl}} \frac{\sin(kr)}{kr}. \quad (3.18)$$

The power spectrum was computed using CAMB.

### 3.6.2 Miscentering correction

Often the centre of a cluster is not properly identified, leading the cluster to be “miscentered”. This means that the weak lensing signal will be suppressed in annuli around the cluster. Following McClintock et al. (2019) we model the recovered  $\Delta\Sigma$  profile as the weighted sum of two independent components:  $\Delta\Sigma_{\text{cen}}$  defines the contribution from properly centred clusters and  $\Delta\Sigma_{\text{mis}}$  from miscentered clusters, resulting in

$$\Delta\Sigma_{\text{model}} = (1 - f_{\text{mis}})\Delta\Sigma_{\text{cen}} + f_{\text{mis}}\Delta\Sigma_{\text{mis}}. \quad (3.19)$$

$f_{\text{mis}}$  here defines the miscentering fraction.  $\Delta\Sigma_{\text{cen}}$  is described in Equ (3.10) and  $\Delta\Sigma_{\text{mis}}$  should be estimated as

$$\Delta\Sigma_{\text{mis}}(R) = \bar{\Sigma}_{\text{mis}}(< R) - \bar{\Sigma}_{\text{mis}}(R). \quad (3.20)$$

If we assume a cluster to be miscentered by some radial offset  $R_{\text{mis}}$ , then the azimuthally averaged surface mass density is (e.g. Yang et al., 2006; Johnston et al., 2007)

$$\bar{\Sigma}_{\text{mis}}(R | R_{\text{mis}}) = \int_0^{2\pi} \frac{d\theta}{2\pi} \bar{\Sigma} \left( \sqrt{R^2 + R_{\text{mis}}^2 - 2RR_{\text{mis}} \cos \theta} \right). \quad (3.21)$$

Letting  $P(R_{\text{mis}})$  be the distribution of radial offsets for miscentered clusters, the corresponding mean miscentered profile  $\bar{\Sigma}_{\text{mis}}$  is

$$\bar{\Sigma}_{\text{mis}}(R) = \int dR_{\text{mis}} P(R_{\text{mis}}) \bar{\Sigma}_{\text{mis}}(R | R_{\text{mis}}), \quad (3.22)$$

where the miscentering distribution  $P(R_{\text{mis}})$  has the form

$$P(R | R_{\text{mis}}) = \frac{R}{R_{\text{mis}}^2} \exp \left( -\frac{R}{R_{\text{mis}}} \right). \quad (3.23)$$

The mean surface mass density inside the radius  $R$  is then

$$\bar{\Sigma}_{\text{mis}}(< R) = \frac{2}{R^2} \int_0^R dR' R' \bar{\Sigma}_{\text{mis}}(R') . \quad (3.24)$$

Rather than actually modelling the recovered  $\Delta\Sigma$  we adopt an already estimated value from the literature. McClintock et al. (2019) found the miscentering fraction  $f_{\text{mis}}$  to be roughly 0.25 and generally we adopt a fixed value for the radial offset,  $R_{\text{mis}} = 0.25$ .

### 3.6.3 Mean mass in redshift-richness bins

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

$$P(M, z | j) = C^{-1} P(j | M, z) \frac{dN}{dz}|_{M,z} , \quad (3.25)$$

where  $C^{-1}$  is given by the condition that  $\int dz \int dM P(M, z | j) = 1$  and  $dN/dz|_{M,z}$  denotes the halo mass function where we assume the fit by Tinker et al. (2008). Then

$$P(j | M, z) = \int d\ln\lambda P(j | \ln\lambda, z) P(\ln\lambda | M, z) , \quad (3.26)$$

where  $P(j | \ln\lambda, z) = 1$  for  $\hat{\lambda}_-^j < \hat{\lambda} < \hat{\lambda}_+^j$  and  $z_-^j < z < z_+^j$  and

$$P(\ln\lambda | M, z) = \mathcal{N}(\ln\lambda | \langle \ln\lambda \rangle(M, z), \sigma_\lambda^2) . \quad (3.27)$$

$\langle \ln\lambda \rangle(M, z)$  is a mass-richness relation defined as (Bleem et al., 2020)

$$\begin{aligned} \langle \ln\lambda \rangle(M, z) = \ln A_\lambda + B_\lambda \ln \left( \frac{M_{500,c}}{3 \cdot 10^{14} M_\odot h^{-1}} \right) + \\ C_\lambda \ln \left( \frac{E(z)}{E(z=0.6)} \right) . \end{aligned} \quad (3.28)$$

with  $A_\lambda = 76.9$ ,  $B_\lambda = 1.02$  and  $C_\lambda = 0.29$ .  $\sigma_\lambda^2 = \sigma_{\ln\lambda}^2 + (\bar{\lambda} - 1)/\bar{\lambda}^2$  where  $\sigma_{\ln\lambda} = 0.23$  and  $\bar{\lambda} = \exp(\langle \ln\lambda \rangle)$ .

Due to the fact that the shear profiles were computed assuming  $\Delta = 200m$  and  $P(M, z | j)$  was computed assuming  $\Delta = 500c$ , we multiply by the Jacobian matrix to transform between the coordinate systems

$$P(M_{200m}, z | j) = \left( \frac{dM_{200m}}{dM_{500c}} \right)^{-1} P(M_{500c}, z | j) . \quad (3.29)$$

Finally, this allows us to compute the weighted  $\Delta\Sigma$  signal for each bin  $j$

$$\Delta\Sigma(j, R) = \int dM \int dz \Delta\Sigma(M, z | j, R) P(M, z | j) . \quad (3.30)$$

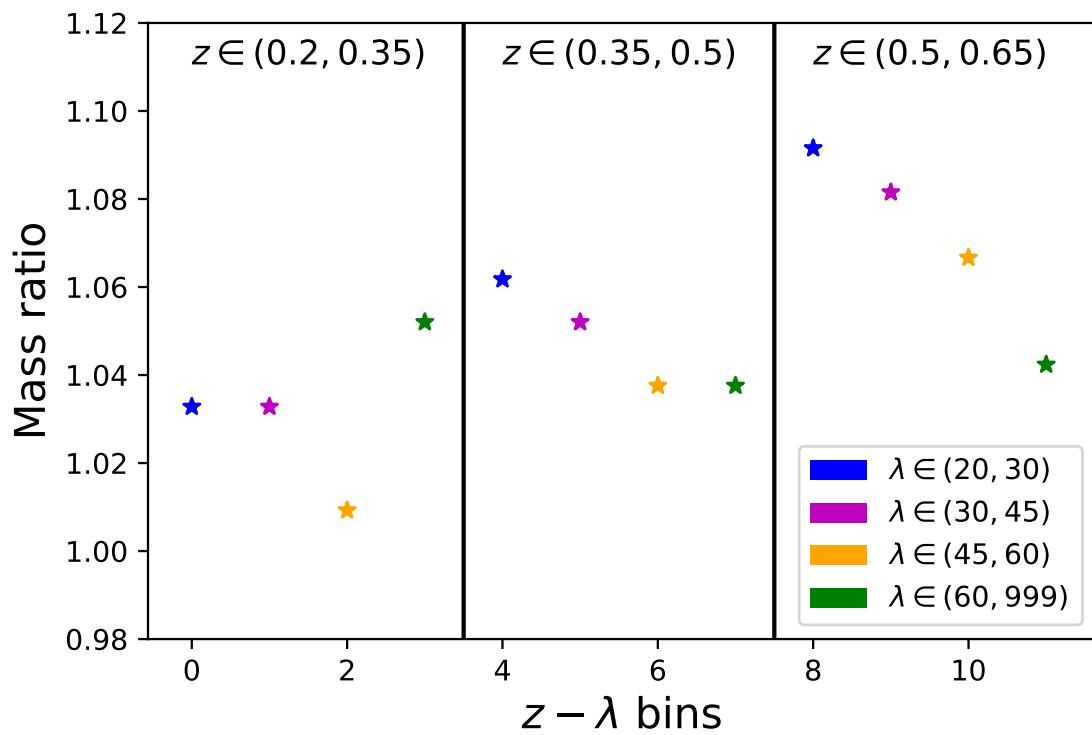


Figure 3.9: We show the ratio between the best fit masses for each redshift-richness bin when assuming the cluster contamination from V19 and that from this work and find, that on average, the masses differ by 5 – 6%.

### 3.6.4 Impact on cluster mass estimation

We now want to find the best fit mass estimates of each redshift-richness bin when considering the cluster contamination measurements from this work and from V19 and compare them, in order to quantify any differences. The cluster contamination is applied to the weighted mock profiles in the following way (see for example McClintock et al., 2019):

$$\Delta\Sigma_{f_{\text{cl}}}(j, R) = (1 - f_{\text{cl}}) \Delta\Sigma(j, R), \quad (3.31)$$

These mock profiles, with a cluster contamination applied, are compared to mean mock profiles, estimated on the mean redshift of the corresponding bins, for an array of masses. We apply the cluster contamination of this work to these profiles as we assume our cluster contamination to be the “true” one. This allows us to find the mass that has the smallest  $\chi^2$  with respect to the mock profile for each of the two cases, which then corresponds to the best fit mass of the respective bin.  $\Delta\Sigma$  profile estimates suffer from statistical uncertainties which include the large intrinsic variations between the shapes of source galaxies (shape noise) and line of sight structures that are not physically connected to the cluster (uncorrelated large scale structure, uLSS). Both of these effects constitute a more dominant source of uncertainty in lensing measurements and are therefore considered in our analysis. Their estimates are taken directly from the covariances measured in Melchior et al. (2017); McClintock et al. (2019). Figure 3.9 shows the ratio between the best fit masses when assuming the cluster contamination from V19 and when assuming  $f_{\text{cl}}$  from this work for each of the redshift-richness bins. We find the difference in mass to increase for the redshift bins and to decrease for richness within these bins. On average there is a deviation of roughly 5–6% between the masses. DES Collaboration (2020) compares the cluster masses predicted through cluster counts using a 3x2pt cosmology (DES Collaboration, 2018) to the masses measured using weak lensing McClintock et al. (2019). Remarkably, they found that the DES 3x2pt cosmology requires that the recovered weak lensing masses must be increased by  $\sim 30 - 40\%$  in the lowest richness bins in order to be consistent with number counts. But this would suggest that our results are not capable of reconciling the tension. We find a very weak richness dependence in our model and the difference found in cluster mass due to the deviating cluster contamination fraction results between this work and V19 is not high enough to close the gap between the very low weak lensing masses recovered at lower richness for weak lensing cosmological analysis and the masses recovered there for the DES 3x2pt cosmology. The systematic error budget on the amplitude of the mass-richness relation estimated in McClintock et al. (2019) for their analysis on the redMaPPer cluster sample with DES Y1 shows that the membership dilution in combination with miscentering contribute only 0.78% to the total error budget, in comparison to photo-z uncertainties which dominate with their 2.6% contribution, closely followed by cluster tri-axiality and line-of-sight projections, each contributing 2%. Therefore, the cluster member contamination correction is already a subdominant effect in the cluster mass estimation and the differences in the cluster contamination fraction results found here are not statistically significant enough to truly have an impact and can most likely not help to explain the tensions found in the cosmological analysis of DES Collaboration (2020).

## 3.7 Summary and Conclusions

We analyse the optically selected redMaPPer cluster sample (5750 clusters) (McClintock et al., 2019) in combination with the shape catalogue of the DES Y1 analysis (Zuntz et al., 2018; Hoyle et al., 2018) in order to build a cluster contamination model. As cluster galaxies can scatter into the background galaxy sample due to photo-z uncertainties, our approach relies on decomposing the redshift distribution  $P(z)$  of the shear galaxies into two parts: cluster member component and a local field component. The latter is extracted from an uncontaminated sky region near each cluster. The actual contamination is modelled as a Gaussian distribution in redshift space, whose amplitude varies with redshift, richness and distance from the cluster centre. The contamination of each cluster is fitted individually, combining the likelihoods across the full cluster sample to extract the best fit parameters and uncertainties of this model.

We compare our best-fit model for cluster contamination to a model independent estimate of the contamination (subsection 3.4.1) and find that our model is a good description of the data, showing that the contamination is well described by a Gaussian and varies in redshift. The redshift dependence of the contamination amplitude appears to follow a rather complex behaviour (see Figure 3.2), instead of a simple power law as one might have expected. We offer a short discussion on possible explanations as to why this trend exists in Section 3.4.2. This behaviour was previously observed in our analysis of the SPT and MARD-Y3 cluster samples (see Figure 3.3) (Paulus et al, in prep.).

We compare our results of the contamination fraction to V19 by adopting the same redshift-richness binning as well as a similar radial range and binning. In our validation plot we compare our best fit model in each bin with the model independent estimate of the contamination (data) and find them to be in good agreement but found the best fit model results of the cluster contamination fraction from V19 to divert from the data, especially for the higher cluster redshift bins across all associated richness bins. The cluster contamination fractions estimated in V19 seem to be significantly lower in these bins and are clearly not consistent with the data. This might indicate that these results are subject to some yet unidentified bias.

Finally, we quantify the impact on the cluster mass estimates within each redshift-richness bin due to the different results of the cluster member contamination. We find that the cluster masses differ on average by roughly 5 – 6%. DES Collaboration (2018) found that the DES 3x2pt cosmology requires that their recovered weak lensing masses must be increased by  $\sim 30 - 40\%$  in the lowest richness bins in order to be consistent with number counts. This suggests that our results are not capable of reconciling this tension. We find a very weak richness dependence in our model and the difference found in cluster mass due to the deviating cluster contamination fraction results between this work and V19 is not high enough to close the gap between the very low weak lensing masses recovered at lower richness for weak lensing cosmological analysis and the masses recovered there for the DES 3x2pt cosmology. Cluster member dilution is generally found to be a subdominant effect in the cluster mass estimation of McClintock et al. (2019). Nonetheless, the cluster contamination fractions found in V19 are inconsistent with the data and our own model

estimations and it should be understood where these differences are coming from.

# Chapter 4

## Conclusions and Outlook

The analyses presented in this thesis are motivated by the need to estimate galaxy cluster masses with high accuracy as their calibration via gravitational weak lensing is dominated by systematic uncertainties impacting precision cosmology. We focus on the so called cluster contamination correction, an important systematic uncertainty impacting the cluster mass estimation. In this final chapter we summarise our findings and conclusions of our analyses as well as give an outlook on future works.

In Chapter 2 we analyse the X-ray selected MARD-Y3 cluster sample, consisting of 208 clusters in the redshift range  $0.15 < z_{\text{cl}} < 0.9$  (Klein et al., 2019), and the SZE selected SPT cluster sample, consisting of 250 clusters in the redshift range  $0.15 < z_{\text{cl}} < 1.2$  (Bleem et al., 2015b), in combination with the shear galaxy catalogue extracted in the DES Y1 analysis (Zuntz et al., 2018; Hoyle et al., 2018) to build a cluster contamination model. The necessity to account for this contamination stems from the fact that cluster own galaxies can scatter into the background shear galaxy sample due to photo-z uncertainties. Therefore, our approach relies on decomposing the redshift distribution  $P(z)$  of the shear galaxies into two parts: a cluster member component and a true background component. The latter is extracted from an uncontaminated region of sky near each cluster, whereas the cluster contamination is modelled as a Gaussian in redshift space that is offset from the cluster redshift and whose amplitude varies with redshift, richness and distance from the cluster centre. In the following bullet points we want to highlight the most important aspects of this work and its findings:

- The  $P(z)$  decomposition method has previously been used by Stern et al. (2019) to aid the SPT weak lensing mass calibration over the DES Science Verification (SV) area. Now we are looking at a significantly larger SPT cluster sample as well as the greatly extended footprint of DES Y1 compared to DES SV. In addition, we are studying the X-ray selected cluster sample MARD-Y3 (Klein et al., 2019) for the first time over the DES Y1 footprint.
- The contamination is modelled for each cluster individually, as opposed to examining

the behaviour of large "stacks" of clusters in richness and redshift bins, allows for the dependencies of the contamination on the physical parameters of the cluster to be extracted for the full ensemble rather than estimated on average within binned sub-samples. Our method is motivated by the need to quantify the cluster contamination in rather small cluster samples of  $10^2$  to  $10^3$  systems where dividing into binned sub-samples is undesirable.

- For the first time we added a variable normalisation in cluster redshift to our model description of the cluster contamination, allowing us to gain deeper understanding of the redshift behaviour of cluster samples.
- We devised a method to estimate the contamination in a model independent way by subtracting a re-normalised local field distribution from the observed background distribution of the cluster. This allows us to test if our model is an accurate description of our data. We find that the cluster contamination is well described by a Gaussian, and that the amplitude of this contamination varies with redshift, richness and distance from the cluster centre in a manner that is described by our model.
- We found that the redshift dependence of the cluster contamination amplitude does not follow a simple power law trend but instead exhibits a rather complex behaviour, with both cluster samples exhibiting similar behaviour in this regard. Generally, this has not been observed before in other studies as most do not consider clusters on an individual basis, but rather work with stacks of clusters. We discussed in detail possible reasons for these redshift dependent features in the amplitude of the contamination.
- We quantified the impact that the cluster contamination constraints have on the systematic mass uncertainty. The systematic uncertainties were derived for three redshift snap shots as well as two different inner fitting radii. In each case the resulting systematic mass uncertainties were of the order of 1% .

In the first half of Chapter 3 we repeat the cluster contamination analysis on the optically selected redMaPPer cluster sample (5750 clusters in the redshift range  $0.2 < z_{\text{cl}} < 0.65$ ), whereas in the second half we focus on comparing our results to that of a similar work by adopting the same parameters, including a specific redshift - richness binning for the cluster sample. Again we want to summarise the most important aspects of this work in the following bullet points:

- The redMaPPer cluster sample is roughly  $10x$  larger than then SPT or MARD-Y3 samples mentioned above, reducing signal to noise in our measurements.
- We found once again that the redshift dependence of the contamination amplitude appears to follow a rather complex behaviour and shows a similar redshift trend as the previous two samples. This might indicate that this behaviour is something intrinsic to DES.

- We compared our results of the contamination fraction to V19 by adopting the same redshift-richness binning as well as a similar radial range and binning. We found discrepancies for the higher cluster redshift bins across all associated richness bins, where the cluster contamination fraction found in V19 seems to be significantly lower in these bins than what the data and our model prediction expects and is clearly inconsistent with the data.
- Finally, we quantified the impact on the cluster mass estimates within each redshift-richness bin due to the different results of the cluster member contamination. We found that the cluster masses differ on average by roughly 5–6%. DES Collaboration (2020) found that their recovered weak lensing masses must be increased by  $\sim 30\text{--}40\%$  in the lowest richness bins in order to be consistent with the masses found through number counts (DES Collaboration, 2018). This would suggest that our results are not capable of reconciling the tension. Our cluster contamination model shows only a very weak dependence on richness and the difference in cluster masses due to the deviating contamination fraction results in this work and V19 is not high enough to boost the very low weak lensing masses recovered at lower richness to the required amount. Nonetheless, the inconsistencies found between V19 and the data and our own model estimations of the contamination fraction should be understood and found where they are coming from.

In the near future, our measured cluster contamination will be employed to correct for the associated biases from cluster member contamination in the mass calibration and associated cosmological analysis for the SPT cluster sample (Bocquet et al., in prep.). The next generation weak lensing shear catalogue from DES Y3 has recently been published (Gatti et al., 2020), which means exciting new weak lensing studies are underway. In comparison to Y1, where the shear catalogue covered roughly  $1500 \text{ deg}^2$  of the DES footprint, the effective area of the survey now covers  $4143 \text{ deg}^2$  after masking for foreground objects and other problematic regions. This represents a significant increase in total area but with a similar depth. Due to slightly modified settings of the DES pipeline processing, the threshold for detection was lowered (Sevilla-Noarbe et al., 2020), resulting in an increase in the number of objects, more than expected from the increased area and depth alone. The DES Y3 weak lensing shape catalogue consists of approximately 100,000,000 galaxies resulting in a weighted source number density of  $n_{\text{eff}} = 5.59 \text{ gal/arcmin}^2$ . This is roughly 3 times more source galaxies than what was found in Y1. All these improvements will aid the cluster mass calibration via gravitational weak lensing and the subsequent cosmological constraining power. Systematic uncertainties can be understood even better and to higher precision. In terms of the cluster contamination correction, the larger the covered area of a wide field imaging survey, the larger the cluster sample that can be used to extract information from, increasing its statistical power. As background source distributions are extracted in radial bins from the cluster centre, a large source number density is beneficial to constrain the data equally in every bin. Especially high redshift clusters are limited by the lack of background source galaxies at even higher redshifts. Deeper survey are

required to provide these. So far, Y3 has not increased its depth in comparison to Y1, but this should change with Y5. This is really important as so far the cluster contamination model is limited by the noise present at higher cluster redshifts, making it difficult to properly constrain the data. With the launch of Euclid in 2022, these current issues will be a problem of the past as it is expected to observe at a depth of roughly 30 galaxies per arcmin<sup>2</sup> (roughly 3 times more than DES Y5).

# Appendix A

## Appendix to Chapter 2

### A.1 Mock catalogue

We want to validate that our analysis pipeline is able to recover mock input values for the parameters of our contamination fraction model. In order to do this a mock catalogue was created by first choosing certain input parameters and then drawing Poisson realisations of the contaminated redshift distributions for each cluster in our sample. The mock has a 10 times larger sample size than the cluster sample in order to achieve a significantly smaller statistical uncertainty than on our real data. We then analyse those mock catalogues, recovering the input parameters with good accuracy. Figure A.1 shows the results for the combined SPT + MARD-Y3 cluster sample.

### A.2 Model-free estimation of the cluster contamination for different binning

We compare the model prediction of the cluster contamination with a model-independent way of determining the cluster contamination. The latter is done by subtracting a renormalised local field distribution,  $P_{\text{field}}(z)$ , from the observed background distribution of the cluster,  $P_{\text{meas}}(z, R)$ . These distributions produced within bins of redshift or richness, stacked in  $z - z_{\text{cl}}$  space, are used both for validation of the model we fit and to extract an average cluster contamination for all the clusters in a particular redshift or richness bin by simply integrating over the excess associated with the cluster contamination. We show example field-subtracted redshift distributions (blue) along with the best fit models (orange) stacked in various cluster redshift bins for the first radial bin (Figure A.2), the second radial bin (Figure A.3), the third radial bin (Figure A.4) and the fourth radial bin (Figure A.5). Furthermore we show example field-subtracted redshift distributions along with the best-fit models stacked in various richness bins for the first radial bin (Figure A.6), the second radial bin (Figure A.7), the third radial bin (Figure A.8) and the fourth radial bin (Figure A.9). All plots are shown for the SPT + MARD-Y3 cluster sample.

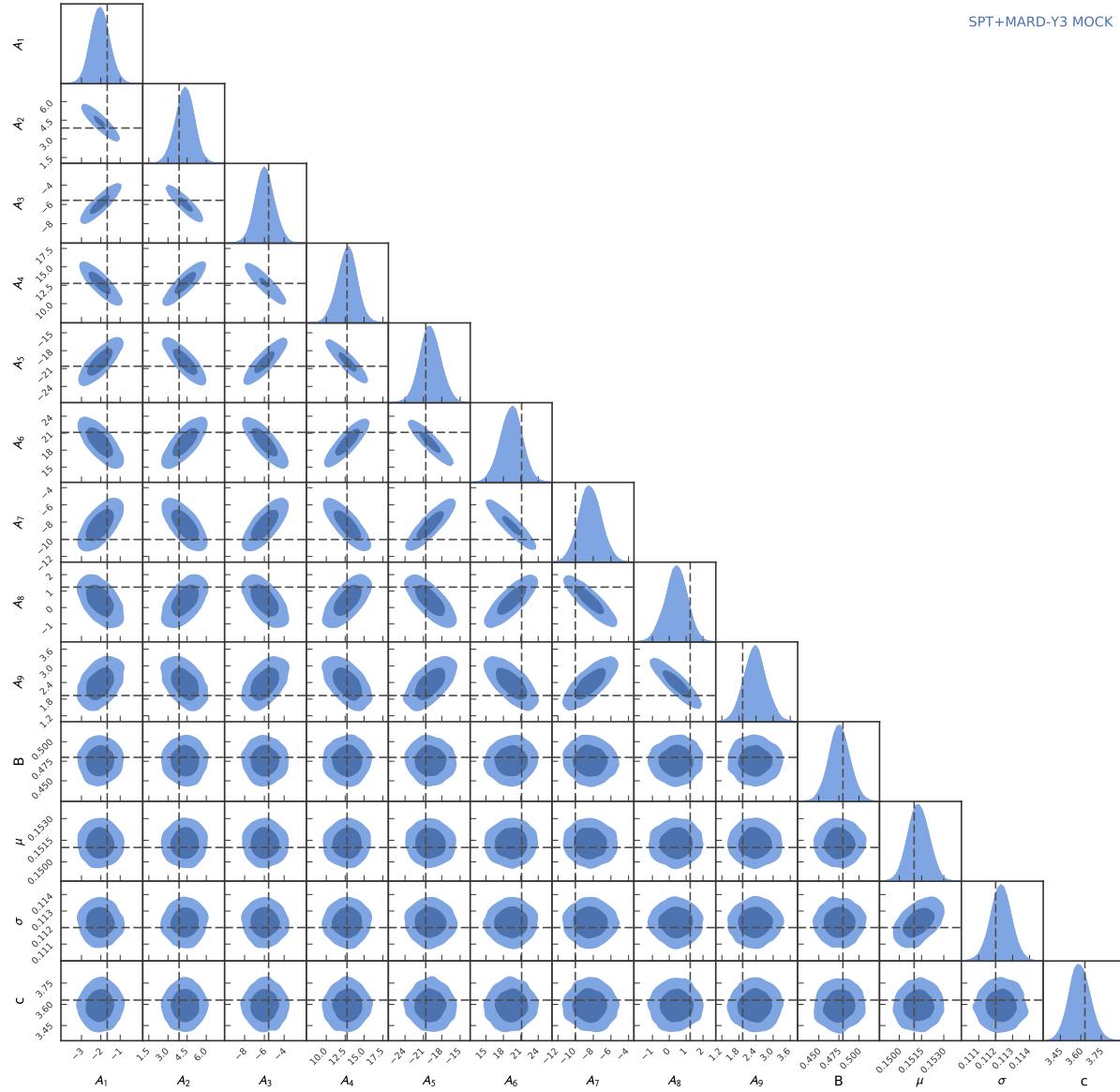


Figure A.1: Mock simulation for the combined SPT + MARD-Y3 cluster sample. The intersection of the dotted lines indicate the assumed value in the mock, that we wish to recover. As the contours indicate, all input parameters are recovered.

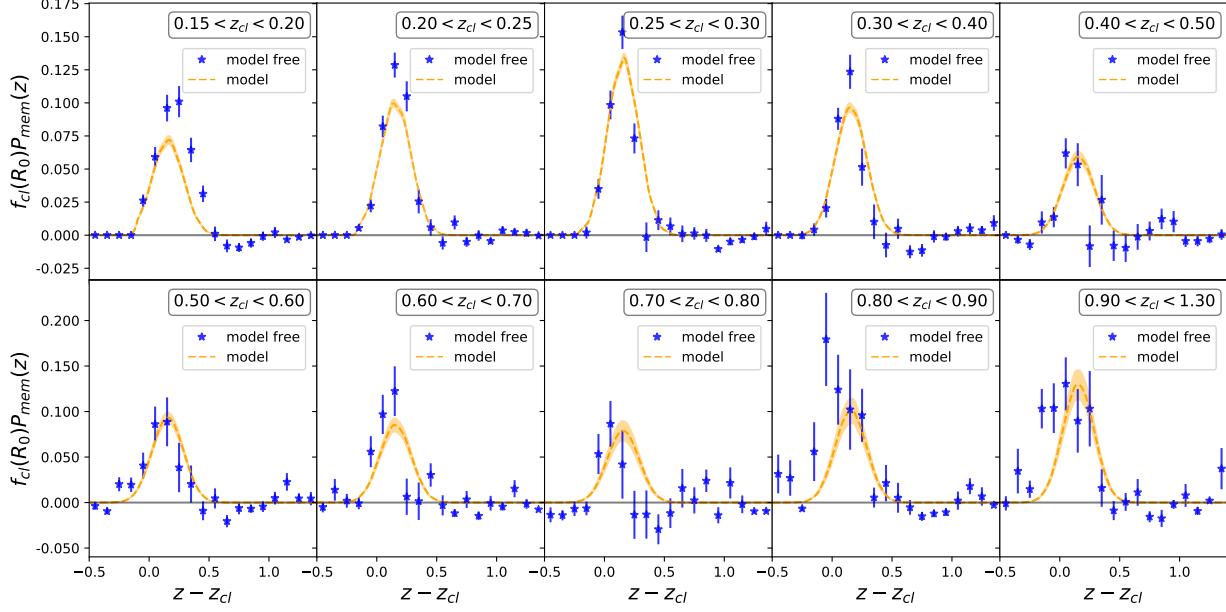


Figure A.2: The field-subtracted redshift distribution stacked in various cluster redshift bins in blue along with the average best-fit model in orange. We show the results for the first radial bin (0.25-0.38 Mpc). The complex redshift evolution behaviour is clearly visible. The results are shown for the SPT + MARD-Y3 cluster sample.

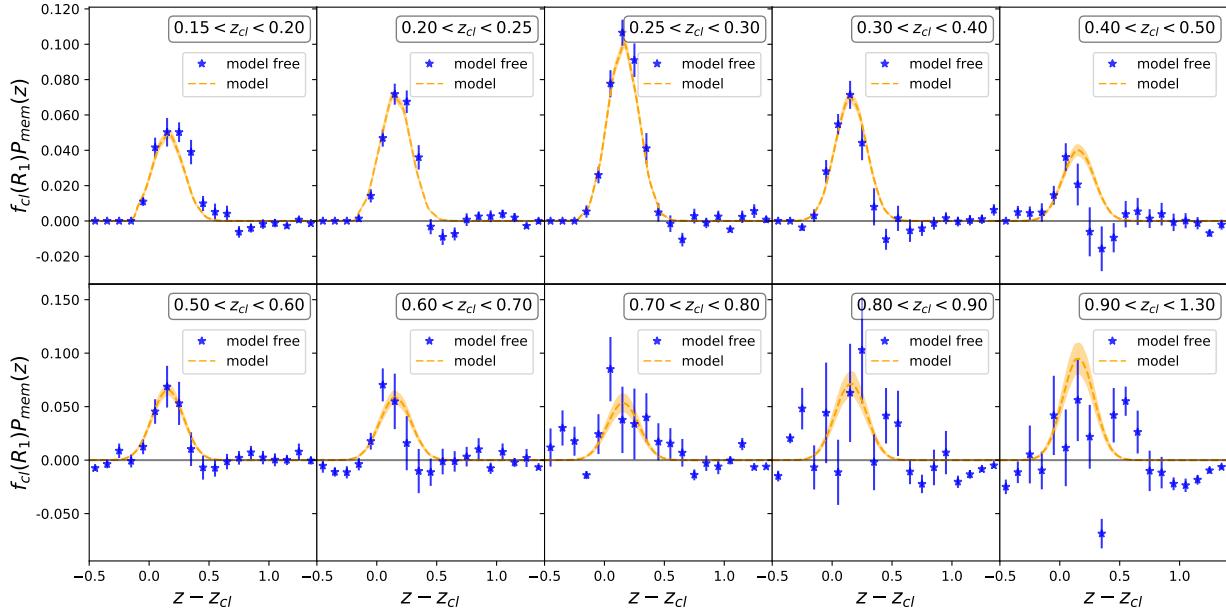


Figure A.3: Same as Figure A.2 but for the second radial bin (0.38-0.57 Mpc).

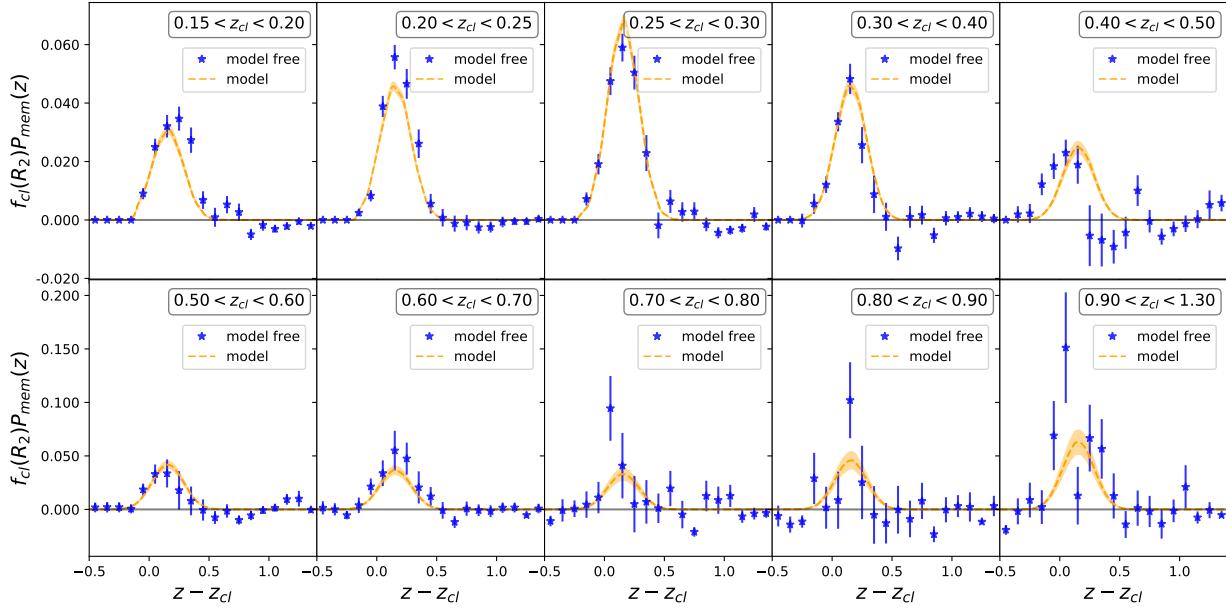


Figure A.4: Same as Figure A.2 but for the third radial bin (0.57-0.87 Mpc).

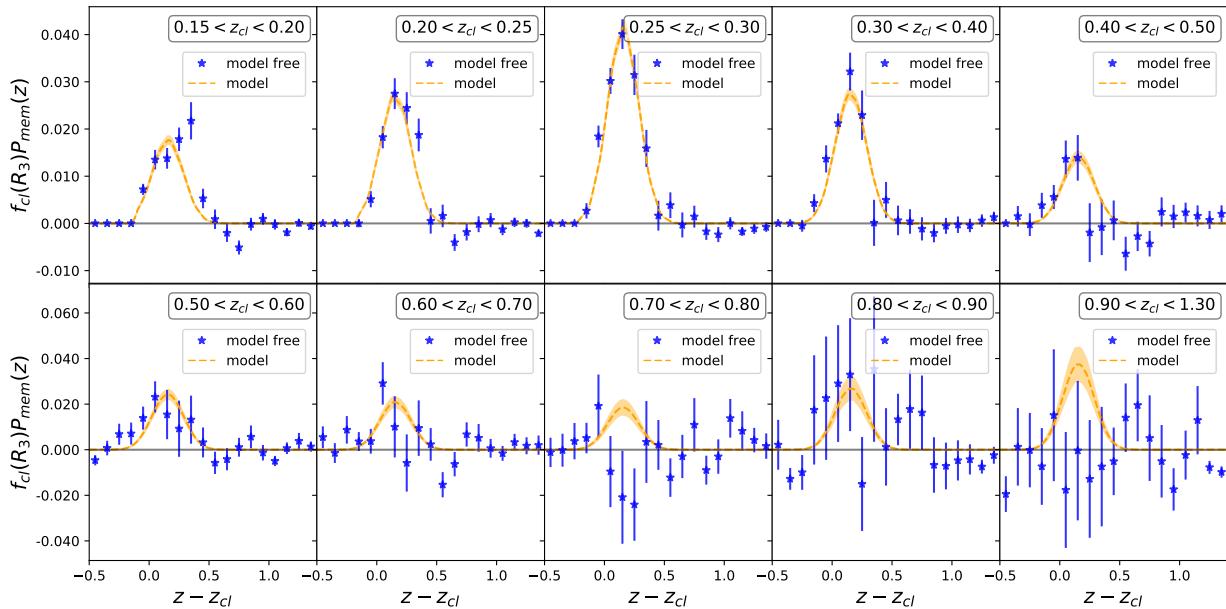


Figure A.5: Same as Figure A.2 but for the fourth radial bin (0.87-1.31 Mpc).

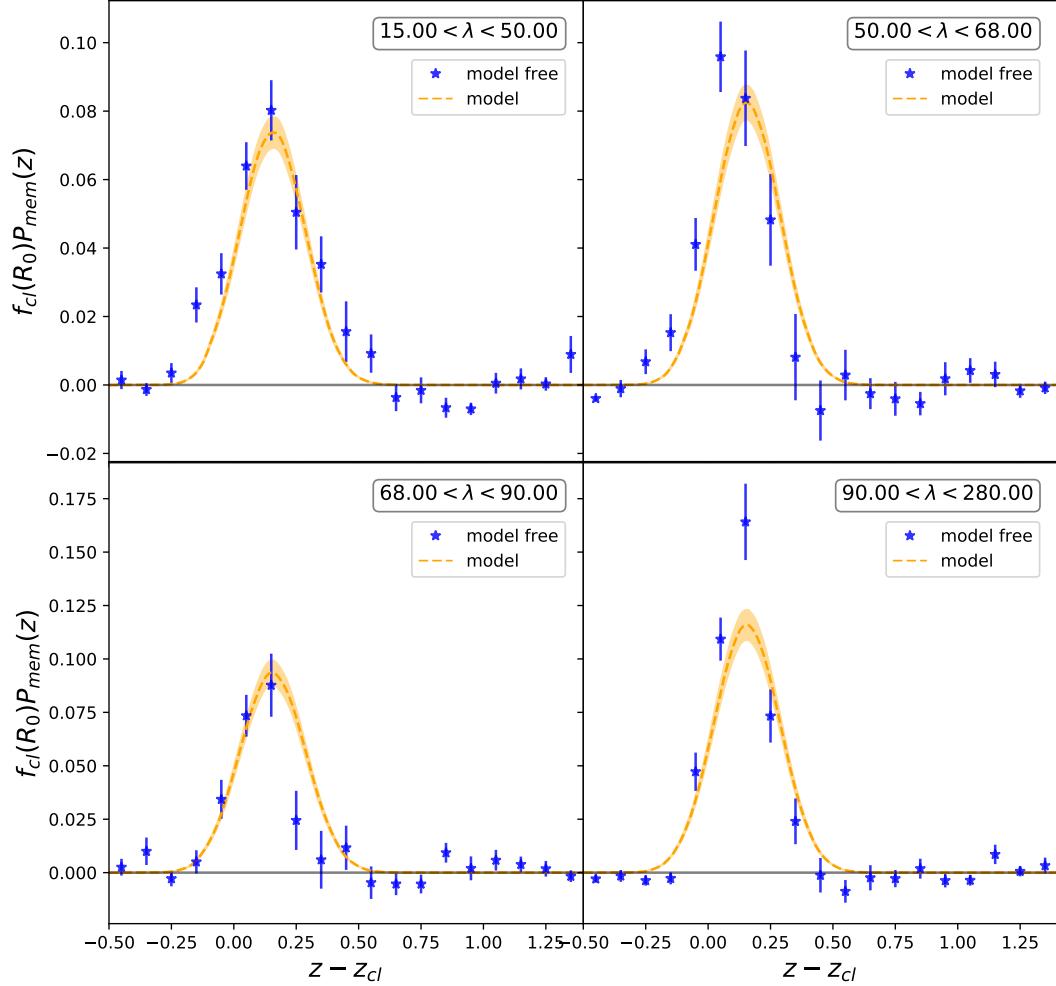


Figure A.6: The field-subtracted redshift distribution stacked in various richness bins in blue along with the best fit average model in orange. Shown here are the results for the first radial bin (0.25-0.38 Mpc). A modest increase in amplitude is visible with increasing richness. The results are shown for the SPT + MARD-Y3 cluster sample.

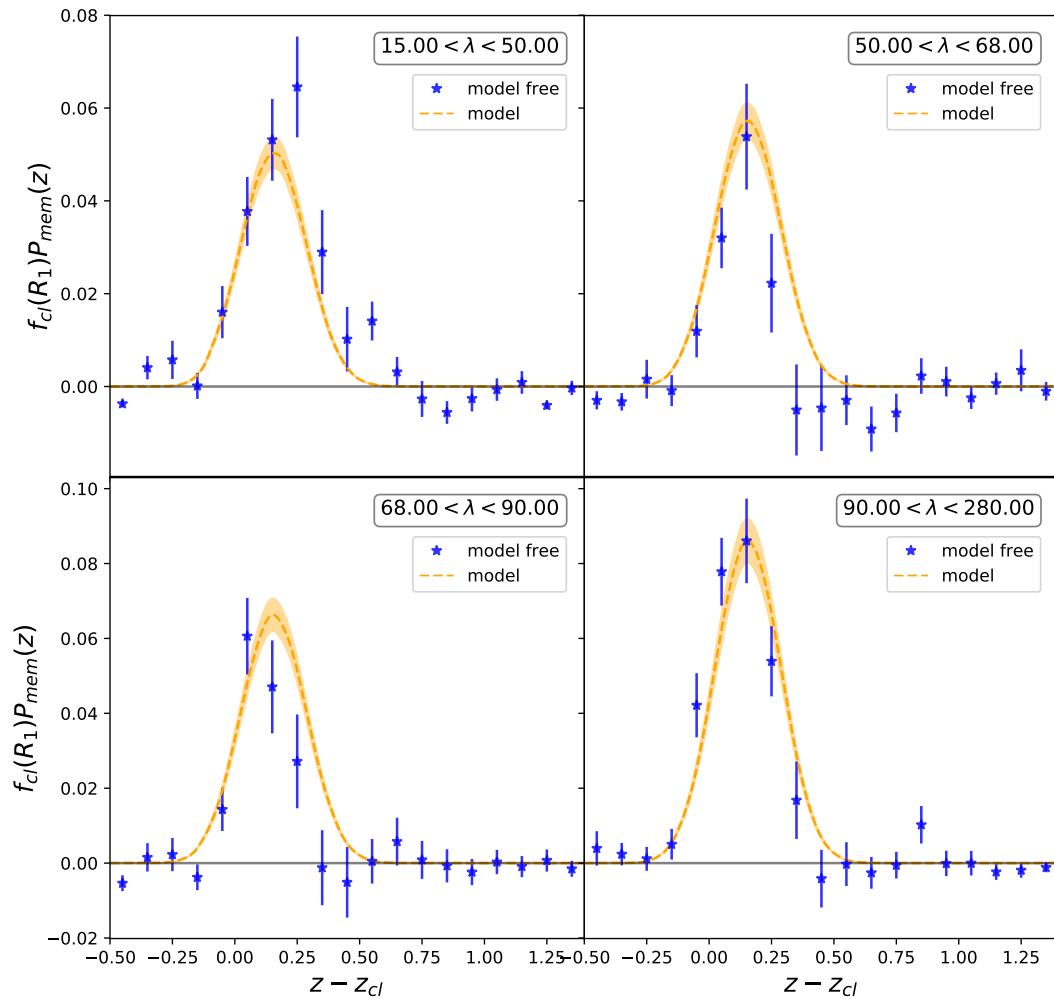


Figure A.7: Same as Figure A.6 but for the second radial bin (0.38-0.57 Mpc).

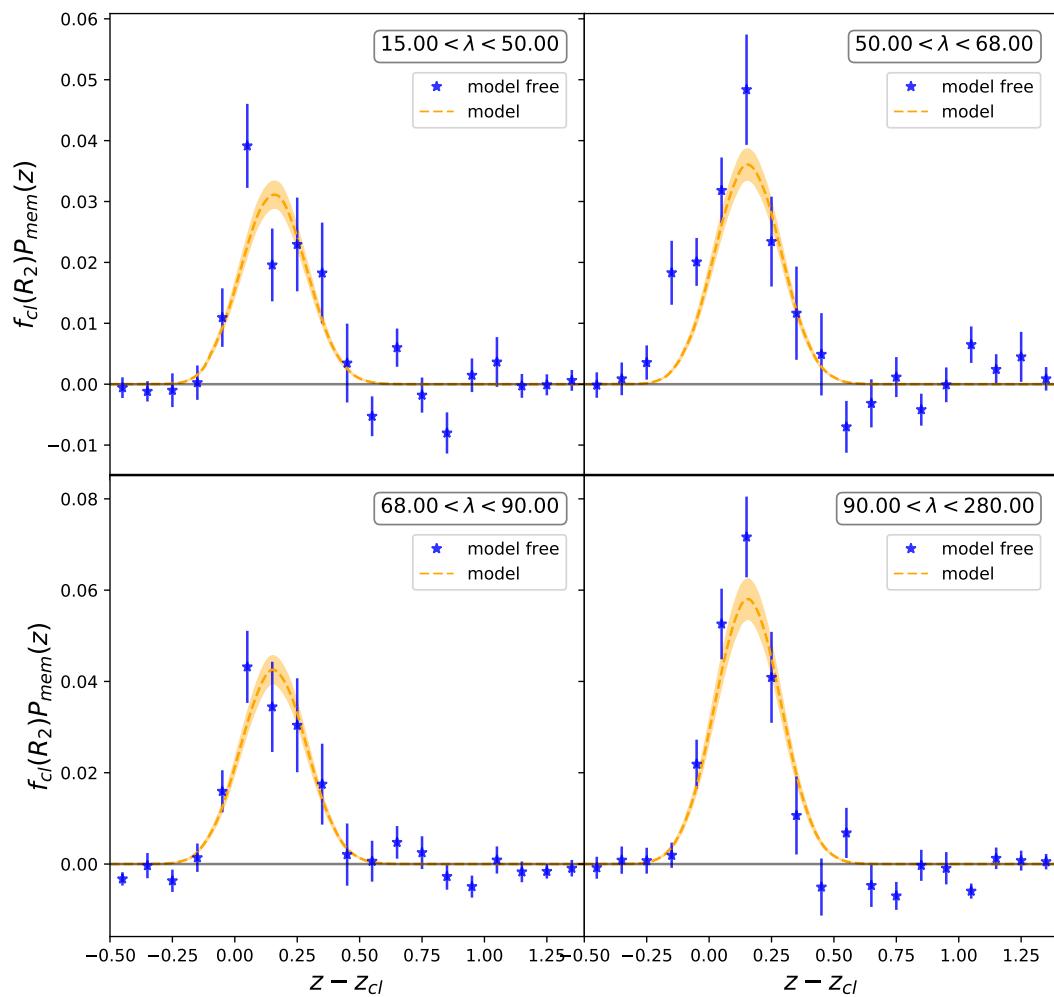


Figure A.8: Same as Figure A.6 but for the third radial bin (0.57-0.87 Mpc).

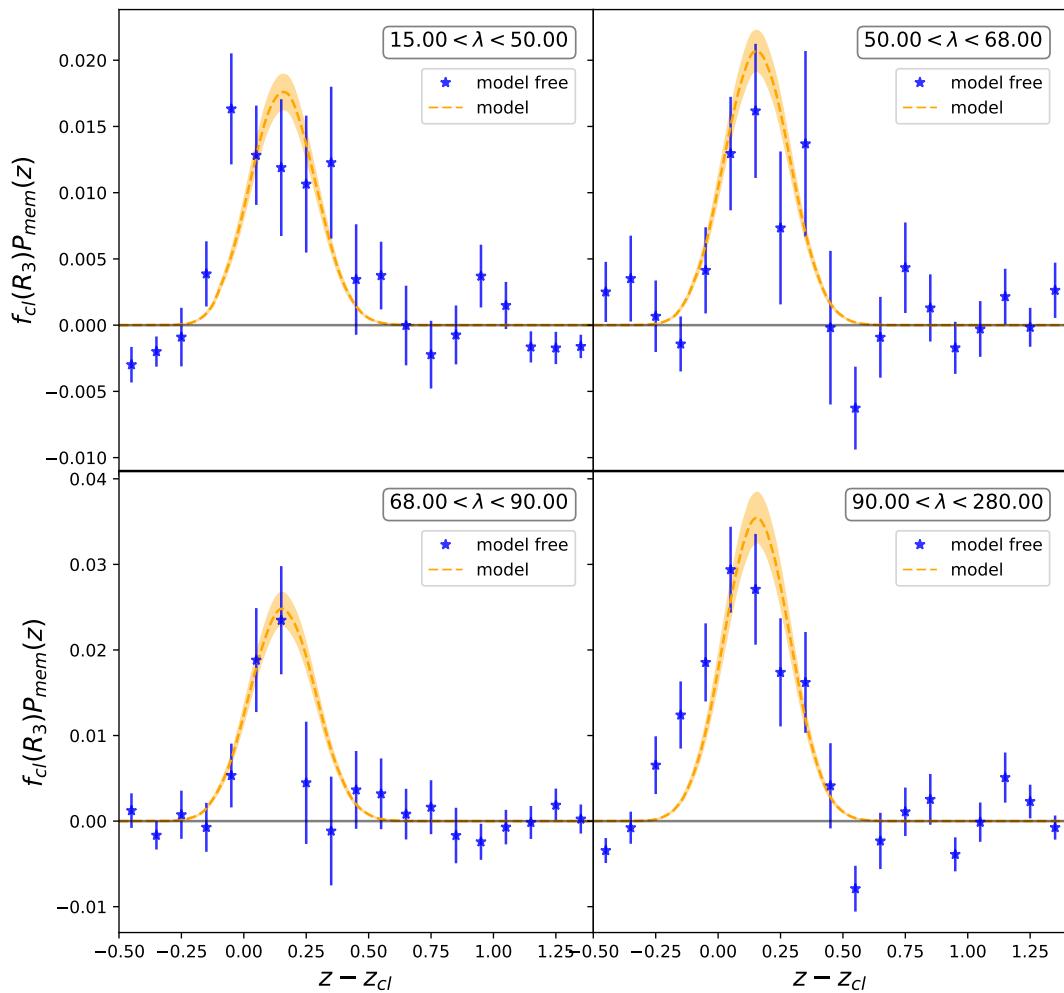


Figure A.9: Same as Figure A.6 but for the fourth radial bin (0.87-1.31 Mpc).

# Bibliography

- Abazajian K., et al., 2003, AJ, 126, 2081
- Abazajian K. N., Adelman-McCarthy J. K., Agüeros M. A., Allam S. S., Prieto C. A., An D., et al. 2009, The Astrophysical Journal Supplement Series, 182, 543–558
- Abbott T. M. C., et al., 2018, The Astrophysical Journal Supplement Series, 239, 18
- Aihara H., et al., 2011, ApJS, 193, 29
- Alcock C., Allsman R., Alves D. a., 1997, The Astrophysical Journal, 486, 697
- Allen S. W., Schmidt R. W., Fabian A. C., 2002, MNRAS, 334, L11–L15
- Allen S. W., Evrard A. E., Mantz A. B., 2011, ARA&A, 49, 409
- Annis J., Soares-Santos M., Strauss M., more 2014, The Astrophysical Journal, 794, 120
- Applegate D. E., et al., 2014, MNRAS, 439, 48
- Bartelmann M., 2010, Classical and Quantum Gravity, 27, 233001
- Bartelmann M., Schneider P., 2001, Phys. Rep., 340, 291
- Battaglia N., et al., 2016, J. Cosmology Astropart. Phys., 8, 013
- Bellagamba F., Roncarelli M., Maturi M., Moscardini L., 2017, MNRAS, 473, 5221
- Benítez N., 2000, ApJ, 536, 571
- Bertin E., Arnouts S., 1996, A&A, supplement, 117, 393
- Bleem L. E., et al., 2015a, ApJS, 216, 27
- Bleem L. E., et al., 2015b, ApJS, 216, 27
- Bleem L. E., Bocquet S., Stalder B., Gladders M. D., more 2020, The Astrophysical Journal Supplement Series, 247, 25
- Bocquet S., Carter F. W., 2016, The Journal of Open Source Software, 1
- Bocquet S., et al., 2015, ApJ, 799, 214
- Bocquet S., et al., 2019, The Astrophysical Journal, 878, 55
- Bocquet S., et al., 2020, in prep.
- Böhringer H., et al., 2004, A&A, 425, 367
- Boller T., Freyberg M. J., Trümper J., Haberl F., Voges W., Nandra K., 2016, A&A, 588, A103
- Burke D. L., Rykoff E. R., more 2018, The Astronomical Journal, 155, 41

- Böhringer H., Werner N., 2010, *The Astronomy and Astrophysics Review*, 18, 127
- Carlstrom J. E., et al., 2011a, *PASP*, 123, 568
- Carlstrom J. E., et al., 2011b, *PASP*, 123, 568
- Chang C., et al., 2018, *ApJ*, 864, 83
- Coe D., Benítez N., Sánchez S. F., Jee M., Bouwens R., Ford H., 2006, *AJ*, 132, 926
- Cohn J. D., Evrard A. E., White M., Croton D., Ellingson E., 2007, *MNRAS*, 382, 1738
- Costanzi M., et al., 2018, *MNRAS*, 482, 490–505
- Costanzi M., et al., 2019, *MNRAS*, 488, 4779
- DES Collaboration 2005, preprint, ([arXiv:astro-ph/0510346](https://arxiv.org/abs/astro-ph/0510346))
- DES Collaboration 2016, *MNRAS*, 460, 1270
- DES Collaboration 2018, *Phys. Rev. D*, 98, 043526
- DES Collaboration 2020, *Phys. Rev. D*, 102, 023509
- Dar A., 1995, *The Astrophysical Journal*, 449, 550
- Dietrich J. P., et al., 2018, *MNRAS*, 483, 2871–2906
- Dietrich J. P., et al., 2019, *MNRAS*, 483, 2871
- Dodelson S., Heitmann K., Hirata C., Honscheid K., Roodman A., Seljak U., Slosar A., Trodden M., 2016, preprint, ([arXiv:1604.07626](https://arxiv.org/abs/1604.07626))
- Drlica-Wagner A., et al., 2017, preprint, ([arXiv:1708.01531](https://arxiv.org/abs/1708.01531))
- Drlica-Wagner A., Sevilla-Noarbe I., Rykoff E. S., more 2018, *The Astrophysical Journal Supplement Series*, 235, 33
- Du N., Force N., Khatiwada R., et al. 2018, *Physical Review Letters*, 120
- Ebeling H., et al., 2013, *MNRAS*, 432, 62
- Edge A., Sutherland W., Kuijken K., Driver S., McMahon R., Eales S., Emerson J. P., 2013, *The Messenger*, 154, 32
- Farahi A., Evrard A. E., Rozo E., Rykoff E. S., Wechsler R. H., 2016, *MNRAS*, 460, 3900
- Fields B. D., Freese K., Graff D. S., 1998, *New Astronomy*, 3, 347–361
- Flaugher B., et al., 2015, *AJ*, 150, 150
- Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, *PASP*, 125, 306
- Freese K., Fields B., Graff D., 2003, *The First Stars*, p. 18–23
- Gatti M., et al., 2020, preprint ([arXiv:2011.03408](https://arxiv.org/abs/2011.03408))
- Gladders M. D., Yee H. K. C., 2000, *AJ*, 120, 2148
- Grandis S., Bocquet S., Mohr J. J., Klein M., Dolag K., 2021, arXiv e-prints, p. [arXiv:2103.16212](https://arxiv.org/abs/2103.16212)
- Gruen D., Brimioule F., 2017, *MNRAS*, 468, 769
- Gruen D., et al., 2014, *MNRAS*, 442, 1507
- Gunn J. E., et al., 2006, *AJ*, 131, 2332

- Haiman Z., Mohr J. J., Holder G. P., 2001, *ApJ*, 553, 545–561
- Hasselfield M., et al., 2013, *Journal of Cosmology and Astroparticle Physics*, 2013, 008–008
- Hennig C., et al., 2017, *MNRAS*, 467, 4015
- Henry J. P., Mullis C. R., Voges W., Böhringer H., Briel U. G., Gioia I. M., Huchra J. P., 2006, *ApJS*, 162, 304
- Hildebrandt H., et al., 2017, *MNRAS*, 465, 1454
- Hilton M., Hasselfield M., Sifón C., Battaglia N., et al., 2018, *The Astrophysical Journal Supplement Series*, 235, 20
- Hinshaw G., Larson D., Komatsu E. a., 2013, *The Astrophysical Journal Supplement Series*, 208, 19
- Hoekstra H., Herbonnet R., Muzzin A., Babul A., Mahdavi A., Viola M., Cacciato M., 2015, *MNRAS*, 449, 685
- Hoyle B., et al., 2018, *MNRAS*, 478, 592
- Huff E., Mandelbaum R., 2017, preprint, ([arXiv:1702.02600](https://arxiv.org/abs/1702.02600))
- Ivezić Ž., et al., 2019, *ApJ*, 873, 111
- Johnston D. E., et al., 2007, preprint, ([arXiv:0709.1159](https://arxiv.org/abs/0709.1159))
- King I., 1962, *Astronomical Journal*, 67, 471
- Klasen M., Pohl M., Sigl G., 2015, *Progress in Particle and Nuclear Physics*, 85, 1–32
- Klein M., et al., 2018, *MNRAS*, 474, 3324
- Klein M., et al., 2019, *MNRAS*, 488, 739
- Koester B. P., et al., 2007, *ApJ*, 660, 221
- Kravtsov A. V., Borgani S., 2012, *ARA&A*, 50, 353
- Kuijken K., 2011, *The Messenger*, 146, 8
- Kuijken K., et al., 2019, *Astronomy & Astrophysics*, 625, A2
- Laureijs R., et al., 2011, preprint ([arXiv:1110.3193](https://arxiv.org/abs/1110.3193))
- Leauthaud A., et al., 2017, *MNRAS*, 467, 3024
- Li R., et al., 2016, *MNRAS*, 458, 2573–2583
- Liu J., et al., 2015, *MNRAS*, 449, 3370
- Malz A. I., Marshall P. J., DeRose J., Graham M. L., Schmidt S. J., and R. W., 2018, *The Astronomical Journal*, 156, 35
- Mantz A., Allen S. W., Rapetti D., Ebeling H., 2010, *MNRAS*, 406, 1759
- Mantz A. B., et al., 2015, *MNRAS*, 446, 2205
- McClintock T., et al., 2019, *MNRAS*, 482, 1352
- Medezinski E., et al., 2018a, *ASJ*, 70, 30
- Medezinski E., et al., 2018b, *ASJ*, 70, S28
- Melchior P., et al., 2017, *MNRAS*, 469, 4899

- Miyatake H., et al., 2019, *The Astrophysical Journal*, 875, 63
- Moore B., Katz N., Lake G., 1996, *The Astrophysical Journal*, 457, 455
- Motl P. M., Hallman E. J., Burns J. O., Norman M. L., 2005, *ApJ*, 623, L63
- Murata R., Nishimichi T., Takada M., Miyatake H., Shirasaki M., More S., Takahashi R., Osato K., 2018, *American Astronomical Society*, 854, 120
- Narayan R., Bartelmann M., 1997, preprint ([arXiv:astro-ph/9606001](https://arxiv.org/abs/astro-ph/9606001))
- Navarro J. F., Frenk C. S., White S. D. M., 1996, *ApJ*, 462, 563
- Navarro J. F., Frenk C. S., White S. D. M., 1997, *ApJ*, 490, 493
- Oguri M., 2014, *MNRAS*, 444, 147
- Okabe N., Smith G. P., 2016, *MNRAS*, 461, 3794
- Olive K. A., 1997, preprint, pp 60–126 ([arXiv:astro-ph/9707212](https://arxiv.org/abs/astro-ph/9707212))
- Pereira M. E. S., et al., 2020, *MNRAS*, 498, 5450–5467
- Planck 2020, *Astronomy & Astrophysics*, 641, A6
- Planck Collaboration et al., 2016b, *Astronomy & Astrophysics*, 594, A27
- Planck Collaboration Ade P. A. R., Aghanim N., Arnaud M., Ashdown M., Aumont et al. 2016a, *A&A*, 594, A27
- Prat J., et al., 2018, *Phys. Rev. D*, 98, 042005
- Pratt G. W., Arnaud M., Biviano A., Eckert D., Ettori S., Nagai D., Okabe N., Reiprich T. H., 2019, *Space Science Reviews*, 215
- Predehl P., et al., 2021, *Astronomy & Astrophysics*, 647, A1
- Roszkowski L., Sessolo E. M., Trojanowski S., 2018, *Reports on Progress in Physics*, 81, 066201
- Rozo E., et al., 2009, *ApJ*, 699, 768
- Rykoff E. S., et al., 2012, *ApJ*, 746, 178
- Rykoff E. S., et al., 2014a, *ApJ*, 785, 104
- Rykoff E. S., et al., 2014b, *ApJ*, 785, 104
- Rykoff E. S., et al., 2016a, *ApJS*, 224, 1
- Rykoff E. S., et al., 2016b, *ApJS*, 224, 1
- Sarazin C. L., 1986, *Rev. Mod. Phys.*, 58, 1
- Sarkar S., 1996, preprint, pp 235–249 ([arXiv:astro-ph/9611232](https://arxiv.org/abs/astro-ph/9611232))
- Saro A., et al., 2015, *MNRAS*, 454, 2305
- Sevilla-Noarbe I., et al., 2020, preprint ([arXiv:2011.03407](https://arxiv.org/abs/2011.03407))
- Sheldon E., 2015, NGMIX: Gaussian mixture models for 2D images, *Astrophysics Source Code Library* (ascl:1508.008)
- Sheldon E. S., Huff E. M., 2017, *ApJ*, 841, 24
- Sheldon E., Annis J., Bohringer H., other 2001, *ApJ*, 554, 881

- Sheldon E. S., et al., 2004, AJ, 127, 2544
- Simet M., McClintock T., Mandelbaum R., Rozo E., Rykoff E., Sheldon E., Wechsler R. H., 2017, MNRAS, 466, 3103
- Smith R. E., et al., 2003, MNRAS, 341, 1311
- Song J., et al., 2012, ApJ, 761, 22
- Staniszewski Z., et al., 2009, ApJ, 701, 32
- Stern C., et al., 2019, MNRAS, 485, 69
- Strazzullo V., et al., 2019, A&A, 622, A117
- Sunyaev R. A., Zeldovich Y. B., 1970, Comments on Astrophysics and Space Physics, 2, 66
- Sunyaev R. A., Zel'dovich Y. B., 1972a, Comments on Astrophysics and Space Physics, 4, 173
- Sunyaev R. A., Zeldovich Y. B., 1972b, Comments on Astrophysics and Space Physics, 4, 173
- Takahashi R., Sato M., Nishimichi T., Taruya A., Oguri M., 2012, ApJ, 761, 152
- Tinker J., Kravtsov A. V., Klypin A., Abazajian K., Warren M., Yepes G., Gottlöber S., Holz D. E., 2008, ApJ, 688, 709
- Troxel M. A., et al., 2018, Physical Review D, 98
- Varga T. N., et al., 2019, MNRAS, 489, 2511
- Vikhlinin A., McNamara B. R., Forman W., Jones C., Quintana H., Hornstrup A., 1998, ApJ, 502, 558
- Voges W., et al., 1999, A&A, 349, 389
- Voges W., et al., 2000, VizieR Online Data Catalog, 9029, 0
- White S. D. M., Navarro J. F., Evrard A. E., Frenk C. S., 1993, Nature, 366, 429
- Yang X., Mo H. J., van den Bosch F. C., Jing Y. P., Weinmann S. M., Meneghetti M., 2006, MNRAS, 373, 1159
- Yanny B., et al., 2009, The Astronomical Journal, 137, 4377–4399
- York D. G., et al., 2000, AJ, 120, 1579
- Zu Y., Weinberg D. H., Rozo E., Sheldon E. S., Tinker J. L., Becker M. R., 2014, MNRAS, 439, 1628
- Zu Y., Mandelbaum R., Simet M., Rozo E., Rykoff E. S., 2017, MNRAS, 470, 551–560
- Zuntz J., et al., 2018, MNRAS, 481, 1149
- Zwicky F., 1933, Helvetica Physica Acta, 6, 110
- de Jong J. T. A., et al., 2017, A&A, 604, A134
- von der Linden A., et al., 2014a, MNRAS, 439, 2
- von der Linden A., et al., 2014b, MNRAS, 443, 1973



# Acknowledgements

I want to thank all the people that have been by my side and supported me ever since I embarked on this journey from the beginning to the very end of this PhD life and those I have met along the way that made my life so much better.

To my supervisor, Joseph Mohr, for the opportunity to carry out my PhD. Thank you for your guidance throughout this time and the many opportunities for me to present my work at conferences and meetings. To Sebastian G. goes my special gratitude. Thank you for being my teacher over the years, I would have never made it this far without your continuous help and guidance as well as your patience. To Sebastian B. and Matthias for your help and guidance I thank you. To Raffaella for making my time at USM a happier one.

I want to thank the amazing group of friends I made right at the start of my PhD life, Jere, Souradeep, Francesco, Matías, Rima and Misa, who made me feel at ease and at home in a new country right away. Thank you for the travels, dinners, board games, the continuous support and the laughter we shared. I want to especially thank Souradeep who has been there for me and supported me throughout my darkest times, on a daily basis if needed, and has been a constant source of comfort. I thank Daniel and Nahir, who have welcomed me into their home with open arms, for their friendship and support and all the awesome movie nights. To Vicki and Vaida, who always make me feel at home when I come back to the UK. Thank you for your continuous friendship over the years. It means so much to me. To the friends I made over the course of time, Linda, Vlas, Agnè, Ivan and Chiara. Thank you for being by my side.

To my family, my parents, my brother and his wonderful girlfriend who have been there with me alongside this journey and supported me and built me up whenever I needed it.

To Jere, for holding my hand when I needed it and your support, especially towards these last few steps of this journey. Thank you for going on all these crazy trips and adventures with me and I hope there will be many more to come.