From galaxy clusters to cosmology

Hernández Lang, Daniel Ernesto

München 2024

From galaxy clusters to cosmology

Hernández Lang, Daniel Ernesto

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

vorgelegt von
Hernández Lang, Daniel Ernesto
aus Antofagasta, Chile

München, den 10.12.2024

Erstgutachter: Prof. Dr. Joseph J. Mohr
Zweitgutachter: Prof. Dr. Jochen Weller
Tag der mündlichen Prüfung: 04.02.2025

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ZUSAMMENFASSUNG

Galaxienhaufen sind die größten gravitativ gebundenen Strukturen im Universum. Mit der Halo-Massenfunktion (HMF) lässt sich die Anzahl der Haufen in einem bestimmten Massenbereich für eine feste Rotverschiebung vorhersagen. Die HMF hängt jedoch von kosmologischen Parametern wie der Dichte der gesamten Materie im Universum, Ω_m, und der Amplitude der Dichtefluktuationen, σ_8, ab. Folglich kann die beobachtete Anzahl von Galaxienhaufen Einschränkungen für diese Parameter liefern. Da die Massen der Haufen nicht direkt beobachtbar sind, sind Skalierungsrelationen, die beobachtbare Eigenschaften mit den tatsächlichen Massen der Galaxienhaufen verknüpfen, von entscheidender Bedeutung. In diesem Zusammenhang ist das Verständnis der Erstellung von Galaxienhaufenkatalogen—einschließlich der Auswahl- und Bestätigungsprozesse—und die präzise Bestimmung der Parameter der Skalierungsrelationen grundlegend für die Nutzung der Haufenzahl als kosmologische Probe.

Meine erste Studie konzentriert sich auf die Analyse des verschmelzenden Galaxienhaufens SPT-CL J0307-6225. Durch die Untersuchung der Verschmelzungsdynamik ist es möglich Substrukturen zu trennen und ein wahrscheinliches Massenverhältnis von 1,3 zu finden. Zudem deutet die Analyse der Galaxienpopulation auf eine frühere Verschmelzung in einer der Substrukturen hin.

In meiner zweiten Studie verwende ich Galaxienhaufenkandidaten, die durch den thermischen Sunyaev-Zeldovich-Effekt mit Daten des *Planck*-Satelliten (bis zu S/N= 3) ausgewählt wurden, und suche mithilfe von photometrischen Daten aus der dritten Datenveröffentlichung der Dark Energy Survey nach optischen Gegenstücken. Der endgültige Katalog, PSZ-MCMF, enthält über 800 bestätigte Haufen mit einer Reinheit von 90%.

In meiner dritten Studie zeige ich, wie ein röntgenbasierter und optisch bestätigter Galaxienhaufenkatalog (RASS-MCMF) genutzt werden kann, um kosmologische Parameter einzuschränken. Mit einer simulierten Haufensammlung, die ähnliche Eigenschaften wie der 99% reine Teilkatalog von RASS-MCMF (~5000 Haufen) aufweist, prognostiziere ich eine parametergenauigkeit von 0,026, 0,033 und 0,15 (1σ) für Ω_m, σ_8 und w voraus.

Schließlich erweitere ich in meiner vierten Studie die Analyse aus der dritten Arbeit, indem ich die Modellierung des RASS-MCMF-Samples verbessere. Diese Verbesserungen umfassen eine neue Methode für die Likelihood der Haufenhäufigkeit und die explizite Einbeziehung der Massenkalibrierung durch den schwachen Gravitationslinseneffekt. Die Ergebnisse zeigen, dass diese Verbesserungen eine Einschränkungsgenauigkeit liefern, die mit den neusten Ergebnissen von SPT and eRASS1.

ABSTRACT

Galaxy clusters are the largest collapsed structures in the universe. Using the halo mass function (HMF), we can predict the number of clusters within a mass range for a fixed redshift. The HMF, however, depends on cosmological parameters such as the total matter density, Ω_m, and the amplitude of matter density fluctuations, σ_8. Consequently, the observed number of galaxy clusters can provide constraints on these parameters. Since cluster masses are not directly observable, scaling relations that link observable properties to true cluster masses are crucial. In this context, understanding the creation of galaxy cluster catalogs—including selection and confirmation processes—and accurately constraining the parameters of the observable-mass relation are fundamental for the use of cluster number counts as cosmological probes.

My first study focuses on the analysis of the merging galaxy cluster SPT-CL J0307-6225. Through an analysis of its merging dynamics, I separate the substructures and find a likely mass ratio of ∼1.3. On the other hand, the analysis of the galaxy population hints towards a previous merger in one of the substructures.

In my second study, I use galaxy cluster candidates, selected using the thermal Sunyaev-Zeldovich effect with data from the *Planck* satellite (down to S/N= 3), and look for optical counterparts using photometric data from the Dark Energy Survey data release 3. The final catalog, PSZ-MCMF, contains over 800 confirmed clusters with a purity of 90%.

In the third study I demonstrate how to use a X-ray selected and optically confirmed galaxy cluster sample (RASS-MCMF) to get cosmology constrains. Using a mock cluster sample with properties similar to the 99% pure subset of RASS-MCMF (∼5000 clusters), I forecast constraining powers of 0.026, 0.033, and 0.15 (1σ) for the parameters Ω_m, σ_8, and w, respectively.

Finally, in my fourth study, I expand the analysis from the third study by improving the modeling of the RASS-MCMF sample. These improvements include a new method for the abundance likelihood and the explicit inclusion of weak-lensing mass calibration. The results indicate that these improvements yield constraining power comparable to the latest results from SPT and eRASS1.

CHAPTER 1

INTRODUCTION

Galaxy clusters, immense associations of hundreds to thousands of galaxies, orbiting a common gravitational potential, are the most massive collapsed structures in the universe. These galaxies are not only embedded in hot gas but also enveloped in the immense dark matter halos that govern their dynamics. Galaxy clusters have fascinated astronomers since the 18th century, when F. Wilhelm Herschel first described the Coma cluster as "*that remarkable collection of many hundreds of nebulæ which are to be seen in what I have called the nebulous stratum of Coma Berenices*" (Herschel, 1785).

Given a theory of their formation and evolution, it is possible to construct what is called the halo mass function. This mass function predicts the number of clusters within mass intervals per volume unit. The predicted number depends on the different cosmological parameters that describe the growth of structures in the universe. Using cluster catalogs, one can count the number of galaxy clusters with different masses to constrain cosmological parameters. The biggest challenges are a comprehensive understanding of the theory and a precise model that can reconstruct the observed number of clusters for a given cluster catalog.

In this chapter, we will explore the formation of these structures, focusing on the initial conditions of the universe that led to the formation of galaxy clusters, as well as the comparison between observations and theoretical predictions. Their general properties, scaling relations and different detection methods will also be discussed. All of this will serve as pillars to understand the goals of this thesis: using galaxy clusters to constrain cosmology.

1.1 Formation and evolution of halos

1.1.1 Early Universe

The cosmological principle states that the Universe is homogeneous and isotropic at large scales. This principle has been supported by various observations, including measurements of the Cosmic Microwave Background (CMB) anisotropies (Smoot et al., 1992; Spergel et al., 2003; Planck

Collaboration et al., 2014a), and the large scale distribution of galaxies (e.g., Peebles, 1980).

In the 1970s, astronomers had a major problem regarding the nature of the early Universe: photons coming from opposite parts of the Universe, originating when the Universe was only a few hundred thousand years old, appeared to have almost identical properties. In fact, photons coming from the early Universe, and observed from any direction, have nearly the exact same temperature. This was known as the *horizon* problem (Rindler, 1956; Weinberg, 1972; Misner et al., 1973). How could structures so far apart, with no causal contact, exhibit similar properties? The theory of inflation, introduced in 1981 (Guth, 1981), has been widely accepted to address such problem, among others (e.g., the lack of magnetic monopoles; Zeldovich & Khlopov, 1978; Guth & Tye, 1980).

Inflation refers to a phase in the early universe, a fraction of a second after the Big Bang, when there was an exponential expansion that increased the size of the universe by at least a factor of about 10^{26}. After inflation ended, the universe entered the radiation-dominated era, where photons were scattered by free electrons in a hot plasma. In this chaotic environment, some overdensities began to form, creating perturbation waves that would later collapse. However, the gravitational collapse of these clumps of baryonic matter was counteracted by radiation pressure, preventing further clustering. Nonetheless, the existence of dark matter (DM) allowed structures to form. Dark matter, which does not interact with baryonic matter except through gravity, was not affected by radiation pressure. As a result, perturbation waves involving dark matter could grow until collapsing into halos of galaxies and galaxy clusters. This is true if we consider DM to be non-relativistic, so that DM particles can interact with each other through their gravity. This non-relativistic DM is referred to as cold dark matter (CDM).

This era continued until the Universe cooled sufficiently for electrons to combine with photons during the recombination era, ∼378 thousand years after the Big Bang. This cooling process was a result of the continued expansion of the universe following the inflationary phase, although the expansion was not as rapid as during inflation. The formation of atoms (along with photons not being scattered), allowed for them to be influenced by the gravitational potential of dark matter clumps, leading to the coupling of dark matter and baryonic matter. This coupling eventually gave rise to the formation of the first stars, star associations, and galaxies.

1.1.2 Redshift and cosmological parameters

To get a better picture of the formation of halos, it is useful to define a few quantities that govern the growth and behaviour of the universe and structures.

1.1.2.1 Expansion rate

Let us define the cosmic scale factor $a(t)$ as a measure of how big or small the universe is at any given moment compared to its size in the past or future. This definition of the scale factor arises from the need of describing the position of objects without taking into account the expansion of space: comoving coordinates. The position of an object at any given time, $\mathbf{r}(t)$, is then defined using the scale factor, such that

$$\mathbf{r}(t) = a(t)\mathbf{x} \tag{1.1}$$

where $\mathbf{x}$ is the position of the object at present time (t_0). Thus, the scale factor is defined such that $a(t_0) = 1$.

The velocity of the comoving object can be determined from the time derivative of Eq. 1.1

$$\mathbf{v}(\mathbf{r}, t) = \frac{\mathrm{d}}{\mathrm{dt}}\mathbf{r}(t) = \frac{\mathrm{da}}{\mathrm{dt}}\mathbf{x} \equiv \dot{a}\mathbf{x} = \frac{\dot{a}}{a}\mathbf{r} \tag{1.2}$$

where we will define the expansion rate as

$$H(t) = \frac{\dot{a}}{a} \tag{1.3}$$

with $H(t)$ defined as the Hubble parameter and has units of km s^{-1} Mpc^{-1}. The Hubble parameter determines the rate at which the universe expands at different times. At present day, $H_0 = H(t_0)$, it represents the current expansion and is used to estimate the age of the universe.

Hubble (1929) determined the first value of $H_0 = 500$ km s^{-1} Mpc^{-1} using, among others, Cepheid variables[1], in order to study the speed at which distant nebulæ were travelling away from Earth

$$v = H_0 D \tag{1.4}$$

where v is the recessional velocity in km s^{-1} and D is the proper distance in Mpc. Eq. 1.4 is often referred to as "Hubble's law".

Ever since this discovery, deriving a precise value of the Hubble's constant has become one of the main drivers of current cosmological analysis. In 2011, Saul Perlmutter, Brian P. Schmidt, and Adam G. Riess received the nobel prize in physics for discovering that the universe is not only expanding, but in accelerated expansion (Riess et al., 1998). They use type Ia supernovae, a type of supernova that involves a binary system in which one of the star companions is a white dwarf that gradually accretes mass from it's companions until they reach a critical mass limit (Chandrasekhar, 1931) and they explode as supernova. Type Ia supernovae are known as standard candles because their luminosity peak is fairly consistent, and thus can be used to measure the distance to their hosts galaxies.

Nowadays, the value of H_0 ranges from 66-75 km s^{-1} Mpc^{-1}, depending on the model and observations used (e.g., Riess et al., 2019; Planck Collaboration et al., 2020b)

1.1.2.2 Redshift

Redshift is the shift in the emitted light from a source with respect to the observed light, which is redder. In general, a photon travelling from any astronomical source will have it's wavelength either red-shifted or blue-shifted depending on if the source is moving away or towards the observer. In the case of cosmological redshift, given that the universe is expanding, a photon from any astronomical source will get its wavelength stretched by the stretch of space-time,

[1] variable stars with a regular period-luminosity relation.

meaning that the observed wavelength will always be larger than the emitted one, thus getting red-shifted. This is known as the Hubble's flow. It is important to note that all objects have peculiar velocities or intrinsic motions that would affect the measurement of the redshift.

The equation to estimate the redshift is

$$z = \frac{\lambda_o - \lambda_e}{\lambda_e} \tag{1.5}$$

where λ_o and λ_e are the observed and emitted wavelengths, respectively.

In cosmology it is common to define the redshift through the scaling factor $a(t)$, so that

$$1 + z = \frac{a(t_0)}{a(t)} = \frac{1}{a}, \tag{1.6}$$

where t is the cosmic time at which the source emitted the photon.

1.1.2.3 The Friedmann–Lemaître Equations

The Friedmann–Lemaître equations are a set of dynamical relations, first derived by the physicist and mathematician Alexandr Friedmann (Friedmann, 1922), based on the Einstein field equations of general relativity. These equations describe the time evolution of the cosmic scale factor $a(t)$ under the assumption of a homogeneous and isotropic universe, as described by the cosmological principle. At first, Friedmann's formulation did not include the cosmological constant, Λ, which had been introduced earlier by Albert Einstein in his field equations (Einstein, 1917)[2]. Later, Georges Lemaître rederived the Friedmann equations, this time incorporating the cosmological constant (Lemaître, 1927). The final forms of the Friedmann–Lemaître equations are

$$\left(\frac{\dot{a}}{a}\right)^2 + \frac{kc^2}{a^2} = \frac{8\pi}{3}G\rho + \frac{\Lambda c^2}{3}, \tag{1.7}$$

$$2\frac{\ddot{a}}{a} + \left(\frac{\dot{a}}{a}\right)^2 + \frac{kc^2}{a^2} = -\frac{8\pi}{c^2}Gp + \Lambda c^2 \tag{1.8}$$

where c is the speed of light in vacuum, G is the universal gravitational constant, ρ is the mass density and p the pressure. The constant k defines the spatial curvature, with typical values of $k = -1$ to define an open universe, $k = 0$ to define a flat universe and $k = +1$ to define a closed universe.

From Eq. 1.7, we can see that the latter term corresponds to an energy density with

$$\Lambda = \rho_{\text{vac}}\frac{8\pi G}{c^2} \tag{1.9}$$

with ρ_{vac} defined as the density of the vacuum. Thus, the latter corresponds to the vacuum energy density, where for positive values of Λ, this translates into a repulsive force that counteracts the gravitational attraction.

[2]The cosmological constant will be discussed in more detail in a following section.

1.1.2.4 Critical density

The critical density, ρ_c in cosmology is the density required for the universe to be flat ($k = 0$). It can be derived from the first Friedmann–Lemaître equation, assuming $\Lambda = 0$. The critical density today, $\rho_{c,0}$, can be derived by setting $t = t_0$ and $H = H_0$ in Eq. 1.7 as

$$\rho_{c,0} = \frac{3H_0^2}{8\pi G} = 1.88 \times 10^{-29} h^2 \mathrm{g/cm^3} \tag{1.10}$$

1.1.2.5 Power spectrum

A random field, such as the initial density field of the Universe (that we will discuss in a few sections), can be described by its power spectrum $P(k)$, where k is a wave number. Two particular power spectrums are the primordial power spectrum and the matter power spectrum. The former characterizes the distribution of fluctuations in the early universe that eventually grew into the large-scale structures we observe today, such as galaxies and galaxy clusters. The latter characterizes the distribution of matter density fluctuations in the late universe, after the primordial fluctuations have evolved due to cosmic expansion, gravitational growth, and other processes (such as the influence of dark matter). In this subsection I will focus on the matter power spectrum.

The matter power spectrum of the Universe describes the level of structure as a function of the length-scale, $L \approx 2\pi/k$, where for larger values of $P(k)$, the larger the amplitude of the fluctuations on such length-scale. In cosmology, the power spectrum $P(k)$ is related to the matter correlation function $\xi(r)$, through a Fourier transform, such that,

$$P(k) = 4\pi \int_0^\infty \mathrm{d}r r^2 \frac{\sin kr}{kr} \xi(r), \tag{1.11}$$

where the correlation function is mathematically defined as

$$\xi(r) = \langle \delta(\mathbf{x}) \delta(\mathbf{x}') \rangle = \frac{1}{V} \int \mathrm{d}^3\mathbf{x} \delta(\mathbf{x}) \delta(\mathbf{x} - \mathbf{r}) \tag{1.12}$$

with $\mathbf{r} = \mathbf{x} - \mathbf{x}'$ and

$$\delta(\mathbf{x}) = \frac{\rho(\mathbf{x}) - \bar{\rho}}{\bar{\rho}}, \tag{1.13}$$

corresponding to a matter overdensity with respect to the mean matter density of the Universe $\bar{\rho}$.

Eq. 1.11 corresponds to the integral over the correlation function with a weight factor which depends on the wave number $k \sim 2\pi/L$. Fig. 1.1 shows the matter power spectrum described by different probes from small scales (higher k) to higher scales (smaller k), using data from different surveys.

1.1.2.6 Cosmological parameters

Cosmological parameters refer to a set of fundamental quantities that describe the properties and behavior of the universe on a large scale. These parameters help us understand the origin, evolution, and overall structure of the universe.

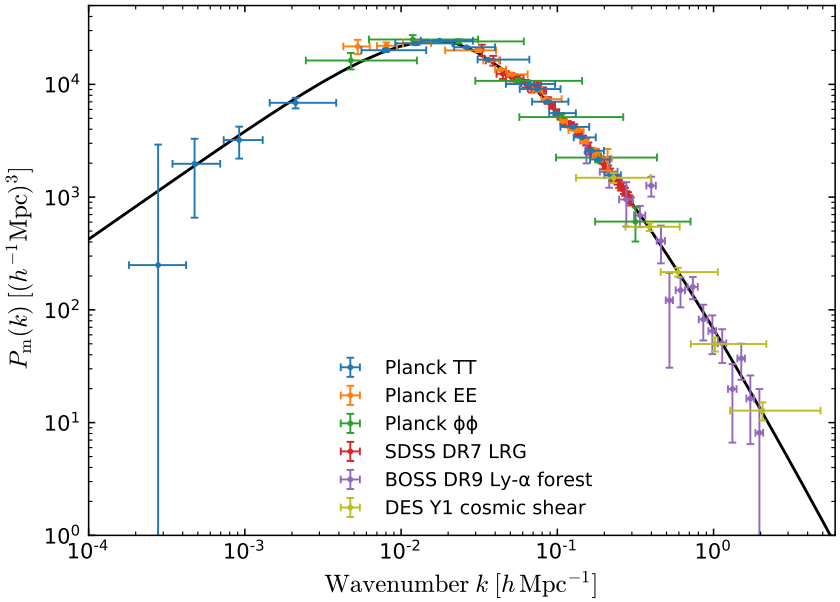

Figure 1.1: Linear-theory matter power spectrum (at $z = 0$), assuming ΛCDM cosmology, inferred from different cosmological probes. Taken from (Planck Collaboration et al., 2020a)

Among the cosmological parameters we can find:

1) Density Parameters (Ω): These parameters quantify the relative amounts of different types of matter and energy in the universe. They include the density of dark energy (Ω_Λ) and of matter (Ω_m), which is subdivided into the density of baryonic matter (Ω_b) and dark matter ($\Omega_{\rm DM}$). The radiation density is also included (Ω_γ), however it is only relevant at redshifts around or higher than the radiation dominated era ($z \gtrsim 3400$). For models of the universe which are not flat ($k \neq 0$), a curvature component, Ω_k, is also considered.

   To derive the density parameters, we normalize the first Friedmann–Lemaître queation (Eq. 1.7) by $H^2 = (\dot{a}/a)^2$ as

$$1 = \frac{8\pi}{3H^2}G\rho + \frac{\Lambda c^2}{3H^2} - \frac{kc^2}{a^2H^2} \tag{1.14}$$

$$1 = \Omega_m + \Omega_\Lambda + \Omega_{\rm curv} \tag{1.15}$$

   Particularly, the mass density parameter of the universe today, can be measured by replacing $\rho = \rho_0$ and $H = H_0$ in Eq 1.15

$$\Omega_{m,0} = \frac{8\pi}{3H_0^2}G\rho_0 = \frac{\rho_0}{\rho_{c,0}} \tag{1.16}$$

2) Dark Energy Equation of State (w): Dark energy is a hypothetical form of energy that is thought to drive the accelerated expansion of the universe (Riess et al., 1998; Perlmutter

et al., 1999). The change in the energy density of dark energy is modelled through it's equation of state, w, which is defined as

$$w = \rho/P \tag{1.17}$$

where P and ρ are the pressure and energy density of dark energy, respectively. Particularly, for a value of $w = -1$, the accelerated expansion comes in the form of a cosmological constant; Λ.

This is derived from the first law of thermodynamics, $dU = -P\mathrm{d}V$, where U is the total energy and can be related to the energy density as $U = \rho V$, and V is the comoving volume $V \propto a(t)^3$. Substituting one can find that

$$\dot{\rho} + 3H(\rho + P) = 0, \tag{1.18}$$

where, for $w = -1$, then $\dot{\rho} = 0$, meaning that the energy density is constant over time.

3) Neutrino Masses (m_ν): Neutrinos are elementary particles that have a small but non-zero mass, and come in three different flavours; electron neutrino (ν_e), muon neutrino (ν_μ) and tau neutrino (ν_τ). Although the exact mass of neutrinos is not known, upper limits on the total neutrino mass, Σm_ν, have been found (e.g. $\Sigma m_\nu < 0.11 eV/c^2$, Planck Collaboration et al., 2020b). Given that neutrinos have a non-zero mass, they affect the formation and growth of structures in the universe, particularly on small scales (such as galaxies and galaxy clusters).

1.1.3 Non-linear evolution of galaxy clusters

In the early universe, when density fluctuations were relatively small, linear perturbation theory could be applied to predict the initial growth of larger structures, such as galaxy clusters. This mathematical framework describes how small density perturbations evolve under the influence of gravity, using linear equations like the Newtonian description of gravity and the Poisson equation. However, as these structures evolve, they enter a non-linear regime where gravitational interactions become significantly more complex.

For galaxy clusters, this transition is particularly pronounced. Clusters form through the accumulation of matter in regions with higher initial density fluctuations, but as they grow, the interplay between galaxies, dark matter, and other matter leads to non-linear dynamics. In this regime, the gravitational forces increase, driving significant deviations from linear predictions. Consequently, while linear perturbation theory lays the groundwork for understanding cosmic structure formation, accurately modeling the evolution of galaxy clusters necessitates numerical simulations to account for the intricate, non-linear interactions that dominate their dynamics.

A particular model, called the model of spherical collapse, is simple enough to understand the fundamental principles of gravitational collapse in the scales of galaxy clusters. But first, some background regarding linear evolution is needed. Following the discussion in Section 1.1.1, in the early universe, before recombination, the distribution of matter was similar to a Gaussian

field, with regions having small overdensities and underdensities compared to the mean density of the universe at any given time or redshift. The density contrast of any region in the universe during this stages can be expressed as

$$\delta(\mathbf{r}, t) = \frac{\rho(\mathbf{r}, t) - \langle\rho(t)\rangle}{\langle\rho(t)\rangle} \tag{1.19}$$

where $\langle\rho(t)\rangle$ is the mean cosmic matter density in the Universe at time t.

Considering an overdense region where the numerator, $\rho(\mathbf{r}, t) - \langle\rho(t)\rangle > 0$, then $\delta(\mathbf{r}, t) > 0$. In this overdense region, the gravitational field will be stronger than that which corresponds to the mean Hubble expansion. This translates into an overall slower expansion with respect to the average Hubble expansion in the Universe. This will again increase the value of $\delta(\mathbf{r}, t)$, thus decreasing the expansion even further. The opposite process will happen in the case of underdense regions, which will experiment an above average Hubble expansion.

Now, lets consider the model of spherical collapse, where a spherical symmetric region has a density higher than the mean cosmic density

$$\rho(t) = [1 + \rho(t)]\,\langle\rho(t)\rangle \tag{1.20}$$

In this model, the spherically-symmetric density fluctuation has a mass of

$$M = \frac{4\pi}{3} R_i^3\,\langle\rho(t)\rangle\,(1 + \delta_i) \tag{1.21}$$

where R_i is the initial radius of the density fluctuation with an initial amplitude $\delta_i > 0$.

Considering $\delta_i \ll 1$, we can evolve this overdensity assuming a closed universe with $\Omega_m = 1 + \delta_i$. This particular cosmology is similar to the Einstein-de Sitter model of the universe, with $\Omega_m = 1$ and $\Omega_\Lambda = 0$ (Einstein & de Sitter, 1932). By substituting $\Lambda = 0$ and $\Omega_k < 0$, with $k > 0$ in the Friedmann–Lemaître equations, we find that, for our particular model, the density fluctuation will reach a maximum value of the scale factor $a_{\max}$, and then it will collapse. Thus, the density perturbation will increase its size from R_i to a maximum $R_{\max}$ at a time $t = t_{\max}$, before collapsing to a single point at $t = t_{\rm coll}$.

The Friedmann–Lemaître equation for our model is thus

$$\left(\frac{\dot{a}}{a}\right)^2 = H_0^2 \Omega_m a^{-3} + H_0^2 (1 - \Omega_m) a^{-2} \tag{1.22}$$

where we have used the knowledge that $\Omega_m + \Omega_k = 1$ and $\Omega_m > 1$. A parametric solution for Eq. 1.22 can be found as

$$R = A(1 - \cos(\theta)),\ t = B(\theta - \sin(\theta)) \tag{1.23}$$

where θ is a parameter that helps to describe how the scale factor evolves with time in the spherical collapse model. For $t = 0$ we have $\theta = 0$, at turnaround $t = t_{\max}$ we get $\theta = \pi$ and at $t = t_{\rm coll}$ we will have a value of $\theta = 2\pi$. Normalising Eq. 1.23 by the values $R_{\max}$ and $t_{\max}$ at $\theta = \pi$, we get

$$\frac{R}{R_{\max}} = \frac{1}{2}(1 - \cos(\theta)),\ \frac{t}{t_{\max}} = \frac{1}{\pi}(\theta - \sin(\theta)) \tag{1.24}$$

Using the Maclaurin expansions for $\cos(\theta)$ and $\sin(\theta)$ in Eq. 1.24, we get:

$$\frac{R}{R_{\rm max}} \simeq \frac{\theta^2}{4} - \frac{\theta^4}{48}, \; \frac{t}{t_{\rm max}} \simeq \frac{1}{\pi}\left(\frac{\theta^3}{6} - \frac{\theta^5}{120}\right) \tag{1.25}$$

Finally, combining both expresions in Eq. 1.25, and substituting $R = a$, we get the linearised scale factor, $a_{\rm lin}$, as

$$\frac{a_{\rm lin}}{a_{\rm max}} \simeq \frac{1}{4}\left(6\pi\frac{t}{t_{\rm max}}\right)^{\frac{2}{3}}\left[1 - \frac{1}{20}\left(6\pi\frac{t}{t_{\rm max}}\right)^{\frac{2}{3}}\right] \tag{1.26}$$

with the first term corresponding to the expansion of the background in our model of the universe $a_{\rm bkg}/a_{\rm max}$.

Given that we consider a matter dominated universe, the mass-energy density, Ω_m, varies as a^{-3}. Thus, we can substitute using our δ value from before as

$$1 + \delta_{\rm lin} = \left(\frac{a_{\rm bkg}}{a_{\rm lin}}\right)^3 \tag{1.27}$$

Substituting 1.25 into 1.27, and assuming that for values of $\delta \ll 1$, $(1+\delta)^{-1/3} \simeq 1 - \frac{1}{3}\delta$, we get

$$\delta_{\rm lin} = \frac{3}{20}\left(6\pi\frac{t}{t_{\rm max}}\right)^{\frac{2}{3}} \tag{1.28}$$

so that, at turnaround ($t = t_{\rm max}$), we get $\delta_{\rm lin}^{\rm turn} = 1.06$. After turnaround, collapse proceeds symmetrically and finishes at $t = 2t_{\rm max}$, which corresponds to $\delta_{\rm lin}^{\rm coll} = 1.686$.

This is only valid in the linear regime, however at values of $\delta \approx 1$ this is no longer valid. Considering non-linear effects at turn-around, we instead get

$$1 + \delta_{\rm nonlin}^{\rm turn} = \left(\frac{a_{\rm bkg}}{a_{\rm max}}\right)^3 = \frac{(6\pi)^2}{4^3}\frac{a_{\rm max}}{a_{\rm max}} = 5.55 \tag{1.29}$$

However, so far we have been considering that this density perturbation will collapse to a single point with $\delta = \infty$ at $t = 2t_{\rm max}$. In reality, the perturbation will reach virial equilibrium at $t = t_{\rm max}$. The virial theorem states that

$$\frac{1}{2}U_{\rm vir} = -T_{\rm vir} \tag{1.30}$$

where U is the potential energy and T the kinetic energy. At turnaround, $T = 0$, because of energy conservation we get that $U_{\rm turn} + T_{\rm turn} = E = \frac{1}{2}U_{\rm vir}$, and thus $U_{\rm turn} = \frac{1}{2}U_{\rm vir}$, which translates to $2R_{\rm vir} = R_{\rm max}$. This means that the density of a virialized object is 8 times higher.

On the other hand, given that $a_{\rm bkg} \propto t^{2/3}$, by $t = 2t_{\rm max}$, the background scale factor will go down by a factor of 4.

Combining this with the number found in Eq. 1.29, we get for a virialized object

$$\Delta_c = 1 + \delta_{\rm nonlin}^{\rm vir} = 5.55 \times 8 \times 4 \approx 178 \tag{1.31}$$

This value is in good agreement with those found by simulations (e.g., White, 2001), and thus a typical value to define virial quantities (such as mass or radius) is $\Delta_c = 200$.

1.1.4 Halo mass function

The halo-mass function at a given redshift z, $n(M, z)$, is used to predict the abundance of collapsed halos (such as galaxies, galaxy clusters or dark matter halos) within a mass range $[M, M + \Delta M]$ and a given volume of the Universe. A pioneering and common mathematical model is the so called Press–Schechter formalism (Press & Schechter, 1974), based on an analytical formulation. In this model, the initial density field features fluctuations $\delta_0(\mathbf{x})$ on all scales. This field is then smoothed by convolving with a top-hat filter with a smoothing size given by a sphere of mass M centered at $\mathbf{x}$. Peaks found in this smoothed density field, $\delta_M(\mathbf{x})$, will have a size $\geq R$, with masses of $M \sim (4\pi R^3/3)\rho_0$. If the amplitude of these peaks, δ_M, are sufficiently large, $\delta_M(\mathbf{x}) > \delta_c$, then a sphere with mass $\geq M$ will start to grow non-linearly, come to a halt, and re-collapse, similar to the spherical collapse model.

Under the previous assumptions, and since the perturbations δ_M are Gaussian distributed with mean 0 and variance $\sigma(M)$, the probability for a smoothed linearly-evolved density field to have peaks that exceed the critical density contrast δ_c is given by

$$p_{>\delta_c}(M, z) = \frac{1}{\sqrt{2\pi}\sigma_M(z)} \int_{\delta_c}^{\infty} \mathrm{d}\delta_M \exp\left(-\frac{\delta_M^2}{2\sigma_M(z)^2}\right) = \frac{1}{2}\mathrm{erfc}\left(\frac{\delta_c}{\sqrt{2}\sigma_M(z)}\right), \tag{1.32}$$

where $\delta_c = 1.686$ is the critical density contrast predicted by linear theory and erfc is the complementary error function. The variance, $\sigma_M(z)$, given by Press & Schechter (1974) is defined as

$$\sigma_M(z)^2 = \frac{\Sigma^2}{M^2} = \frac{\sigma^2}{M\rho} \tag{1.33}$$

where Σ is the standard deviation of the mass and σ^2 is the variance per volume.

The problem with eq. 1.32, is that integrating for the whole range of masses gives a values of 1/2 instead of 1. This is because the probability of density peaks with $\delta < \delta_c$ is assigned to zero, meaning that the Press-Schechter model does not account for the fact that smaller halos can be embedded within larger structures that will collapse later. To account for this, Press & Schechter (1974) simply multiplied their equation by a factor of 2. Since eq. 1.32 gives the volume of objects within a mass range, we can get the number of objects by dividing by the volume as

$$\begin{aligned} \frac{\mathrm{d}n(M, z)}{\mathrm{d}M} &= \frac{2}{V_M}\frac{\partial p_{>\delta_c}(M, z)}{\partial M} \\ &= \sqrt{\frac{2}{\pi}}\frac{\bar{\rho}}{\sigma_M(z)}\left|\frac{\mathrm{d}\log\sigma_M(z)}{\mathrm{d}\log M}\right| \exp\left(-\frac{\delta_c^2}{2\sigma_M(z)^2}\right) \end{aligned} \tag{1.34}$$

This is the final form of the Press-Schechter halo mass function. Cosmological parameters enter through the mass variance σ_M, through the linear perturbation growth factor and through the critical density δ_c. Since then, other models of the halo mass function have been introduced (e.g., Jenkins et al., 2001; White, 2002; Tinker et al., 2008). With the advent of numerical simulations,

comparisons with the predictions from analytical solutions became crucial in the development of an analytical formulation. Fig. 1.2 shows the halo mass function, expressed by the halo multiplicity function, $M^2\rho^{-1}\mathrm{d}n/\mathrm{d}M$, determined from the Millenium simulation (Springel et al., 2005), where the comparison to the Press-Schechter model is shown as dashed blue lines at $z = 0$ and $z = 10.07$. The solid black lines show the fit of the halo mass function to a number of N-body numerical simulations done by Jenkins et al. (2001), which shows a much better agreement than the Press-Schechter model.

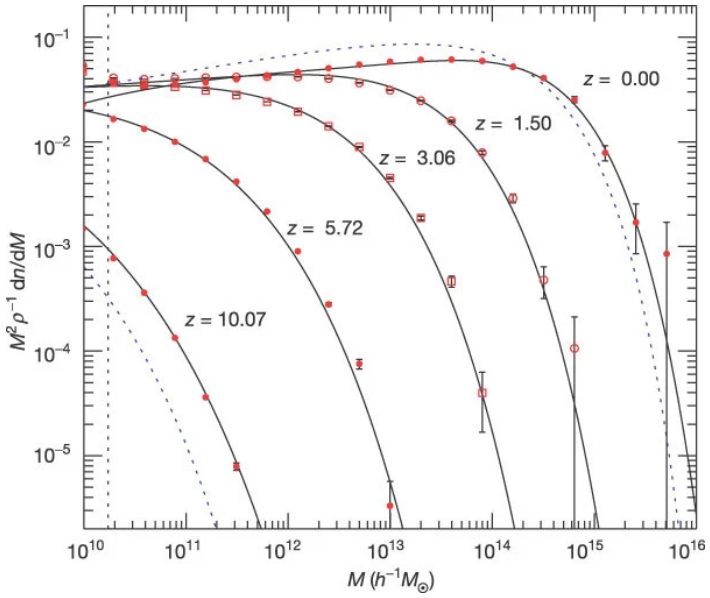

Figure 1.2: Halo multiplicity function $M^2\rho^{-1}\mathrm{d}n/\mathrm{d}M$ (symbols with 1σ error bars) taken from Springel et al. (2005), where ρ is the mean density of the Universe. Solid lines are predictions from an analytic fitting proposed by Jenkins et al. (2001), and the dashed blue lines give the Press–Schechter model (Press & Schechter, 1974) at $z = 0$ and 10.07.

Tinker et al. (2008) measured the mass function of dark matter halos in cosmological simulations with a flat ΛCDM cosmology, investigating its evolution at $z \leq 2$. They describe the halo abundance in their simulations as

$$\frac{\mathrm{d}n}{\mathrm{d}M} = f(\sigma)\frac{\bar{\rho}_m}{M}\frac{\mathrm{d}\ln\sigma^{-1}}{\mathrm{d}M} \tag{1.35}$$

In this model, the behavior of the overdensity, δ_M, as a function of smoothing scale can be described as a random walk. Each step of the walk corresponds to a change in mass scale (or smoothing scale), and the excursion occurs when δ_M crosses the critical threshold δ_c. The function $f(\sigma)$ describes the weighted distribution of the initial crossings of these random walks over a barrier, which separates collapsed regions from those that remain uncollapsed (such as

when the randomly varying overdensity first surpasses the critical value δ_c). The function $f(\sigma)$ is parameterized as

$$f = A\left[\left(\frac{\sigma}{b}^{-a}\right) + 1\right] e^{-c/\sigma^2}, \tag{1.36}$$

where

$$\sigma^2 = \int dk P(k)\hat{W}(kR)k^2 \tag{1.37}$$

is the matter variance, $P(k)$ is the power spectrum at wave number k, and $\hat{W}$ is the Fourier transform of the real-space smoothing scale defined as a top hat window function of radius $R = (3M/4\pi\bar{\rho})^{1/3}$. For higher masses M (increasing smoothing scale), the value of σ decreases, so there are fewer large-scale fluctuations. The best-fit values of the parameters A, a, b and c were determined by fitting Eq. 1.36 to the $z = 0$ simulations.

1.1.5 Amplitude of mass fluctuations

It is interesting to note that eq. 1.37 describes a fundamental cosmological parameter, σ_8, which corresponds to a measure of the rms (root mean square) fluctuation of matter density at a scale of 8 h^{-1} Mpc at $z = 0$, and it specifies how matter is distributed (clumped) in the present day Universe.

Although σ_8 characterizes the matter fluctuations at present, its value can help trace the history of structure formation. Comparing measurements of σ_8 at different epochs allows us to infer the rate at which structures grew. The present-day abundance of massive galaxy clusters depends exponentially on this parameter, assuming Gaussian initial fluctuations.

1.2 Properties and selection of galaxy clusters

Galaxy clusters are the largest collapsed structures in the universe. Their sizes vary from 1 to 3 h^{-1} Mpc, with a range of masses between 10^{13}-$10^{15.5}$ $M_\odot$ and between 50 to over 10000 galaxy members (Schneider, 2006). Following the hierarchical structure formation model in a cosmology with cold dark matter and a cosmological constant (ΛCDM), galaxy clusters evolve through a sequence of mergers with structures of similar mass and accretion of smaller mass systems (Kravtsov & Borgani, 2012). Their principal components are: galaxies (composed of gas, stars and dust), gas from the intra-cluster medium (hereafter ICM) and a DM halo. The baryonic matter accounts for up to 10% of the cluster's total mass, while the remaining 90% is in the form of DM. The 10% baryonic matter is mostly composed of the ICM ($\sim$ 9%) in which the galaxies ($\sim$ 1%) are embedded, which is in itself also embedded within the DM halo (for more details see Kravtsov & Borgani, 2012; Bykov et al., 2015).

Galaxy clusters are primarily composed of early-type galaxies. The pioneering work of Dressler (1980) found a relation between galaxy morphology and local density, with the fraction of elliptical (spiral) galaxies increasing towards higher (lower) densities. This relation has been confirmed up to $z \sim 1.3$ (Postman et al., 2005; Holden et al., 2007; van der Wel et al., 2007;

Desai et al., 2007; Mei et al., 2012). A similar trend has been observed between passive and star-forming galaxies (Fritz et al., 2014; Haines et al., 2015; Paccagnella et al., 2016). One mechanism that could explain these environment-based relationships is ram-pressure stripping (Gunn & Gott, 1972), where the gas in galaxies is removed by the hot plasma of the ICM as they cross the cluster. An extreme example of this process seen in so-called jellyfish galaxies (Poggianti et al., 2016; McPartland et al., 2016).

To constrain cosmological parameters, it is necessary to be able to systematically model the cluster selection, which means that we need to understand how galaxy cluster catalogs are created. We will here focus on three ways of finding galaxy clusters using two of their main components: the ICM and galaxies.

1.2.1 X–ray selected surveys

The ICM consists mainly of a hot plasma composed mostly of fully ionised Hydrogen and Helium atoms, at temperatures between 10^7 to 10^8 K in which electrons travelling near ions are decelerated. This, in turn, causes a loss of kinetic energy in the electron, which is converted into radiation, and also produce a loss of energy in the plasma, causing the overall temperatures of the gas to cool down. This effect is known as thermal *bremsstrahlung*. Given that these electrons often travel at relativistic speeds, the photons that are emitted as a consequence of the decceleration have an energy that peaks in X–ray wavelengths. It is common to express the temperature of the gas in terms of the particle energies, $k_B T$[3], which for galaxy clusters ranges from $\approx 0.1 - 10.0$ keV.

Thus, galaxy clusters can be found by observing with X–ray telescopes. Thargetted X-ray observations such as those made with the XMM-Newton observatory (Jansen, F. et al., 2001) or with the *Chandra* X-ray observatory (Weisskopf et al., 2000), have been used in small patches of the sky to get the deepest X-ray images (e.g., Giacconi et al., 2001; Hornschemeier et al., 2001; Finoguenov et al., 2007; Leauthaud et al., 2010). Shallower all sky survey observations have also been made using the ROSAT mission (Truemper, 1982) and with the recently launched eROSITA mission (Predehl et al., 2021), finding over 8,000 galaxy clusters with a sample purity of $\geq 90\%$ (Klein et al., 2023; Kluge et al., 2024).

1.2.2 The thermal Sunyaev-Zeldovich Effect

The cosmic microwave background (CMB), a radiation coming from the photons of what is referred to as the last scattering surface of the universe during the recombination era. It was first (accidentally) discovered in 1965 by Arno Penzias and Robert Wilson, when they encountered an unexpected noise being registered in their receiver system (Penzias & Wilson, 1965). In the early universe photons were Compton scattered by free electrons until the latter coupled with protons, increasing the mean free path of the photons until they were no longer being scattered ($z \sim 1100$). Considering the gravitational instabilities, density fluctuations on the CMB temperature are

[3]where $k_B = 8.617333262 \times 10^{-5}$ eV K^{-1} is the Boltzmann constant

expected to be of the order of $\Delta T/T \approx 10^{-5}$ (Smoot et al., 1992). These anisotropies on the CMB temperature are called *primary anisotropies*.

These photons travelled through the Universe and are now observed at a temperature of 2.7 K, following the black body spectrum. However, they can still be subjected to effects that can alter their energies. As mentioned above, the ICM of galaxy clusters has temperatures up to T$\sim 10^8$ K (for massive clusters). At these temperatures, electrons are moving at relativistic speeds through the ICM. Sunyaev & Zeldovich (1972) found that, as photons (coming from the CMB) travel through the universe, and in particular through galaxy clusters, on their way to us (the observers), they can be inversed Compton scattered by the high-energy electrons of the ICM. This translates into photons receiving a boost in energy by roughly $k_B T_e/m_e c^2$ compared to what we would expect from the black body spectrum of the CMB, where m_e is the mass of the electron and T_e is the electron temperature. This is known as the thermal Sunyaev-Zeldovich effect (SZE). Fig. 1.3 shows the expected black body spectrum of the CMB (dashed lines) and how it shifts after interacting with electrons from a massive cluster (solid lines; Carlstrom et al., 2002). This shift effectively translates into a decrease of the intensity at low frequencies, a negligible effect at exactly 217 GHz and an increase at higher frequencies.

This spectral distortion of the CMB is given by

$$\frac{\Delta T_{\rm SZE}}{T_{\rm CMB}} \propto y = \int {\rm d}l n_e \frac{k_B T_e}{m_e c^2} \sigma_T, \tag{1.38}$$

where y is the Compton y-parameter, σ_T is the Thomson cross-section for electron scattering and n_e is the electron number density.

Satellite missions have primarily driven the study of CMB anisotropies by mapping the full sky, with the latest being the *Planck* mission (Planck Collaboration et al., 2011). Over 1,600 cluster candidates have been found using the *Planck* data (Planck Collaboration et al., 2016a). Ground based observations at sub-mm and mm wavelenghts, such as with the Atacama Cosmology Telescope (ACT; Fowler et al., 2007; Swetz et al., 2011) or the South Pole Telescope (SPT; Carlstrom et al., 2011), have also been used to observe parts of the sky in more depth and with higher resolution. Using a combination of optical data and data from ACT, Klein et al. (2024b) confirmed over 6,000 galaxy clusters in a region of ∼13,000 deg^2.

1.2.3 Sequence of red cluster galaxies

A significant fraction of the members inhabiting galaxy clusters consist of elliptical and lenticular galaxies (Oemler, 1974; Dressler, 1980). These early-type galaxies feature a well defined linear relation between their colour and magnitude (Visvanathan & Sandage, 1977; Bower et al., 1992; Kodama & Arimoto, 1997), the so called red cluster sequence (RS). Such galaxies in the RS have null or little ongoing star formation (Gladders & Yee, 2005; Gladders et al., 2007; De Propris et al., 2016), and their colour evolution can be remarkably well described by simple evolutionary models (Stanford et al., 1998). In fact, such models are so successful that they have been used to identify clusters (e.g. Gladders & Yee, 2000; Rykoff et al., 2014; Bleem et al., 2015) and to provide robust photometric redshifts up to ∼1.5 (e.g., Song et al., 2012b; Bleem et al., 2020), with a precision better than $\sim 0.01 \times (1 + z)$ up to $z \sim 1.0$ (e.g. Rykoff et al., 2016; Klein et al.,

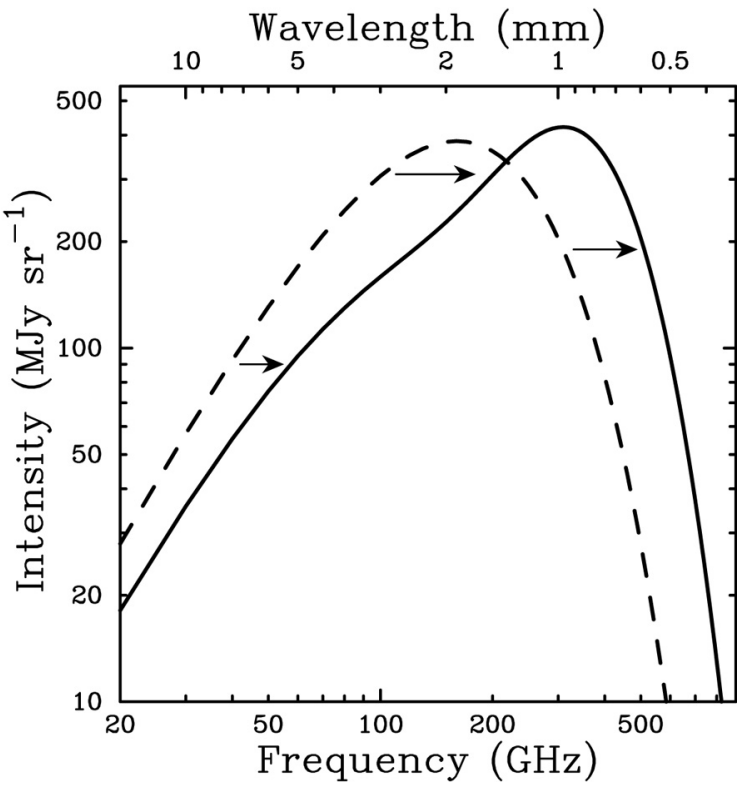

Figure 1.3: The Cosmic Microwave Background (CMB) spectrum, undistorted (dashed line) and distorted by the Sunyaev-Zel'dovich effect (SZE) (solid line), taken from Carlstrom et al. (2002). The SZE distortion shown is for a fictional cluster 1000 times more massive than a typical massive galaxy cluster. The SZE causes a decrease in the CMB intensity at frequencies $\lesssim 218$ GHz and an increase at higher frequencies.

2018, 2019; Hernández-Lang et al., 2022). Fig. 1.4 shows the models from Gladders & Yee (2000) that predict the slopes of the RS for clusters up to $z = 1.0$.

Koester et al. (2007a) used the maxBCG red-sequence method (Koester et al., 2007b) to find over 13,000 clusters using photometric data from the Sloan Digital Sky Survey (SDSS; York et al., 2000) within $0.1 < z < 0.3$. Rozo & Rykoff (2014) used spectroscopically confirmed clusters to model the RS at different redshifts, finding ~25,000 clusters within $0.08 < z < 0.5$ using the SDSS DR8 data. Wen & Han (2024) used the recently released DESI Legacy Imaging Surveys DR10 and found over 1.5 million clusters. RS fitting algorithms have also been used to confirm cluster candidates drawn from other methods, such as X–ray catalogs (e.g., Klein et al., 2019, 2022, 2023; Kluge et al., 2024) or via the SZE (Hernández-Lang et al., 2022; Klein et al., 2024a,b).

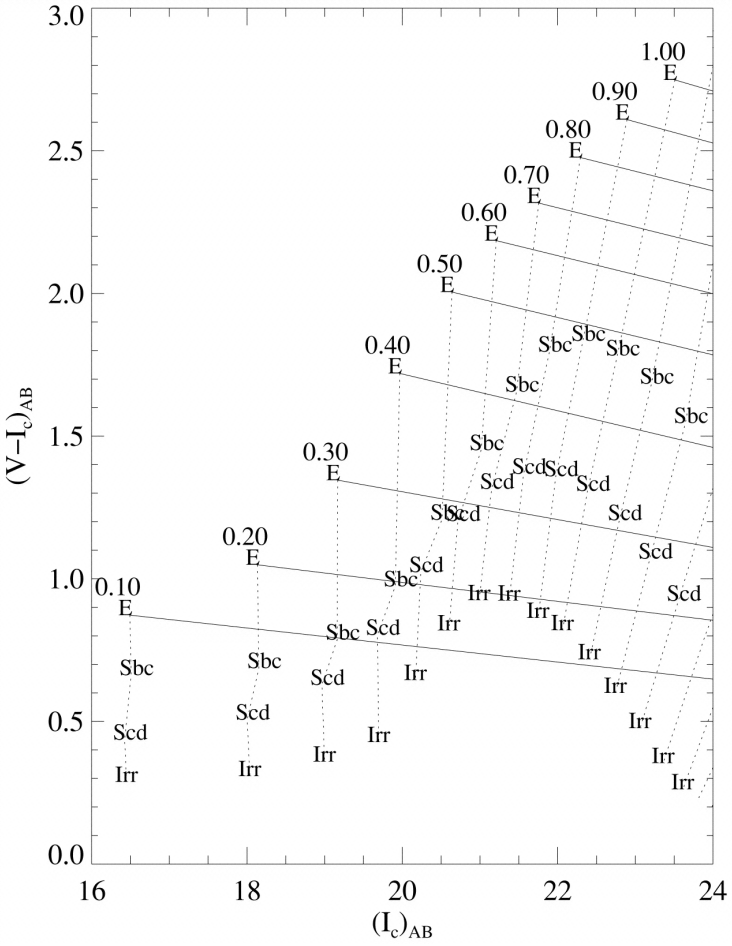

Figure 1.4: Colour-magnitude diagram showing the predicted slope of the red cluster sequence for clusters at different redshfits as black lines. Figure taken from Gladders & Yee (2000).

1.3 Galaxy clusters as cosmological probes

1.3.1 Overview

From the previous chapter we have now some understanding on how galaxy clusters can be selected, but how can we use them to infer the cosmological parameters that govern our universe? That's where the halo mass function becomes relevant. The halo mass function predicts the number of clusters, for a redshift, mass and volume range, given a set of cosmological parameters. Given a sample of galaxy clusters, with known masses, redshifts and selection function, we can use the abundance of observed clusters and compare with the theoretical halo mass function to get insight on the underlying cosmological parameters that best describe our data sets (e.g., Haiman et al., 2001). First analysis using low redshift galaxy clusters were able to constrain σ_8 with a ~25% accuracy (Evrard, 1989; Frenk et al., 1990; White et al., 1993; Viana & Liddle, 1996). Given that at low redshift σ_8 and Ω_m are strongly degenerate (Huterer & White, 2002; Spergel

et al., 2003), using cluster samples that spawn to higher redshift has proben useful to break this degeneracy and put constrains on Ω_m (Bahcall & Fan, 1998; Blanchard & Bartlett, 1998; Bahcall & Bode, 2003).

However, it is not possible to directly observe the masses of the clusters, but rather mass proxies (or observables) that are correlated to the mass through scaling relations (such as velocity dispersion of galaxies, X-ray fluxes, SZE fluxes, etc). Indeed, it has been demonstrated that surveys with a sufficiently large sample of galaxy clusters–typically in the range of hundreds to thousands of clusters–can provide the necessary data to constrain both cosmological parameters and the scaling relation parameters (Levine et al., 2002; Majumdar & Mohr, 2003, 2004; Bocquet et al., 2015, 2019; Chiu et al., 2023).

Recent surveys like the South-Pole Telescope Sunyaev-Zeldovich (SPT-SZ; Bleem et al., 2015) survey, in combination with SPTpol ECS (Bleem et al., 2020), and SPTpol 500d (Bleem et al., 2024) surveys, and the eROSITA all sky survey (eRASS; Merloni et al., 2024), have been used to get the most precise values of cosmological parameters to date using the number counts (or abundance) of galaxy clusters (Bocquet et al., 2024; Ghirardini et al., 2024). Fig. 1.5 shows the constraining power from the abundance of SPT clusters with respect to other probes, such as the constraints derived from the CMB maps by Planck Collaboration et al. (2020b).

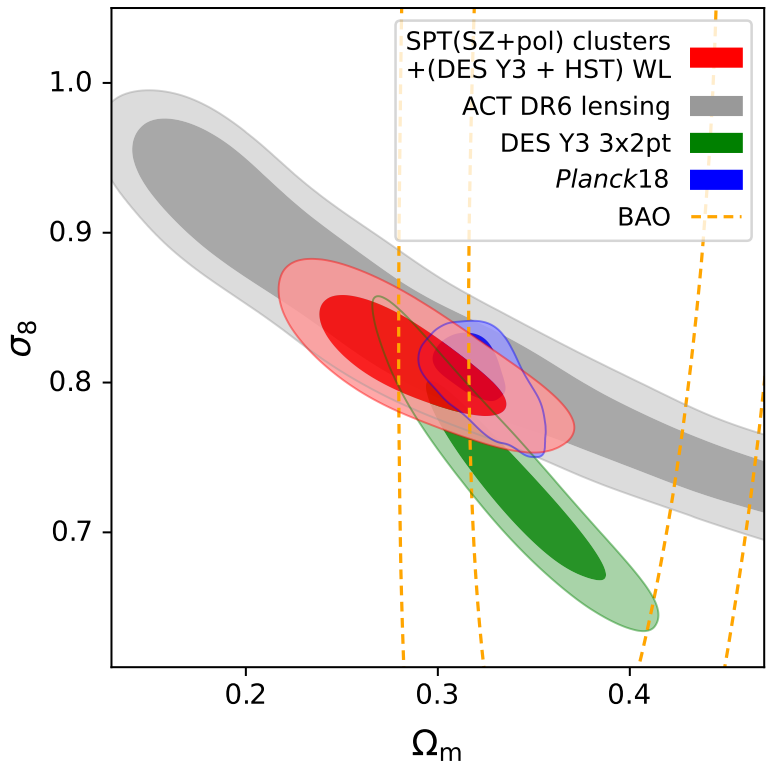

Figure 1.5: Comparison of constraints on Ω_m and σ_8 (68% and 95% credible regions) in a ΛCDM universe with massive neutrinos from different probes. Taken from Bocquet et al. (2024).

1.3.2 Cluster masses

The abundance of massive galaxy clusters can be used to constrain the mass function and, in particular, the cosmological parameters that describe it (such as Ω_m, σ_8, and w). To achieve this, appropriately calibrated scaling relations between cluster masses and their observables are required. These observables include, for example, the integrated SZ signal or the X-ray temperature.

The mass definition I will use from here on is

$$M_{\Delta_c} = \frac{4\pi}{3}\Delta_c \rho_c R^3_{\Delta_c} \tag{1.39}$$

where M_{Δ_c} is defined as a the mass of a sphere of radius $R^3_{\Delta_c}$ that encloses a density which is Δ_c times higher than the critical density of the universe. In general, the values of Δ_c are chosen as 200 (which yields the approximate virial mass and radius of a cluster) or 500 when methods that depend on the gas content of the cluster are used (see Evrard et al., 1996).

It is expected that, for example, the larger the luminosity and the integrated SZ signal of a cluster are, the more massive the cluster is. In general, the scaling relations are derived from different equations that relate the different observables to the cluster masses following different assumptions (e.g. hydrostatic equilibrium), and are characterised by the probability of getting the observable given the mass and redshift of the cluster. Often these probabilities are modelled as a lognormal distribution, such that

$$P(\ln O|\ln M, z) = \mathcal{N}(\langle \ln O|\ln M, z\rangle, \sigma^2_{\ln O}) \tag{1.40}$$

with mean $\langle \ln O|\ln M, z\rangle$, described by the scaling-relation between an observable, O, and the mass and redshift of the cluster, with variance $\sigma^2_{\ln O}$.

Below I will name a few observables that are used in the literature, and in this thesis, to estimate cluster masses.

Integrated SZ signal

As mentioned earlier, the hot ICM of galaxy clusters interacts with photons from the CMB via inverse Compton scattering. This interaction has a magnitude which is proportional to the Compton y-parameter, y, and to the total integrated SZ signal (over the cluster extent), $Y_{\rm SZ}$, which is defined as

$$Y_{\rm SZ} D^2_{\rm A} = \frac{\sigma_{\rm T}}{m_e c^2}\int P_e \mathrm{d}V \tag{1.41}$$

where $D_{\rm A}$ is the angular distance. Since the pressure is linked to the temperature and density of the gas ($P_e = n_e k_B T_e$), which are themselves functions of the total cluster mass, it is expected that $Y_{\rm SZ} D^2_{\rm A}$ is tightly related to the mass with low-intrinsic scatter (Barbosa et al., 1996; Motl et al., 2005; Nagai et al., 2007).

Richness

The optical richness, often characterized as λ, of a cluster is roughly referring to the number of galaxies in a cluster, computed through different means (e.g., Zwicky et al., 1961; Rozo et al., 2009a; Rykoff et al., 2008, 2014). The richness of clusters have been shown to follow a tight correlation with their masses (Yee & Ellingson, 2003). For example, Gladders et al. (2007), using a sample of ∼1000 optically selected galaxy clusters, estimated constrains on the cosmological parameters of $\Omega_m = 0.31^{+0.11}_{-0.10}$ and $\sigma_8 = 0.67^{+0.18}_{-0.13}$, following the self-calibrating method from Majumdar & Mohr (2003) using a mass-richness scaling relation.

Throughout this thesis I will use the richness definition from Klein et al. (2018), where richness is defined as the excess of the weighted sum of galaxies (within a given radius), with respect to the expected number galaxies in the absence of a cluster.

X-ray luminosity, flux and count rate

As noted in the section earlier, the total X-ray luminosity emitted by galaxy clusters comes from thermal bremsstrahlung (free-free) emission, which is proportional to the temperature, gas volume and gas density, such that

$$L_X \propto \rho_g^2 r_{\Delta_c}^3 T^{1/2} \propto \rho_g^2 M_{\Delta_c} T^{1/2}, \tag{1.42}$$

where the mass is related to the luminosity by the radius (Sarazin, 1986). If we define the gas density within a sphere of radius $r^3_{\Delta_c}$ so that $\rho_g \sim M_g r^{-3}_{\Delta_c}$, with M_g being the gas mass, and define the gas fraction as $f_g = M_g/M_{\Delta_c}$, then we can also relate the gas density to the mass as $\rho_g \sim f_g M_{\Delta_c}$.

Reiprich & Böhringer (2002) used a sample of over 100 galaxy clusters with X-ray luminosities from RASS, finding a a tight relation between the X-ray luminosities of clusters and the virial mass M_{200_c} (derived using intracluster gas density profiles and temperatures). Fitting a log-normal scaling relation they find a intrinsic scatter of $\sigma_{\ln L_X} \approx 0.3$.

Moreover, the X-ray luminosity is related to its flux by the following relation

$$L_X = 4\pi D_L^2 f_X \tag{1.43}$$

where D_L^2 is the luminosity distance to the cluster, $D_L = (1+z)r(z)$, which accounts for the cosmological distance (comoving distance; $r(z)$) between the observer and the source.

Another X-ray observable are the count rates (commonly referred to as ctr, C_R or η), which are defined as the number of X-ray photons observed divided by the exposure time, $\eta = n_\gamma t_{\rm exp}$. X-ray fluxes and count rates are tightly correlated, with a dependency on other quantities such as the hydrogen column density (N_H) or the cluster gas temperature (Cao et al., 1999; Reiprich & Böhringer, 2002). Thus, X-ray fluxes, count rates and luminosities are all good proxies of cluster masses.

1.4 Thesis outline

The previous sections are meant to give an introduction to what this thesis aims to do: understand the usage of galaxy clusters to probe cosmology, from the creation of a cluster sample to the development and application of a proper likelihood model.

In Chapter 2, my collaborators and I analyzed optical photometric and spectroscopic data, along with X-ray and SZ observations, of the merging galaxy cluster SPT-CL J0307-6225. I identified new galaxy members, separated the two merging systems, estimated the masses of the substructures, and characterized the merger orbit by comparing with simulations. Additionally, I analyzed the galaxy population and its relation to the merging environment.

In Chapter 3 I, along with collaborators, created the PSZ-MCMF catalogue, with over 1000 galaxy clusters with a purity > 90% and a completeness of 80% (Hernández-Lang et al., 2022). PSZ-MCMF is SZ selected and optically confirmed, where the SZ selection was done by collaborators. I modified and used the MCMF cluster confirmation algorithm to define the sample and understand it's properties.

In Chapter 4, we created an all sky, X-ray selected cluster catalogue: RASS-MCMF (Klein et al., 2023). For the catalogue creation, I calibrated the different photometric surveys used for the optical confirmation of the X-ray candidates. My main contribution however, was on the forecast of the constraining power of such a sample (see § 4.6) using cluster number counts. For this I created a mock cluster catalogue that follows the properties of RASS-MCMF, and then I followed a similar framework as in previous works (e.g., Bocquet et al., 2023, 2024; Mazoun et al., 2024; Vogt et al., 2024) to get constrains on the cosmological parameters σ_8, Ω_m and w.

Finally, in Chapter 5 I re-derive the likelihood model that we used in Chapter 4. I explain in details how the new likelihood is derived while also explaining the differences with our previous model. Additionally, we explicitly implement the weak-lensing mass calibration likelihood (Singh et al., 2024). I generate new realistic mocks following the RASS-MCMF cosmology sample and present the latest results and validation of the pipeline. I also present a comparison with the latest cosmology results from SPT (Bocquet et al., 2024) and eRASS1 (Ghirardini et al., 2024).

CHAPTER 2

CLASH OF TITANS: A MUSE DYNAMICAL STUDY OF THE EXTREME CLUSTER MERGER SPT-CL J0307-6225

Hernández-Lang, D., Zenteno, A., Diaz-Ocampo, A., Cuevas, H., Clancy, J., Prado P., H., Aldás, F., Pallero, D., Monteiro-Oliveira, R., et al., *Monthly Notices of the Royal Astronomical Society* 517, 4355-4378, 2022

We present MUSE spectroscopy, Megacam imaging, and *Chandra* X–ray emission for SPT-CL J0307-6225, a z=0.58 major merging galaxy cluster with a large BCG-SZ centroid separation and a highly disturbed X–ray morphology. The galaxy density distribution shows two main overdensities with separations of 0.144 and 0.017 arcmin to their respective BCGs. We characterize the central regions of the two colliding structures, namely 0307-6225N and 0307-6225S, finding velocity derived masses of $M_{200,N} = 2.44 \pm 1.41 \times 10^{14}$ $M_\odot$ and $M_{200,S} = 3.16 \pm 1.88 \times 10^{14}$ $M_\odot$, with a line-of-sight velocity difference of $|\Delta v| = 342$ km s^{-1}. The total dynamically derived mass is consistent with the SZ derived mass of $7.63\ h_{70}^{-1} \pm 1.36 \times 10^{14}$ $M_\odot$. We model the merger using the Monte Carlo Merger Analysis Code, estimating a merging angle of 36^{+14}_{-12} degrees with respect to the plane of the sky. Comparing with simulations of a merging system with a mass ratio of 1:3, we find that the best scenario is that of an ongoing merger that began $0.96^{+0.31}_{-0.18}$ Gyr ago. We also characterize the galaxy population using Hδ and [OII] λ3727 Å lines. We find that most of the emission-line galaxies belong to 0307-6225S, close to the X–ray peak position, with a third of them corresponding to red-cluster sequence galaxies, and the rest to blue galaxies with velocities consistent with recent periods of accretion. Moreover, we suggest that 0307-6225S suffered a previous merger, evidenced through the two equally bright BCGs at the center with a velocity difference of ~674 km s^{-1}.

2.1 Introduction

Galaxy clusters are located at the peaks of the (dark) matter density field and, as they evolve, they accrete galaxies, galaxy groups, and other clusters from the cosmic web. Some of those merging events are among the most energetic and violent events in the Universe, releasing energies up to 10^{64} ergs (Sarazin, 2002, 2004), providing extreme conditions to study a range of phenomena, from particle physics (e.g. Markevitch et al., 2004; Harvey et al., 2015; Kim et al., 2017) to cosmology (e.g. Clowe et al., 2006; Thompson et al., 2015), including galaxy evolution (e.g. Ribeiro et al., 2013; Zenteno et al., 2020).

The cluster assembly process affects galaxies via several physical processes, including harassment, galaxy-galaxy encounters (e.g., Toomre & Toomre, 1972), tidal truncation, starvation, and ram pressure stripping (Gunn & Gott, 1972), which act upon the galaxies at different cluster centric distances (e.g., Treu et al., 2003). Such events not just change the galaxies in terms of stellar populations and morphologies (e.g., Kapferer et al., 2009; McPartland et al., 2016; Poggianti et al., 2016; Kelkar et al., 2020), but also by destroying them, as indicated by a Halo Occupation Number index lower than 1 (e.g., Lin et al., 2004; Zenteno et al., 2011, 2016; Hennig et al., 2017).

In such extreme environments, galaxies are exposed to conditions that may quench (e.g. Poggianti et al., 2004; Pallero et al., 2022) or trigger star formation (e.g. Ferrari et al., 2003; Owers et al., 2012). For example, Kalita & Ebeling (2019) found evidence of a Jellyfish galaxy in the dissociative merging galaxy cluster A1758N ($z \sim 0.3$), concluding that it suffered from ram-pressure striping due to the merging event. Pranger et al. (2014) studied the galaxy population of the post-merger system Abell 2384 (z∼0.094), finding that the population of spiral galaxies at the center of the cluster does not show star formation activity, and proposing that this could be a consequence of ram-pressure stripping of spiral galaxies from the field falling into the cluster. Ma et al. (2010) discovered a fraction of lenticular post-starburst galaxies in the region in-between two colliding structures, in the merging galaxy cluster MACS J0025.4-1225 (z∼0.59), finding that the starburst episode occurred during the first passage (∼0.5-1 Gyr ago), while the morphology was already affected, being transformed into lenticular galaxies because of either ram-pressure events or tidal forces towards the central region.

On the other hand, Yoon & Im (2020) found evidence of increase in the star formation activity of galaxies in merging galaxy clusters, alleging that it could be due to an increment of barred galaxies in this systems (Yoon et al., 2019). Stroe et al. (2014) found an increase of Hα emission in star-forming galaxies in the merging cluster "Sausage"(CIZA J2242.8+5301) and, by comparing the galaxy population with the more evolved merger cluster "Toothbrush" (1RXS J0603.3+4213), concluded that merger shocks could enhance the star formation activity of galaxies, causing them to exhaust their gas reservoirs faster (Stroe et al., 2015). Furthermore, Stroe et al. (2017) using a sample of 19 clusters, at $0.15 < z < 0.31$, found excess of Hα emission in merging clusters with respect to relaxed cluster, specially closer to the cluster's core. Such results were further confirmed with an spectroscopic examination of 800 Hα-selected cluster galaxies (Stroe & Sobral, 2021).

To understand how the merger process impacts cluster galaxies, it is crucial to assemble large samples of merging clusters and determine their corresponding merger phase: pre, ongoing

or post. The SZ-selected samples are ideal among the available cluster samples, as they are composed of the most massive clusters in the Universe and are bound to be the source of the most extreme events. The South Pole Telescope (SPT, Carlstrom et al., 2011) has completed a thermal SZ survey, finding 677 cluster candidates (Bleem et al., 2015), providing a well understood sample to study the impact of cluster mergers on their galaxy population. There is rich available information on those clusters, including the gas centroids (via SZ and/or X–ray), optical imaging, near-infrared imaging, cluster masses, photometric redshifts, etc. Furthermore, as the SPT cluster selection is nearly independent of redshift, a merging cluster sample will also allow evolutionary studies to high redshifts.

Using SPT-SZ selected clusters and optical imaging, Song et al. (2012b) reported the brightest cluster galaxy (BCG) positions on 158 SPT cluster candidates and, by using the separation between the cluster BCG and the SZ centroid as a dynamical state proxy, found that SPT-CL J0307-6225 is the most disturbed galaxy cluster of the sample, i.e., with the highest separation. Recently, Zenteno et al. (2020) employed optical data from the first three years of the Dark Energy Survey (DES, Abbott et al., 2018; Morganson et al., 2018; Collaboration: et al., 2016) to use the BCG in 288 SPT SZ-selected clusters (Bleem et al., 2015) to classify their dynamical state. They identified the 43 most extreme systems, all with a separation greater than 0.4 r_{200}, including once again SPT-CL J0307-6225.

SPT-CL J0307-6225 is a merger candidate at $z = 0.5801$ (Bayliss et al., 2016), with a mass estimate from SPT data of $M_{500} = 5.06 \pm 0.90 \times 10^{14} h_{70}^{-1}$ $M_\odot$ (Bleem et al., 2015). SPT-CL J0307-6225 has (1) *gri* optical data observed with the Megacam instrument on the Magellan Clay telescope (Chiu et al., 2016), (2) X–ray data obtained with the *Chandra* telescope (McDonald et al., 2013), and (3) spectroscopic information taken with the Gemini Multi-Object Spectrograph (GMOS; Bayliss et al., 2016). Dietrich et al. (2019) used the Megacam data to measure the weak lensing mass density and, although the cluster was observed under the best seeing conditions in the sample (0.55-0.65 arcsec), the resulting WL mass distribution is of low significance, with the recovered center located away from the gas distribution or the galaxies (see their Fig. B.4).

In the absence of precise WL measurements, the galaxy-gas offset can be used to constrain self-interacting dark matter models as shown by Wittman et al. (2018). The separation between the X–ray centroid of SPT-CL J0307-6225, estimated using *Chandra* data (McDonald et al., 2013), and the BCG (Zenteno et al., 2020) is 1.98 arcmin (~790 kpc). This would be the largest gas-galaxy offset within the Wittman et al. (2018) sample of merging galaxy clusters, implying a high potential for SPT-CL J0307-6225 to constrain such models. Using GMOS spectroscopic data, Bayliss et al. (2016) studied the velocity distribution of the SPT-GMOS sample (62 galaxy clusters), finding SPT-CL J0307-6225 to be one of the 9 clusters with a non-Gaussian (i.e., disturbed) velocity distribution (2-σ level). Nurgaliev et al. (2017) used the *Chandra* data to make an estimate of the X–ray asymmetry for this system, finding it to be the second[1] most asymmetric system in the full SPT-Chandra sample (over 90 galaxy clusters), with an X–ray morphology as disturbed as El Gordo, a well-known major merger (Williamson et al., 2011;

[1]In Zenteno et al. (2020), the most asymmetric system, SPT-CL J2332-5053, was said to be a cluster in pre-merger state with a close companion, which would then contaminate the estimated asymmetry index. Excluding SPT-CL J2332-5053 would make SPT-CL J0307-6225 the most asymmetric system in the sample.

Menanteau et al., 2012), making this cluster an interesting system to test the impact of a massive merging event in galaxy evolution, the goal of this paper.

We use VLT/MUSE integral field and Gemini/GMOS spectroscopy, X–ray data from *Chandra*, and Megacam optical imaging to characterize the SPT-CL J0307-6225 merger stage, and its impact on galaxy population. The paper is organized as follow: in §2.2 we provide details of the observations and data reduction. In §2.3 we show the analysis for the spectroscopic and optical data, while in §2.4 we report our findings for both the merging scenario and the galaxy population. In §2.5 we propose an scenario for the merging event and connect it to the galaxy population. In §2.6 we give a summary of the results. Throughout the paper we assume a flat Universe, with a ΛCDM cosmology, $h = 0.7$, $\Omega_m = 0.27$ (Komatsu et al., 2011). Within this cosmology, 1 arcsec at the redshift of the cluster ($z \approx 0.58$) corresponds to ∼6.66 kpc.

2.2 Observations and Data Reduction

2.2.1 Optical Imaging

Chiu et al. (2016) obtained optical images using Magellan Clay with Megacam during a single night on November 26, 2011 (UT). They reduced and calibrated the data following High et al. (2012). Megacam has a 24′x 24′field-of-view, which at redshift ∼0.58 correspond to ∼10 Mpc. Several dithered exposures were taken in g, r, and i filters for a total time of 1200 s, 1800 s, and 2400 s respectively. The median seeing of the images was approximately 0.79 arcsec or about 5 kpc, with a better seeing in r-band, averaging 0.60 arcsec. The 10σ limit magnitudes in gri are 24.24, 24.83, and 23.58, respectively (Chiu et al., 2016). In Fig. 2.1 we show the gri pseudo-color image, centered on the SZ cluster position of SPT-CL J0307-6225, with the white bar on the bottom right showing the corresponding scale.

The catalogs for the photometric calibration were created following High et al. (2012) and Dietrich et al. (2019) including standard bias subtraction and bad-pixel masking, as well as flat fielding, illumination, and fringe (for i-band only) corrections. To calibrate the zeropoint of the data, the stellar locus regression technique was used (High et al., 2009), together with constraints by cross-matching with 2MASS catalogs (Skrutskie et al., 2006), giving uncertainties in absolute magnitude of 0.05 mag and in color of 0.03 mag (Desai et al., 2012; Song et al., 2012b).

For the creation of the galaxy photometric catalogs, we use a combination of Source Extractor (`SExtractor`; Bertin & Arnouts, 1996) and the Point Spread Function Extractor (PSFEx; Bertin, 2011) softwares. SExtractor is run in dual mode, using the i-band image as the reference given the redshift of the cluster[2]. We extract all detected sources with at least 6 pixels connected above the 4σ threshold, using a 5 pix Gaussian kernel. Deblending is performed with 64 sub-thresholds and a minimum contrast of 0.0005. Galaxy magnitudes are SExtractor's `MAG_AUTO` estimation, whereas colors are derived from aperture magnitudes.

The star-galaxy separation in our sample is performed following Drlica-Wagner et al. (2018), by using the SExtractor parameter SPREAD_MODEL, and its corresponding error, SPREAD-ERR_MODEL, derived from the i-band image, for objects within R_{200} from the SZ center (R_{200} =

[2]At $z \approx 0.58$, the i-band is located redwards the 4000Å break.

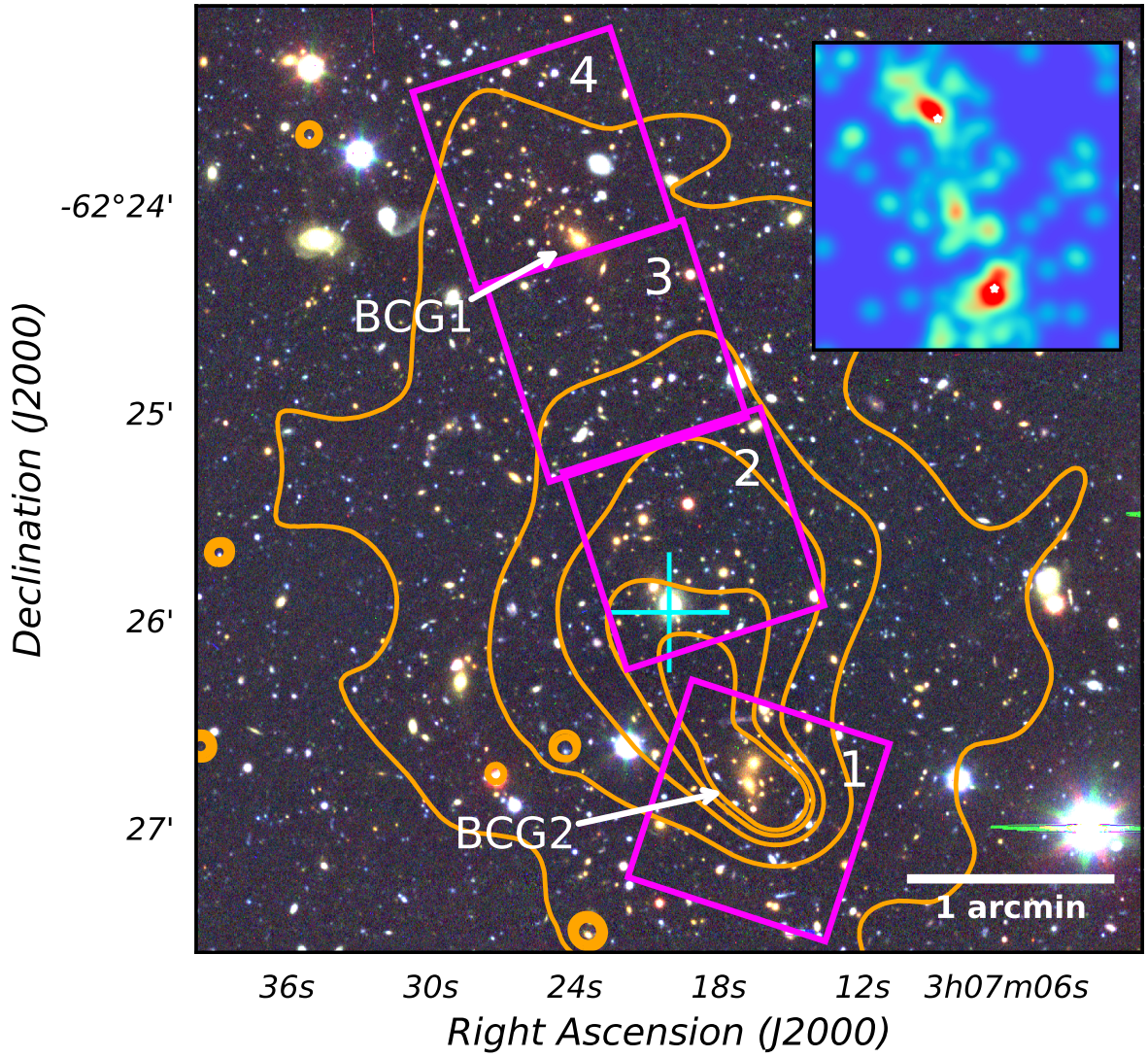

Figure 2.1: Pseudo-color image, from *gri* filters combination, of the central area of SPT-CL J0307-6225. Magenta squares show the MUSE footprints, where the numbers on the top-right corner of each square shows the cube's number. Orange contours where derived from archival *Chandra* images. The cyan plus-sign marks the X–ray centroid (McDonald et al., 2013). The arrows show the positions of the two brightest galaxies of the cluster. The white bar on the bottom shows the scale of 1 arcmin. The inset shows the 2D galaxy number density (which matches the size of the main figure), where the two highest intensity areas correspond to the areas around the BCGs, which are shown as white stars.

3.84′; Song et al., 2012b; Zenteno et al., 2020). Drlica-Wagner et al. (2018) classified a source as a star if it satisfies

$$|\text{SPREAD\_MODEL} + \left(\frac{5}{3}\right) \times \text{SPREADERR\_MODEL}| < 0.002 \tag{2.1}$$

With this, we remove stars from our catalogue and, to improve upon this selection, we apply a magnitude cut, such that $i_{\rm auto} < 18.5$ mag, which is ~ 0.5 mag brighter than the BCG. On the faint end the cut is set at $i_{\rm auto} < m^* + 3 = 23.39$, which is beyond the limit of our spectroscopic catalogue (see Appendix 2.7). With this we obtain 639 photometric galaxies.

2.2.2 Spectroscopic data

2.2.2.1 MUSE data

The Multi Unit Spectroscopic Explorer (MUSE, Bacon et al., 2012) observations were taken on August 22nd, 23rd and 24th, 2016 (program id: 097.A-0922(A), PI: Zenteno), and November 10 and December 20, 2017 (program id: 100.A-0645(A), PI: Zenteno). The observations consisted of four pointings, with a total exposure time of 1.25 hours per data cube, with an airmass = 1.4 (see Table 2.1). MUSE in nominal mode covers the wavelength range 4800-9300 Å, with resolution of 1700 < R < 3500, covering redshifted emission lines such as [OII] λ3727 Å and [OIII] λ5007 Å, as well as absorption lines such as the Hydrogen Balmer series Hδ, Hγ and Hβ. The positions of the pointings were selected to cover the two BCGs (labeled as BCG1 and BCG2 on Fig. 2.1) and the area between them. The MUSE footprints for the 4 observed data cubes are shown as magenta squares on Fig. 2.1, with the cubes enumerated in the top right corner of each square. We use these numbers to refer to the cubes throughout the paper.

Table 2.1: Central coordinates and seeing conditions of the observed MUSE fields.

CUBE	Program ID	Coordinates R.A. (J2000)	Coordinates Dec. (J2000)	Seeing (arcsec)
1	097.A-0922(A)	$03^{\rm h}\ 07^{\rm m}\ 16.34^{\rm s}$	$-62^\circ\ 26'\ 54.98''$	0.56
2	097.A-0922(A)	$03^{\rm h}\ 07^{\rm m}\ 19.052^{\rm s}$	$-62^\circ\ 25'\ 36.430''$	0.70
3	0100.A-0645(A)	$03^{\rm h}\ 07^{\rm m}\ 22.271^{\rm s}$	$-62^\circ\ 24'\ 42.140''$	0.68
4	0100.A-0645(A)	$03^{\rm h}\ 07^{\rm m}\ 25.302^{\rm s}$	$-62^\circ\ 23'\ 46.570''$	0.97

The data was taken in WFM-NOAO-N mode, with a position angle of 18 deg for three of the cubes and 72 deg for the one to the south, and using the dithering pattern recommended for best calibration: 4 exposures with offsets of 1 arcsec and 90 degrees rotations (MUSE User Manual ver. 1.3.0). The raw data were reduced through the MUSE pipeline (Weilbacher et al., 2014, 2016) provided by ESO.

We construct 1D spectra from the MUSE cube using the `MUSELET` software (Bacon et al., 2016). `MUSELET` finds source objects by constructing line-weighted (spectrally) 5x1.25 Å wide narrow band images and running `SExtractor` on them. In order to create well fitted masks to their respective sources, the parameter `DETECT_THRESH` is set to be 2.5. If the chosen value is below that, `SExtractor` will detect noise and output wrong shapes in the segmentation map. We proceed to use the source file to extract the `SExtractor` parameters `A_WORLD`, `B_WORLD` and `THETA_WORLD` to create an elliptical mask centered in each source.

Finally, we use the `MUSELET` routines `mask_ellipse` and `sum` to create the 1D weighted spectra of the sources. To make sure the objects fit into their apertures, the SExtractor parameter

PHOT_FLUXFRAC is set at 0.9, which means that 90% of the source's flux will be contained within the mask's radius.

2.2.2.2 GMOS data

We complement MUSE redshifts with Gemini/GMOS data published by Bayliss et al. (2016). The Bayliss galaxy redshift sample consists in 35 galaxies redshifts, with 8 not present in our MUSE data. The spectroscopic data from their sample can be found online at the VizieR Catalogue Service (Ochsenbein et al., 2000), with the details on the data reduction described in Bayliss et al. (2016) and Bayliss et al. (2017). For SPT-CL J0307-6225, they used 2 spectroscopic masks with an exposure time of 1 hour each. The target selection consisted mostly of galaxies from the red sequence (selected as an overdensity in the color-magnitude and color-color spaces) up to $m^* + 1$, prioritising BCG candidates.

2.2.3 X–ray data

SPT-CL J0307-6225 was observed by *Chandra* as part of a larger, multi-cycle effort to follow up the 100 most massive SPT-selected clusters spanning $0.3 < z < 1.8$ (McDonald et al., 2013, 2017). In particular, this observation (12191) was obtained via the ACIS Guaranteed Time program (PI: Garmire). A total of 24.7 ks was obtained with ACIS-I in VFAINT mode, centering the cluster $\sim 1.5'$ from the central chip gap. The data was reprocessed using CIAO v4.10 and CALDB v.4.8.0. For details of the observations and data processing, see McDonald et al. (2013). The derived X–ray centroid is shown as a cyan plus-sign on Fig. 2.1.

An image in the 0.5–4.0 keV bandpass was extracted and adaptively smoothed using CSMOOTH[3]. This smoothed image, shown as orange contours in Fig. 2.1, reveals a highly asymmetric X–ray morphology, with a bright, dense core offset from the large-scale centroid by $\sim 1'$ ($\sim$400 kpc).

2.3 Analysis

2.3.1 Color-Magnitude Diagram and RCS selection

The color-magnitude diagram (CMD) for the cluster is shown in Fig. 2.2, where the magenta triangles and the blue squares are galaxies from the MUSE and GMOS spectroscopic samples, respectively, and the dots represent galaxies from our photometric sample (selected as described in § 2.2.1). For the selection of the red cluster sequence (RCS) galaxies, which consist mostly of passive galaxies which are likely to be at the redshift of the cluster (Gladders & Yee, 2000), we examine the location of the galaxies from our spectroscopic sample in the CMD. We then select all galaxies with $r - i > 0.65$ and perform a 3σ-clipping cut on the color index to remove outliers. We keep all the galaxies from our previous magnitude cut in § 2.2.1 ($i_{\rm auto} < 23.39$). Finally, we fit a linear regression to the remaining objects, which is shown with a red dashed line in Fig. 2.2. The green dotted lines denote the limits for the RCS, chosen to be ± 0.22 [mag] from the fit, which

[3] https://cxc.harvard.edu/ciao/ahelp/csmooth.html

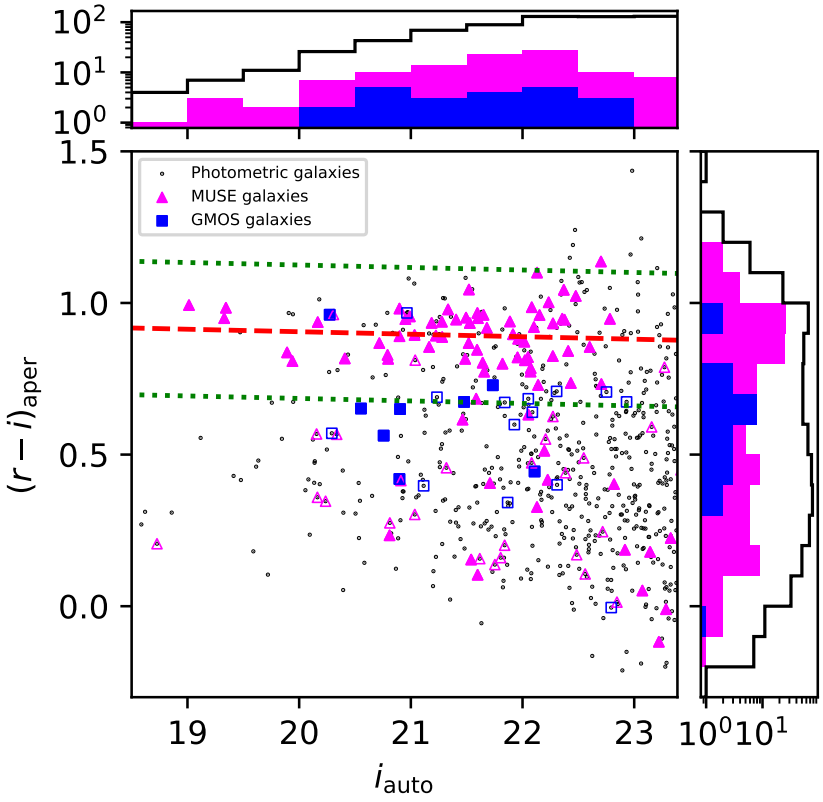

Figure 2.2: Color-magnitude diagram (CMD) of SPT-CL J0307-6225 from Megacam data within R_{200}. The y-axis shows the color index $r-i$ estimated from aperture magnitudes, with a fix aperture of ~40 kpc (~6 arcsec) at the cluster redshift, while the x-axis shows SExtractor's `MAG_AUTO`. Magenta triangles and blue galaxies represent galaxies from our MUSE and GMOS data, respectively, filled for those that belong to the cluster, whereas black dots are galaxies from our photometric sample. The red cluster sequence (RCS) estimated for the cluster is shown as a red-dashed line, while the green dotted lines are the 0.22 mag width established for the RCS.

corresponds to the average scatter of the RCS at 3σ (López-Cruz et al., 2004). This gives us a total of 210 optically selected RCS galaxy candidates, with 64 of those being spectroscopically confirmed members.

2.3.2 Spectroscopic catalog

2.3.2.1 Galaxy redshifts

To obtain the redshifts, we use an adapted version of MARZ (Hinton et al., 2016) for MUSE spectra[4]. MARZ takes the 1D spectra of each object as an input, obtaining the spectral type (late-type galaxy, star, quasar, etc.) and the redshift that best fits as an output. The results are examined visually for each of the objects, calibrating them using the 4000Å break and the

[4] `http://saimn.github.io/Marz/#/overview` (Hinton, private communication)

Calcium H and K lines. Heliocentric correction was applied to all redshifts using the RVCORRECT task from IRAF. The upper panel of Fig. 2.3 shows the stacked spectra of a couple of blue and red galaxies.

There are three sources in the cube 4 region which appeared to be part of the cluster, but were not well fitted by MARZ. These sources are shown in the bottom panel of Fig. 2.3, with their spectra shown in black and the cutouts of the galaxies in the left. The cyan spectra shows a galaxy with an estimated redshift higher than that of the cluster but with a $r - i$ color within our RCS selection. We manually estimate the redshifts of these 3 sources using MARZ.

In total we estimate spectroscopic redshifts for 117 objects within the MUSE fields, with 4 of them classified as stars. In Table 2.8 we show the redshifts and magnitudes for this objects. For details of the different columns please refer to Appendix 2.9.

In Table 2.8 we show the properties of 22 objects from GMOS, excluding the 12 in common with MUSE and the potential cluster member from our measured redshifts. In Appendix 2.8 we give further details into the estimation of the GMOS spectra redshifts, the comparison to our estimates with MUSE and the exclusion of potential members. GMOS redshifts in Table 2.8 correspond to the ones measured using FXCOR. Our final spectroscopic catalog is composed of 139 objects; 134 galaxies and 5 stars.

2.3.2.2 Cluster redshift estimation

The cluster's redshift is estimated following the biweight average estimator from Beers et al. (1990), using the median redshift from all objects with measured redshift in our sample. This estimated redshift is then used instead of the median in their equation, in order to estimate a new redshift. This process is iterated 3 times. We select only spectroscopic sources with a peculiar velocity within ±5000 km s^{-1} from the cluster's estimated redshift, in order to exclude most of the foreground and background objects (eg. Bösch et al., 2013; Pranger et al., 2014). We then estimate the velocity dispersion (σ_v) using the biweight sample variance presented in Ruel et al. (2014), so that

$$\sigma_{\rm bi}^2 = N \frac{\sum_{|u_i|<1}(1-u_i^2)^4(v_i-\bar{v})^2}{D(D-1)} \tag{2.2}$$

$$D = \sum_{|u_i|<1}(1-u_i^2)(1-5u_i^2) \tag{2.3}$$

where the peculiar velocities of the galaxies, v_i, and the biweight weighting, u_i, are estimated as

$$v_i = \frac{c(z_i - z_{\rm cl})}{1+z_{\rm cl}} \tag{2.4}$$

$$u_i = \frac{v_i - \bar{v}}{9{\rm MAD}({\rm v_i})} \tag{2.5}$$

with c being the speed of light, MAD corresponds to the median absolute deviation and z_i, z being the redshifts of the galaxies and the biweight estimation of the redshift of the sample, respectively. Then, the velocity dispersion is estimated as the square root of $\sigma_{\rm bi}^2$, with its uncertainty estimated

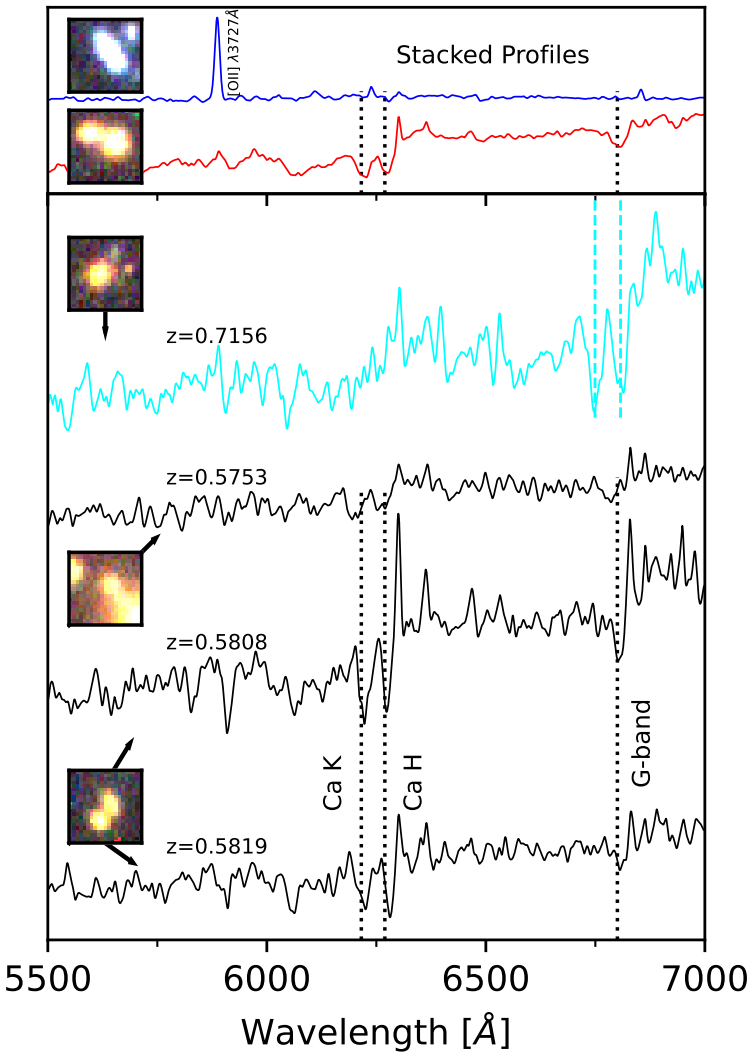

Figure 2.3: *top*: Stacked spectra of a couple of blue and red galaxies at the cluster's redshift, shown in blue and red, respectively. The cutout on the left shows an example of a galaxy from each profile. Black dotted lines mark the Calcium H and K lines, together with the G-band feature at 4304 Å, redshifted to $z = 0.58$. *bottom*: Spectrum of the sources with redshifts estimated manually (black) and that of a galaxy with similar characteristic to those of the cluster, but at $z = 0.716$. A small cutout of 5×5 arcsec^2 is shown on the left for each galaxy, with a black arrow pointing at the respective spectra. The redshift found with MARZ of each source is written on top of each spectrum. Dotted lines are the same as in the upper panel, with the cyan dashed lines marking the Calcium H and K lines redshifted to $z = 0.716$.

as $0.92\sigma_{\rm bi} \times \sqrt{N_{\rm members} - 1}$. To obtain a final redshift for the cluster we use a 3σ-clipping iteration (with $\sigma = \sigma_v$), obtaining $z_{\rm cl} = 0.5803 \pm 0.0006$, where the error is estimated as the standard error, i.e., the standard deviation over the square root of the number of cluster members.

2.3.2.3 Cluster member selection

Observationally, galaxies belonging to a cluster are selected by imposing restrictions on their distance to the center of the cluster and their relative velocities to the BCG. In this section, we studied the appropriate cut in the Line of Sight (LoS) velocity for a theoretical cluster with the same mass and the same redshift than SPT-CLJ0307-6225 using the Illustris TNG300 simulations. Illustris TNG is a suite of cosmological-magnetohydrodynamic simulation which aims to study the physical processes that drive galaxy formation (Nelson et al., 2018; Pillepich et al., 2018; Springel et al., 2018; Naiman et al., 2018; Marinacci et al., 2018). We used the TNG300 because it is the simulation with the largest volume, having a side length of $L \sim 250h^{-1}$ Mpc. This volume contains 2000^3 Dark Matter (DM) particles and 2000^3 baryonic particles. The relatively large size of the simulated box allow us to identify a significant number of massive structures. The mass resolution of TNG300 is $5.9 \times 10^7 M_\odot$, and $1.1 \times 10^7 M_\odot$ for the DM and baryonic matter respectively. Also, the adopted softening length is 1 h^{-1} kpc for the DM particles and 0.25 h^{-1} kpc for the baryonic particles (Marinacci et al., 2018).

This simulation have a total of 1150 structures with masses between $10^{14} M_\odot \leq M_{200} \leq 9 \times 10^{14} M_\odot$, in a redshift range $0.1 \leq z \leq 1$. Here M_{200} is the mass within a sphere having a mean mass density equal to 200 times the critical density of the Universe. To ensure that our results are not affected by numerical resolution effects, we only selected subhalos with at least 1000 dark matter particles per galaxy ($M_{\rm DM} \geq 5.9 \times 10^{10} M_\odot$) and at least 100 stellar particles ($M_{\rm stellar} \geq 1.1 \times 10^9 M_\odot$).

We used the criteria proposed by Zenteno et al. (2020) to divide the clusters according their virialization stage. We consider a that a cluster is disturbed when the offset between the position of the BCG and the center of mass of the gas is greater than $0.4 \times R_{200}$ (used as a proxy for the Sunayev-Zeldovich effect) otherwise, we consider them as relaxed. The final sample used in this work is composed by the 150 relaxed clusters and 150 disturbed clusters.

To stack information from the selected clusters we normalize the velocity distributions using the $\sigma_v - M_{200}$ scaling relation from Munari et al. (2013). This scaling relation was obtained from a radiative simulation which included both (a) star formation and supernova triggered feedback, and (b) active galactic nucleus feedback (which they call the AGN-set). The equation is described as follows:

$$\sigma_{\rm 1D} = A_{\rm 1D} \left[\frac{h(z) M_{200}}{10^{15} M_\odot} \right]^\alpha \tag{2.6}$$

where $\sigma_{\rm 1D}$ is the one-dimensional velocity dispersion and h(z) = H(z)/100 km s^{-1} Mpc^{-1}. We choose the values of $A_{\rm 1D} = 1177 \pm 4.2$ and $\alpha = 0.364 \pm 0.0021$, obtained using galaxies associated to subhaloes in the AGN-set simulation (Munari et al., 2013).

To find the intrinsic Line of Sight (LoS) velocity distribution of a simulated cluster with mass $M_{200} = 5 \times 10^{14} M_\odot$, at a given redshift of $z = 0.6$, we followed the following procedure. We first fit the projected 1D velocity distribution of the cluster galaxies relative to the BCG using a Gaussian distribution with mean μ_0 and dispersion σ_0. After, using the Equation 2.6, we compute the value of the 1D velocity dispersion σ_1 that the cluster would have if it had a mass of $M_{200} = 5 \times 10^{14} M_\odot$. Then, we obtain the 1D velocities for each galaxy normalized by the

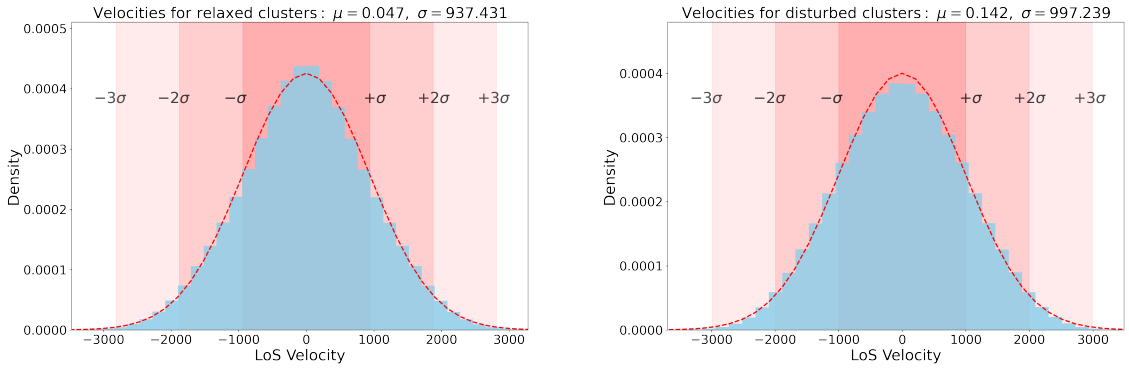

Figure 2.4: Histogram for the LoS satellite velocities distribution for relaxed (left) and disturbed (right) clusters with masses $M_{200} = 7.64 \times 10^{14} M_{\odot}$ at redshift $z = 0.6$, in red the fitted normal distribution and in light red the confidence intervals.

mass and the redshift using the equation 2.7. Finally, we obtained the LoS velocities applying 200 different randomized rotations to each cluster,

$$z = \sigma_1 \left(\frac{x - \mu_0}{\sigma_0} + \mu_0 \right). \tag{2.7}$$

Fig. 2.4 presents the histogram in the LoS velocity for the galaxies associated to the 150 relaxed (top) and disturbed (bottom) clusters stacked in different projections (blue histogram), the best fit normal distribution (red dashed line) and the confidence intervals shaded red areas. We conclude that for a relaxed cluster with mass of $M_{200} = 7.64 \times 10^{14}$ the LoS velocity is distributed with a dispersion $\sigma_v = 940 km\ s^{-1}$. For disturbed clusters the velocities are normally distributed with a dispersion of $\sigma = 1000 km\ s^{-1}$. This means that 95% of the galaxies belonging to a disturbed cluster with $M_{200} = 7.64 \times 10^{14}$ would have LoS velocities lower than 2000 km s^{-1}, and 99% of them have LoS velocities lower than 3000 km s^{-1}. In what follows we adopt a cut of 3,000 km s^{-1}. Our results shows that the distribution of LoS velocity is not significantly affected by the virialized status of the studied cluster.

Applying the ± 3,000 km/s cut we obtain a total number of cluster redshifts of 87, including 25 members from cube 1, 21 from cube 2, 11 from cube 3, 22 from cube 4 and 8 from the GMOS data.

2.3.2.4 Summary of spectroscopic catalog

In total, we obtain 87 galaxies with spectroscopic redshifts for SPT-CL J0307-6225. Out of those, 79 come from the 1D MUSE objects from §2.2.2 and 8 from the GMOS archival spectroscopic data (Bayliss et al., 2016). The final redshift, estimated as the biweight average estimator, is $z_{\rm cl} = 0.5803 \pm 0.0006$. The final galaxy cluster redshift distributions is shown in Fig. 2.5. The inset shows the peculiar velocity of these selected galaxies, with the black dashed lines denoting the velocity cut and the black dotted line marking the velocity of the BCG. The velocity dispersion for the cluster, estimated following Eq. 2.2, is $\sigma_v = 1093 \pm 108$ km s^{-1}.

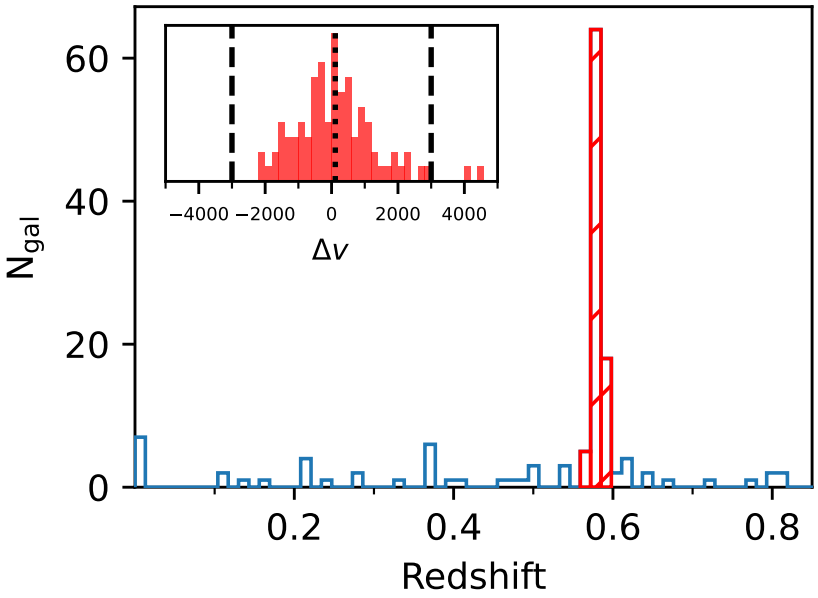

Figure 2.5: Redshift distribution of spectroscopic sources with good measurement from MARZ and FXCOR. Hashed red bars represent the region within a range of ±3000 km s^{-1} in peculiar velocity from the cluster's redshift. The histogram insert on the top left shows the distribution of galaxies within this velocity range, where the black dashed and dotted lines represent the cuts at ±3000 km s^{-1} and the velocity of the BCG, respectively.

Table 2.2: Galaxy population classification.

Type	Criteria
PSB	Galaxies with EW(Hδ) $\geq$ 5 Å and EW(OII) $<$ 5 Å
SSB	Galaxies with EW(Hδ) $<$ 0 Å and EW(OII) $\geq$ 5 Å
EL	Galaxies with EW(OII) $\geq$ 5 Å (SF, SSB and A+em)
NEL	Galaxies with EW(OII) $<$ 5 Å (Passive and PSB)
Red	Galaxies belonging to (or redder than) the RCS
Blue	Galaxies with colors lower than the RCS

2.3.2.5 Spectral classification

To understand if the merger is playing a role in the star formation activity of the galaxies, we make use of two measurements; the equivalent widths (EW) of the [OII] λ3727 Å and Hδ lines. [OII] λ3727 Å traces recent star formation activity in timescales $\leq$10 Myr, while the Balmer line Hδ has a scale between 50 Myr and 1 Gyr (Paulino-Afonso et al., 2020). A strong Hδ absorption line is interpreted as evidence of an explosive episode of star formation which ended between 0.5-1.5 Gyrs ago (Dressler & Gunn, 1983). To measure the equivalent widths of [OII] λ3727 Å, EW(OII), and Hδ, EW(Hδ), the flux spectra for each object is integrated following the ranges

Table 2.3: Fraction of galaxy types at different magnitude ranges. The second column is the total number of galaxies for a given magnitude range, while the third column is the median S/N of the galaxies.

			Photometric		Spectroscopic				
Mag	N_{Total}	S/N	Red	Blue	NEL	EL	Low S/N	SSB	PSB
			%	%	%	%	%	%	%
$i_{auto} < m^*$	6	12.0	100.00	0.00	100.00	0.00	0.00	0.00	0.00
$m^* \leq i_{auto} < m^* + 1$	16	7.8	93.75	6.25	81.25	12.50	6.25	0.00	0.00
$m^* + 1 \leq i_{auto} < m^* + 2$	43	4.0	81.40	18.60	76.75	18.60	4.65	4.65	4.65
$i_{auto} \geq m^* + 2$	14	2.3	50.00	50.00	-	-	-	-	-

Notes. SSB are a subpopulation of the EL galaxies, whereas PSB are a subpopulation of NEL galaxies. The red and blue populations add up to 100% for the photometric classification, while the NEL, EL and Low S/N populations add up to 100% in the spectroscopic classification. We do not use spectral classification for galaxies with $i_{auto} \geq m^* + 2$.

described by Balogh et al. (1999) using the IRAF task SBANDS. Also, we only make use of MUSE galaxies, excluding the 8 GMOS galaxies added, given that the MUSE selection is unbiased. We do not expect this to change our main results since these galaxies are not located along the merger axis.

We use the same scheme defined by Balogh et al. (1999) to classify our galaxies into different categories; passive, star forming (SF), short-starburst (SSB), post-starburst (PSB, K+A in Balogh et al., 1999) and A+em (which could be dusty star-forming galaxies). For this classification we only take into account galaxies with $i_{auto} < m^* + 2$, meaning over 80% completeness (Appendix 2.7), and a signal-to-noise ratio, S/N > 3 (62 galaxies), given that galaxies with low S/N can affect the measurements of lines in crowded sections, like in the region of the [OII] λ3727 Å line (Paccagnella et al., 2019). The median signal-to-noise ratio (S/N) of our MUSE galaxies is shown in Table 2.3 for different magnitude ranges. We estimate the S/N in the entire spectral range of our data by using the DER_S/N algorithm (Stoehr et al., 2007).

For simplicity, we use the following notation (and their combinations) to refer to the different galaxy populations throughout the text; EL for emission-line galaxies (EW(OII) $\geq$ 5 Å), including SSB, star-forming (SF) and A+em, and NEL for non emission-line galaxies (passive and PSB). We also use the RCS selection from § 2.3.1 to separate red and blue galaxies. We also analyze in particular SSB and PSB galaxies. Table 2.2 summarizes the different criteria of each population.

Table 2.3 shows the fraction of galaxies for different magnitude ranges. The fractions are divided by the photometric classification (red or blue) and the spectroscopic classification (EL, NEL, SSB, PSB or Low S/N). Fig. 2.6 shows the sky positions of the galaxy population on top of the X–ray emission map. The results of this classification will be further discussed in §2.4.4.

2.3.3 Galaxies association

One of the most common techniques to estimate the level of substructure in galaxy clusters is to analyze the galaxy velocity distribution on a 1D space, where it is assumed that for a relaxed cluster it should be close to a Gaussian shape (e.g. Menci & Fusco-Femiano, 1996; Ribeiro et al.,

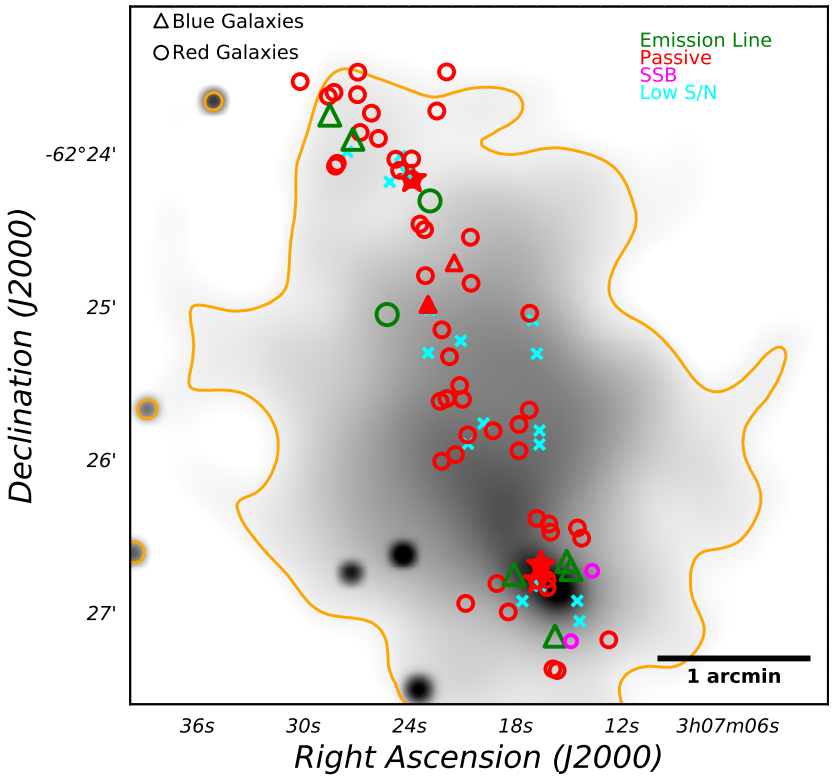

Figure 2.6: Sky positions of the MUSE cluster galaxies on top of the X–ray map, with the orange contour showing the outermost contour in Fig. 2.1. Circles are red galaxies and triangles are blue galaxies, color coded by their spectral type, with cyan crosses being galaxies with S/N < 3 or $i_{\rm auto} \geq m^* + 2$ and the PSB shown as the red filled triangle.

2013). Hou et al. (2009) used Monte Carlo simulations to show that the Anderson-Darling (AD) test is among the most powerful to classify Gaussian (G) and non-Gaussian (NG) clusters.

Hou et al. (2009) use the α value (the significance value of the statistic) to assign the dynamical state of clusters (see Eq. 17 in their paper), where $\alpha < 0.05$ indicates a NG distribution. Nurgaliev et al. (2017) uses the p-value of the statistic ($p_{\rm AD}$) and separates the clusters using $p_{\rm AD} < 0.05/n$ for NG clusters, where n indicates the number of tests being conducted. We divide our data in 4 subsets for the application of the AD test; Cubes 2 and 3 for the middle overdensity, Cubes 1 and 4 to compare the two most overdense regions, all the data cubes and all the data cubes plus GMOS data.

To test for 3D substructures (using the velocities and the on-sky positions), we use the Dressler-Shectman test (DS-test, Dressler & Shectman, 1988), which uses the information of the on-sky coordinates along with the velocity information, and can be used to trace perturbed structures (e.g. Pranger et al., 2014; Olave-Rojas et al., 2018). The DS-test uses the velocity information of the closest (projected) neighbors of each galaxy to estimate a Δ statistic, which is given by

$$\Delta = \sum_{i}^{N_{\rm tot}} \delta_i, \tag{2.8}$$

where $N_{\rm tot}$ corresponds to the total number of members of the cluster and

$$\delta^2 = \frac{N+1}{\sigma_{\rm cl}^2}\left[(\bar{v}_{\rm loc}-\bar{v}_{\rm cl})^2+(\sigma_{\rm loc}-\sigma_{\rm cl})^2\right], \tag{2.9}$$

where δ is estimated for each galaxy. N corresponds to the number of neighbors of the galaxy to use to estimate the statistic, estimated as $N=\sqrt{N_{\rm tot}}$ (Pinkney et al., 1996), $\sigma_{\rm cl}$ and $\sigma_{\rm loc}$ correspond to the velocity dispersion of the whole cluster and the velocity dispersion of the N neighbors, respectively, and $\bar{v}_{\rm cl}$ and $\bar{v}_{\rm loc}$ correspond to the mean peculiar velocity of the cluster and the mean peculiar velocity of the N neighbors, respectively. A value of $\Delta/N_{\rm tot} \leq 1$ implies that there are no substructures on the cluster.

To calibrate our DS-test results, we perform 10^4 Monte Carlo simulations by shuffling the velocities, i.e., randomly interchanging the velocities among the galaxies, while maintaining their sky coordinates (meaning that the neighbors are always the same). The p-value of the statistic (p_Δ) is estimated by counting how many times the simulated Δ is higher than that of the original sample, and divide the result by the total number of simulations. Choosing $p_\Delta < 0.05$ ensures a low probability of false identification (Hou et al., 2012) and is accepted for the distribution to be considered non-random. Both AD and DS test results are shown in Table 2.4.

To test for 2D substructures (sky positions) we build surface density maps (see, e.g., White et al., 2015; Monteiro-Oliveira et al., 2017, 2018, 2020; Yoon et al., 2019). The galaxy surface density map at the top right of Fig. 2.1 implies that there are at least two colliding-structures. To obtain the density map we use the RCS galaxy catalog and the SKLEARN.NEIGHBORS.KERNELDENSITY python module, applying a gaussian kernel with a bandwidth of 50 kpc.

2.4 Results

2.4.1 Cluster substructures

Table 2.4 shows the results of both the AD-test and the DS-test applied to different subsets. The second column corresponds to the number of spectroscopic galaxies belonging to a given subsample. The subset which gives the smallest p-values for both the AD-test and the DS-test is the Cubes 1+4 subset, with these cubes located on top of the two density peaks, enclosing also the area next to the two brightest galaxies (see Fig. 2.1). We find that both the AD-test and the DS-test provide no evidence of substructure. Applying a 3σ-clipping iteration to the samples does not change the results. The results, along the X–ray morphology, show no evidence of substructure along the line of sight, and rather support a merger in the plane of the sky, thus we take a look into the spatial distribution of the galaxies.

Fig. 2.7 shows the contours of the unweighted and flux weighted density maps, top and bottom figures respectively, of the RCS galaxies. The contour levels begin at 100 gal Mpc^{-2} and increase in intervals of 50 gal Mpc^{-2}. Dots correspond to galaxies from our spectroscopic samples. These figures, regardless of whether they are weighted or unweighted, show the core of the two main structures with corresponding BCGs, and a high density of galaxies in-between them.

Table 2.4: Results for the substructure-identification tests applied to different subsamples.

Subsample	N	AD-test		DS-test	
		α	P-value	$\Delta/N_{\rm tot}$	P-value
Cubes 2+3	32	0.264	0.674	0.967	0.421
Cubes 1+4	48	0.383	0.383	1.329	0.097
All Cubes	79	0.234	0.789	1.205	0.138
MUSE+GMOS	87	0.272	0.662	1.203	0.152

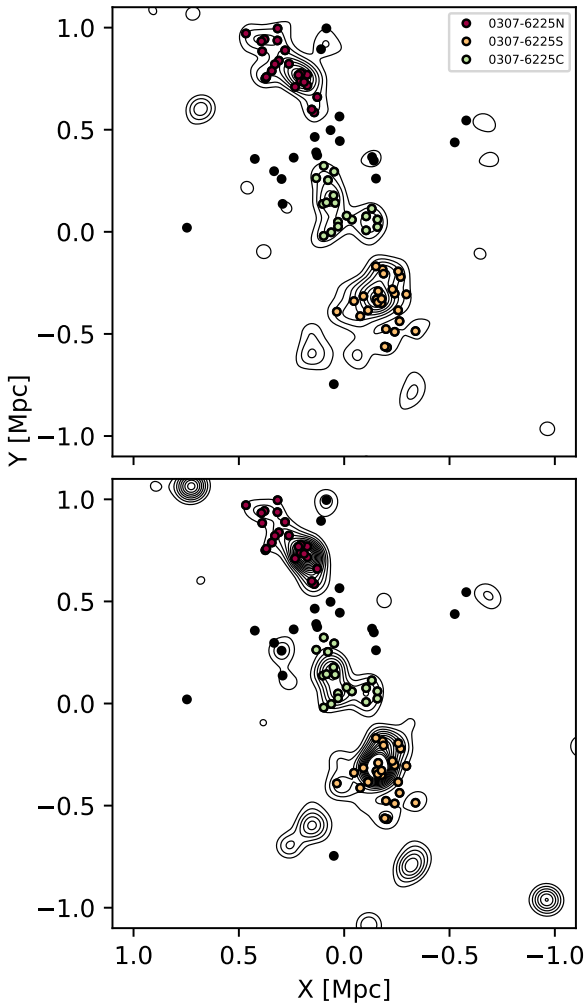

Figure 2.7: Unweighted (top) and flux weighted (bottom) RCS galaxies (photometric and spectroscopic) numerical density map is shown in black contours, where levels begin at 100 galaxies per Mpc^2 and the flux was estimated from the *i* band. Galaxies not close to the density levels or classified as not being part of any structure by the `DBSCAN` algorithm are shown as black dots, while dots in different substructures according to the algorithm are shown with different colors according to the substructure; 0307-6225N (red), 0307-6225S (orange) and a in-between overdensity (green).

For the definition of the substructures we take into account only spectroscopic members within (or near) the limits of our density contours. To distinguish the galaxies with a higher probability of being part of each structure we use the Density-Based Spatial Clustering of Applications with Noise (DBSCAN, Ester et al., 1996) algorithm. The advantage of using this algorithm is that the galaxies are not necessarily assigned to a given group, leaving some of them out. We use a PYTHON-based application of this algorithm, following the work of Olave-Rojas et al. (2018, substructure defined as at least three neighbouring galaxies within a separation of ~140 kpc).

Fig. 2.7 shows the results of the different found structures. Black dots represent galaxies that either were too far from our density contours or were discarded by the DBSCAN algorithm. We name the two most prominent structures, defined by DBSCAN, as 0307-6225N (red dots) and 0307-6225S (orange dots), comprised by 23 members and 25 members, respectively. The BCGs for 0307-6225S and 0307-6225N are marked in Table 2.8 by the upper scripts S_1 and N, respectively. Both structures show a Gaussian velocity distribution when applying the AD test, and the distance between them is: ~1.10 Mpc between their BCGs and ~1.15 Mpc between the peaks of the density distribution.

We also find a third substructure in-between the two colliding ones (green dots), which we name 0307-6225C, with 19 galaxies and no BCG-like galaxy. Fig. 2.8 shows the velocity distribution of the galaxies of each substructure, color coded following Fig. 2.7. Table 2.5 shows the sky coordinates of the substructures (estimated as the peak of the overdensity), along with their estimated redshifts, velocity dispersions and number of members.

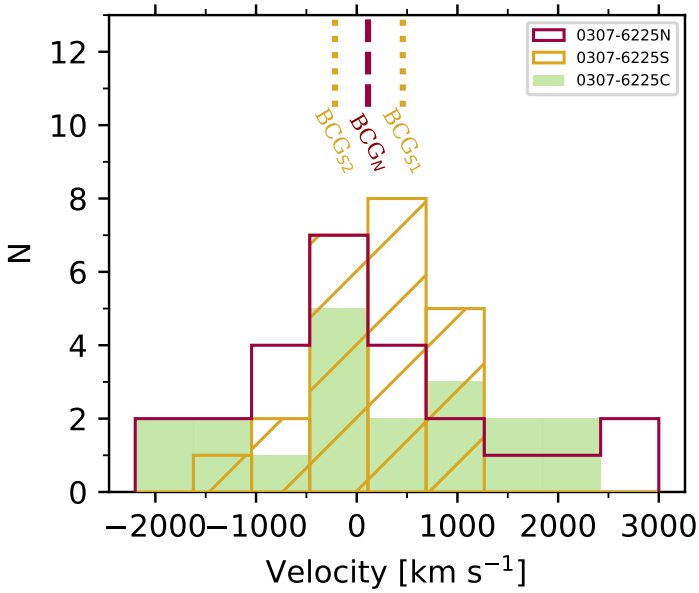

Figure 2.8: Peculiar velocity distribution of the galaxies belonging to the three substructures; 0307-6225N, 0307-6225S and 0307-6225C. The velocity of the BCG of 0307-6225N is shown with a dashed line in the top axis, while the velocity of the two BCGs of 0307-6225S shown with dotted lines.

2.4.2 Cluster dynamical mass

We estimate the masses using Munari et al. (2013) scaling relations between the mass and the velocity dispersion of the cluster (see Eq. 2.6). The Gaussian velocity distribution together with the large separation between the center of both structures (∼1.1 Mpc between the BCGs) and the fact that the velocity difference between them is $\Delta v_{N-S} = 342$ km s^{-1} (at the cluster's frame of reference) strongly suggest a plane of the sky merger (see, e.g. Dawson et al., 2015; Mahler et al., 2020) and could therefore, imply that the overestimation of the masses using scaling relations is minimal (Dawson et al., 2015). We further explore this in §2.5.1.1. In order to minimize the possible overstimation of using scaling relations, we only use RCS spectroscopic galaxies to estimate σ_v, since in clusters with a high accretion rate, blue galaxies tend to raise the value of the velocity dispersion (Zhang et al., 2012). Note that, however, the number of members shown in Table 2.5 also considers blue galaxies.

In Table 2.5 we show the estimated masses of the substructures. The two prominent substructures, 0307-6225S and 0307-6225N, have similar masses with the most probable ratio of $M_S/M_N \approx 1.3$ with large uncertainties. Galaxies selected for the dynamical mass estimation are likely to belong to the core regions of the two clusters. Galaxies in these regions are expected to be virialized and should more closely follow the gravitational potential of the clusters during a collision, giving a better estimation of the masses when using the velocity dispersion.

Table 2.5: Substructure properties.

Structure 0307-6225	R.A. (J2000)	Dec. (J2000)	z	σ_v km s^{-1}	$M_{200,dyn}$ $\times 10^{14}$ M$_\odot$	N
S	46.8195	-62.4463	0.5792 ± 0.0002	756 ± 164	3.16 ± 1.88	25
N	46.8526	-62.4009	0.5810 ± 0.0002	688 ± 145	2.44 ± 1.41	23
C	46.8396	-62.4258	0.5803 ± 0.0004	1415 ± 336	17.67 ± 11.53	19

2.4.3 Cluster merger orbit

With the masses estimated, the merging history can be recovered by using a two-body model (Beers et al., 1990; Cortese et al., 2004; Gonzalez et al., 2018) or by using hydrodynamical simulations constrained with the observed properties of the merging system (e.g. Mastropietro & Burkert, 2008; Machado et al., 2015; Doubrawa et al., 2020; Moura et al., 2021), with the disadvantage being that the latter method is computationally expensive. To understand the merging event, we use the Monte Carlo Merger Analysis Code (`MCMAC`, Dawson, 2013), which is a good compromise between computational time and accuracy of the results, with a dynamical parameter estimation accuracy of about 10% for two dissociative mergers; Bullet Cluster and Musket Ball Clusters. `MCMAC` analyzes the dynamics of the merger and outputs its kinematic parameters. The model assumes a two-body collision of two spherically symmetric halos with a NFW profile (Navarro et al., 1996, 1997), where the total energy is conserved and the impact

parameters is assumed to be zero. The different parameters are estimated from the Monte Carlo analysis by randomly drawing from the probability density functions of the inputs.

The inputs required for each substructure are the redshift and the mass, with their respective errors, along with the distance between the structures with the errors on their positions. We use the values shown in Table 2.5 as our inputs, where the errors for the redshifts are estimated as the standard error, while the errors for the distance are given as the distances between the BCGs and the peak of the density distribution of each structure (0.144 arcmin and 0.017 arcmin for 0307-6225N and 0307-6225S, respectively). The results are obtained by sampling the possible results through 10^5 iterations, and are showed and described in Table 2.6, with the errors corresponding to the 1σ level.

Table 2.6: Output from the `MCMAC` code, with the priors from Table 2.5. Errors correspond to the 1σ level.

Param.	Median	Unit	Description
α	39^{+13}_{-11}	deg	Merger axis angle
$d3D_{\rm obs}$	$1.29^{+0.32}_{-0.15}$	Mpc	3D distance of the halos at $T_{\rm obs}$.
$d3D_{\rm max}$	$1.72^{+0.44}_{-0.22}$	Mpc	3D distance of the halos at apoapsis.
$v3D_{\rm col}$	2300^{+122}_{-96}	km/s	3D velocity at collision time.
$v3D_{\rm obs}$	547^{+185}_{-103}	km/s	3D velocity at $T_{\rm obs}$.
$v_{\rm rad}$	339^{+28}_{-28}	km/s	Radial velocity of the halos at $T_{\rm obs}$.
TSP0	$0.96^{+0.31}_{-0.18}$	Gyr	TSP for outgoing system.
TSP1	$2.60^{+1.07}_{-0.53}$	Gyr	TSP for incoming system.

`MCMAC` gives as outputs the merger axis angle α, the estimated distances and velocities at different times and two possible current stages of the merger; outgoing after first pericentric passage and incoming after reaching apoapsis. The time since pericentric passage (TSP) for both possible scenarios are described as TSP0 for the outgoing scenario and TSP1 for the incoming one. This last two estimates are the ones that we will further discuss when recovering the merger orbit of the system.

To further constrain the stage of the merger we compare the observational features with simulations. We use the Galaxy Cluster Merger Catalog (ZuHone et al., 2018)[5], in particular, the "A Parameter Space Exploration of Galaxy Cluster Mergers" simulation (ZuHone, 2011), which consists of an adaptive mesh refinement grid-based hydrodynamical simulation of a binary collision between two galaxy clusters, with a box size of 14.26 Mpc. The binary merger initial configuration separates the two clusters by a distance on the order of the sum of their virial radii, with their gas profiles in hydrostatic equilibrium. With this simulation one can explore the properties of a collision of clusters with a mass ratio of 1:1, 1:3 and 1:10, where the mass of the primary cluster is $M_{200} = 6 \times 10^{14}$ $M_\odot$, similar to the SZ derived mass of $M_{200} = 7.63 \times h_{70}^{-1} 10^{14}$ $M_\odot$ for SPT-CL J0307-6225 (Bleem et al., 2015), and with different impact parameters (b = 0, 500, 1000 kpc).

[5] `http://gcmc.hub.yt/simulations.html`

We use both a merger mass ratio of 1:3 and 1:1. Since we cannot constrain the impact parameter, we use all of them and study their differences, where, for example, the bigger the impact parameter, the longer it takes for the merging clusters to reach the apoapsis. We also note that for our analysis we use a projection on the z-axis, since evidence suggests a collision taking place on the plane of the sky.

2.4.3.1 Determining TSP0 and TSP1 from the simulations

We use the dark matter distribution of both objects to determine the collision time, focusing on the distance between their density cusps at different snapshots. Also, to determine the snapshots for an outgoing and an incoming scenario, which would be the closest to what we see in our system, we look for the snapshot where the separation between the peaks is similar to the projected distance between our BCGs (∼1.10 Mpc).

Table 2.7: Estimated collision times and times since collision ($TSP0_{sim}$ and $TSP1_{sim}$) for the simulations with different impact parameters b and mass ratios.

b	Mass ratio	Collision time	$TSP0_{sim}$	$TSP1_{sim}$
kpc		Gyr	Gyr	Gyr
0	1:3	1.22 ± 0.02	0.78 ± 0.20	-
500	1:3	1.24 ± 0.02	0.66 ± 0.20	0.96 ± 0.20
1000	1:3	1.34 ± 0.02	0.56 ± 0.20	1.46 ± 0.20
0	1:1	1.32 ± 0.02	0.68 ± 0.20	-
500	1:1	1.34 ± 0.02	0.46 ± 0.20	-
1000	1:1	1.40 ± 0.02	0.80 ± 0.20	1.00 ± 0.20

Notes. No TSP1 value is provided when we cannot separate between the outgoing and incoming scenarios by requiring a distance of ∼1.1 Mpc.

In Table 2.7 we show the results for the different impact parameters, where the second column indicates the mass ratio. The third column shows the simulation time where the distance between the two halos is minimal (pericentric passage time). The errors are the temporal resolution of the simulation at the chosen snapshot. Following the previous nomenclature, the fourth column, $TSP0_{sim}$, corresponds to the amount of time from the first pericentric passage (minimum approach), while the fifth column, $TSP1_{sim}$, corresponds to the amount of time from the pericentric passage, to the first turn around, and heading towards the second passage. Times are either the snapshot time or an average between two snapshots if the estimated separations are nearly equally close to the ∼1.10 Mpc distance.

For $b = 0$ kpc, the maximum achieved distance between the two dark matter halos in the 1:3 mass ratio simulation was 1.05 Mpc, while for the 1:1 mass ratio it was 0.99 Mpc, meaning that we cannot separate between both scenarios when comparing the projected distance of 0307-6225N and 0307-6225S.

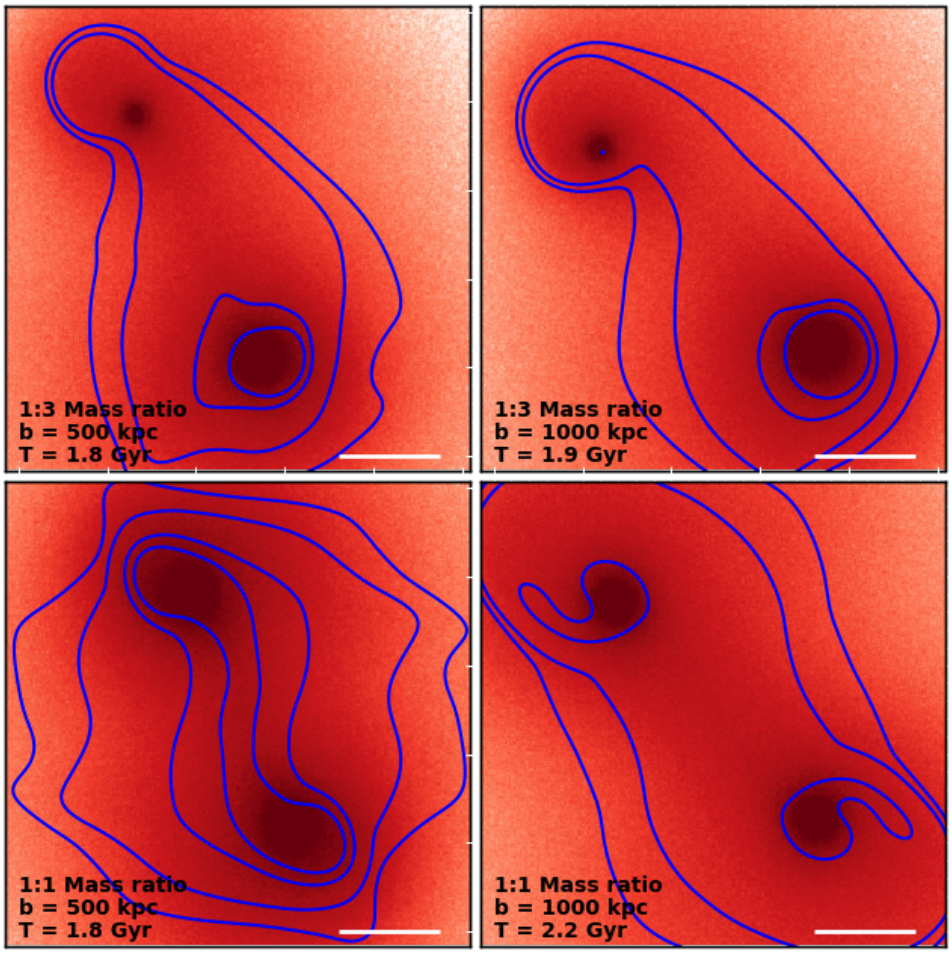

Figure 2.9: Density and X–ray contours of the different simulations. The simulation times are shown on the bottom left corner, and correspond to (or are close to in case of averaging over two snapshots) the collision time plus the TSP0 time since collision (see Table 2.7). The projected total density of the simulations is shown in red in the background, with the contrast starting at 1×10^7 $M_\odot$ kpc^{-2}. Blue contours where derived from the projected X–ray emission, with the levels being $0.5, 1, 5, 10, 15 \times 10^{-8}$ photons/s/cm^2/arcsec^2. Simulations are divided according to their mass ratio (1:3 on top and 1:1 on the bottom) and according to the impact parameter (500 kpc on the left panels and 1000 kpc on the right panels). The used box size is the same to the one used in Fig. 2.1. The white bar also corresponds to the same length of 1 arcmin shown in Fig. 2.1.

2.4.3.2 X–ray morphology

The hydrodynamical simulations render a gas distribution that can be directly compared to the observations. Fig. 2.9 shows the snapshots of the outgoing scenario, while Fig. 2.10 shows the snapshots of the incoming scenario, where the X–ray projected emission is overplotted as blue contours on top of the projected total density, for the simulation snapshots close to the derived TSP (Table 2.7), with the simulation time shown on the bottom left of each panel. Note however that for the 1:1 mass ratio and b = 500 kpc, the system has the ∼1.1 Mpc distance at turnaround,

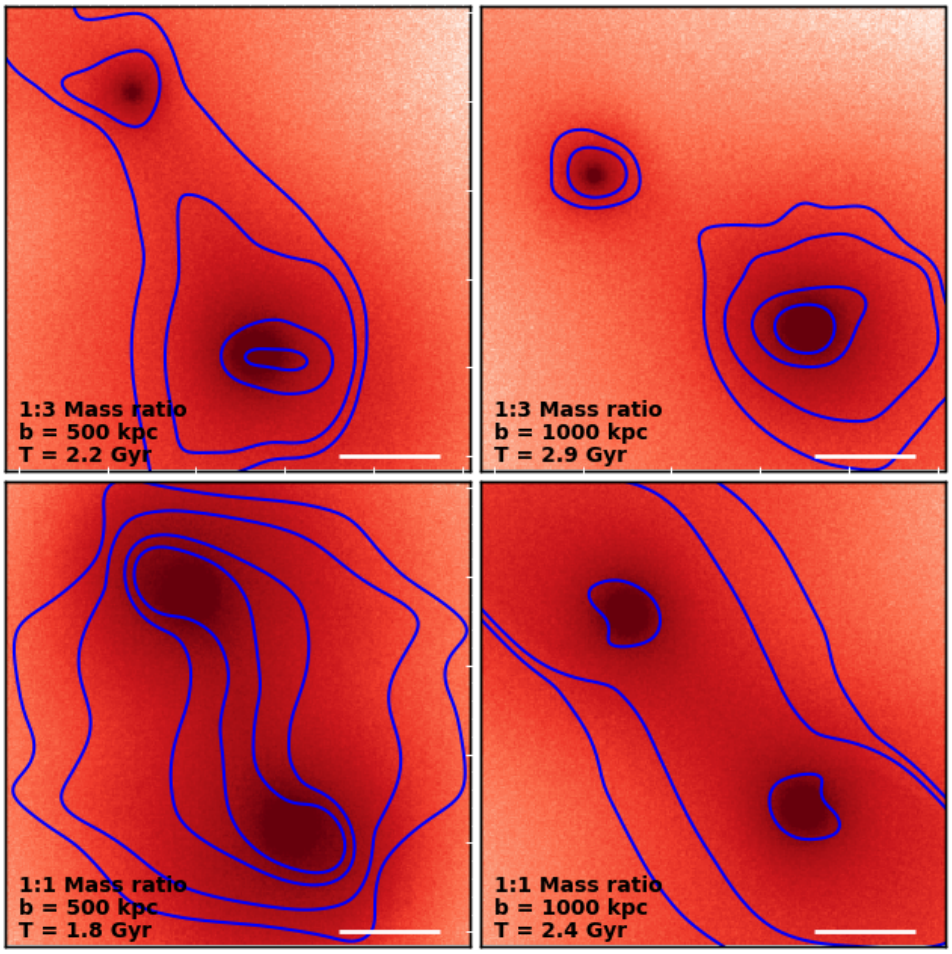

Figure 2.10: Same as Fig. 2.9, but derived from the simulations at the TSP1 times.

which means that we cannot differentiate between and outgoing and incoming scenario. We decide to keep the same snapshot in both Figures 2.9 and 2.10 just for comparison. The scenarios for 1:3 mass ratio closest resemble the gas distribution from our *Chandra* observations (orange contours on Fig. 2.1). We comeback to this in § 2.5.1.3.

2.4.4 The impact of the merging event in the galaxy populations

In Fig. 2.11 we show the CMD for each subsample; all galaxies, galaxies belonging to 0307-6225N and 0307-6225S, and galaxies not belonging to either of them. Galaxies are color coded according to their spectral classification. Most of the star-forming galaxies are located within the two main structures (9 out of 10 SF+SSB galaxies), with some of them being classified as RCS galaxies (4; 2 SF and 2 SSB). Galaxies with S/N < 3 and/or $i_{\rm auto} > m^* + 2$ are plotted as black crosses.

Given that most of the SF galaxies seem to be located in the substructures, especially the red SF galaxies, it is plausible that they were part of the merging event, instead of being accreted after

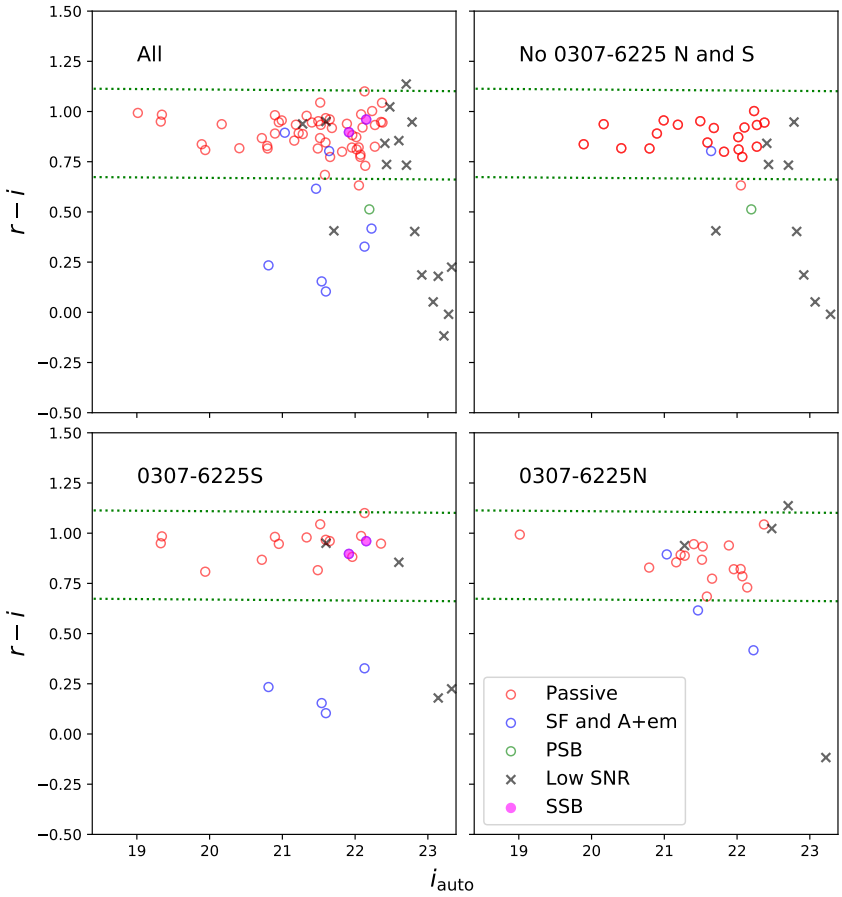

Figure 2.11: CMD of the cluster for the different samples. Galaxies are color-coded depending on their spectral classification described in §2.4.4. *top left*: entire spectroscopic data sample. *top right*: sample comprising galaxies not belonging to 0307-6225N and 0307-6225S, i.e., galaxies from 0307-6225C plus galaxies not belonging to any substructure according to DBSCAN. *bottom*: 0307-6225S and 0307-6225N samples shown in left and right panels, respectively. The green dotted lines are the limits for the RCS zone. Black crosses are galaxies with S/N < 3 or $i_{\rm auto} \geq m^* + 2$. Filled colors are galaxies classified as SSB.

it. In Fig. 2.12 we show a phase-space diagram, with the X-axis being the separation from the SZ-center. Galaxies are color coded following the substructure to which they belong. In Fig. 2.20 we show small crops of 7×7 arcsec^2 (47×47 kpc^2 at the cluster's redshift) of the EL galaxies plus the two blue NEL galaxies, separating by different substructures and with the spectra of each galaxy shown to the right.

2.4.5 The particular case of 0307-6225S

Fig. 2.11 shows that 0307-6225S has (1) the bluest members from our sample and (2) two very bright galaxies with nearly the same magnitudes (galaxies with ID 35 and 46 from the MUSE-1 field in Table 2.8, marked with an upper script S_1 and S_2, respectively). In Fig. 2.13 we provide a zoom from Fig. 2.1, to show in more detail the southern structure. Red circles mark spectroscopic

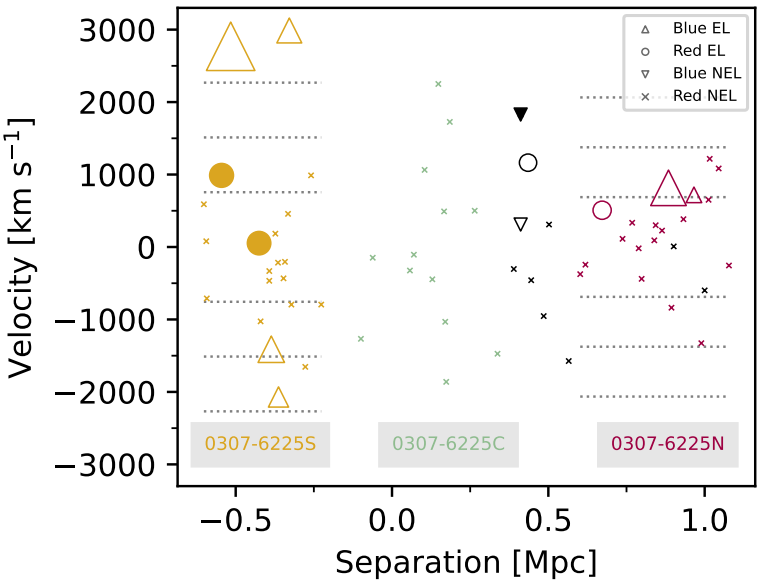

Figure 2.12: Phase-space diagram of spectroscopic members with S/N $\geq$ 3 and $i_{\rm auto} < m^* + 2$. The separation is measured with respect to the SZ-center, negative for objects to the south of it. Galaxies are colored as dark red, dark orange, dark green and black if they were classified as belonging to the 0307-62255N, 0307-6225S, 0307-6225C or to neither of them, respectively. Crosses are galaxies classified as non-emission line galaxies. Emission line galaxies which belong to (or have redder colors than) the RCS are plotted as circles, triangles are galaxies with colors lower than the RCS, whereas inverted triangles are blue post-starburst (filled) or passive (unfilled) galaxies. The sizes of EL galaxies are correlated with their EW(OII) strength. Filled circles correspond to SSB galaxies. Black dotted lines mark $\pm1\sigma_v$, $\pm2\sigma_v$ and $\pm3\sigma_v$ for the two main substructures.

members for this region with S/N > 3 and $i_{\rm auto} < m^* + 2$. The two brightest galaxies are the two elliptical galaxies in the middle marked with red stars, with $\Delta m_i = 0.0152 \pm 0.0063$ and $\Delta v = 600 km/s$. The on-sky separation between the center of them (~41 kpc), suggests that these galaxies could be interacting with each other.

2.5 Discussion

2.5.1 Merging history of 0307-6225S and 0307-6225N

2.5.1.1 Mass estimation of a merging cluster

Being able to recover the merging history of two observed galaxy clusters is not trivial. Most methods require a mass estimation of the colliding components, which is not always an easy task (see merging effect on cluster mass in Takizawa et al., 2010; Nelson et al., 2012, 2014).

The velocity dispersion (along the line-of-sight) of the galaxies of a cluster can be used to infer its mass, using for example the virial theorem (e.g. Rines et al., 2013; White et al., 2015) or scaling relations (e.g. Evrard et al., 2008; Saro et al., 2013; Munari et al., 2013; Dawson et al.,

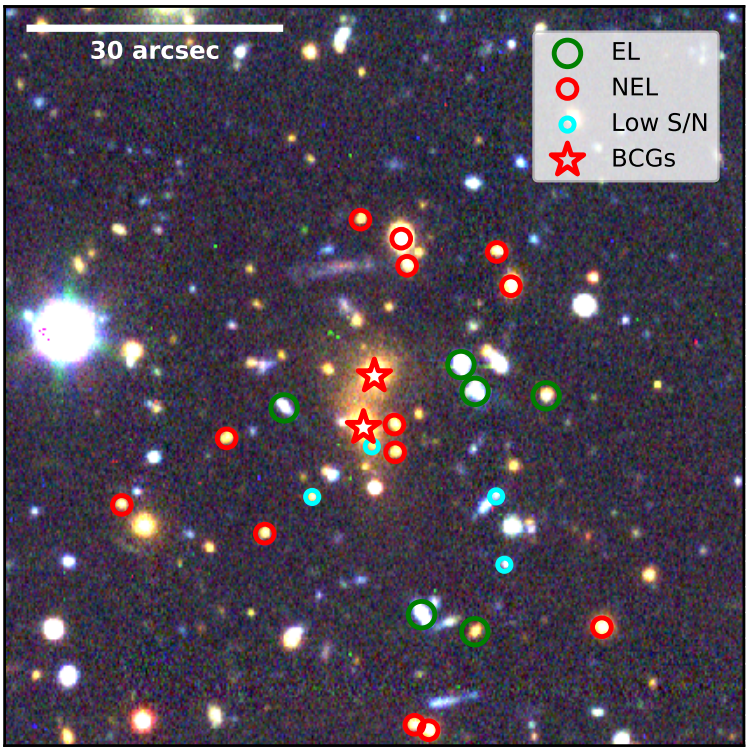

Figure 2.13: Zoom from Fig. 2.1 into 0307-6225S, with the white bar on the top left showing the scale of the image. Spectroscopic members with S/N < 3 or $i_{\rm auto} \geq m^* + 2$ are shown as cyan circles, while red and green circles/stars represent passive and emission-line cluster galaxies, respectively, where emission-line refers SF or SSB galaxies. The 2 brightest galaxies are marked with stars.

2015; Monteiro-Oliveira et al., 2021). For the mass estimations of our structures we use the later one, although it is important to note that these measurements are also affected by the merging event, as colliding structures could show alterations in the velocities of their members. White et al. (2015) argues that the masses of merging systems estimated by using scaling relations can be overestimated by a factor of two. Evidence suggests that the merger between 0307-6225S and 0307-6225N is taking place close to the plane of the sky, with a low velocity difference between the two, similar to what Mahler et al. (2020) find for the dissociative merging galaxy cluster SPT-CLJ0356-5337. The velocity difference between the BCGs and the redshift of each substructure is $\leq$20 km s^{-1} for both substructures, which might indicate that the two merging substructures were not too dynamically perturbed by the merger.

It is worth noting that recently Ferragamo et al. (2020) suggested correction factors on both σ_v and the estimated mass to account for cases with a low number of galaxies. They also apply other correction factors to turn σ_v into an unbiased estimator by taking into account, for example, interlopers and the radius in which the sources are enclosed. However, applying these changes does not change our results drastically, with the new derived masses being within the errors of the previously derived ones.

To check how masses derived from the velocity dispersion of merging galaxy clusters could

be overestimated, we estimate the masses, following the equations from Munari et al. (2013), of the simulated clusters from the 1:3 merging simulation (from §2.4.3) at all times (and b) using their velocity dispersion. It is worth noting that we cannot separate RCS members to estimate the velocity dispersions, since the simulation does not give information regarding the galaxy population. Fig. 2.14 shows the σ_v derived masses at different times for the 1:3 mass ratio simulation for different values of b. The black dotted lines represent the collision time and the dashed lines with the gray shaded areas represent the TSPs and their errors from Table 2.7, respectively. Before the collision and some Gyr after it, the masses are overestimated, especially for the case of the smaller mass cluster. However, near the TSP0 times, the derived masses are in agreement, within the errors, with respect to the real masses. This is true also for the TSP1 with $b = 500$ kpc, but for the same time with $b = 1000$ kpc, the main cluster's mass is actually underestimated. Although we cannot further constrain the masses from the simulation using only RCS members, this information does suggest that our derived masses are not very affected by the merging itself given the possible times since collision.

Bleem et al. (2015) estimated a total Sunyaev-Zeldovich based mass of $M_{500,SZ} = 5.06 \pm 0.90 \times 10^{14}\ h_{70}^{-1}\ M_\odot$, corresponding to $M_{200,SZ} = 7.63 \pm 1.37 \times 10^{14}\ h_{70}^{-1}\ M_\odot$ (Zenteno et al., 2020), which is in agreement to our estimation of the total dynamical mass from scaling relations $M_{200,dyn} = M_S + M_N = 5.55 \pm 2.33 \times 10^{14}\ M_\odot$, at the 1σ level.

2.5.1.2 Recovery of the merger orbit

`MCMAC` gives as a result two different time since collision, TSP0=$0.96^{+0.31}_{-0.18}$ Gyr and TSP1=$2.60^{+1.07}_{-0.53}$ Gyr, for an outgoing and an incoming merger, respectively, after the first pericentric passage. A more detailed analysis of the X–ray could further constrain both the `MCMAC` output, e.g. by constraining the merging angle (Monteiro-Oliveira et al., 2017, 2018) and the TSP (Dawson, 2013; Ng et al., 2015; Monteiro-Oliveira et al., 2017) from shocks (if any), and also the merging scenario from hydrodynamical simulations, e.g. by comparing the temperature maps or by running a simulation which recovers the features (both of the galaxies and of the ICM) of this particular merger. This is particularly interesting given that the simulations that we use to compare have a merger axis angle of $\alpha = 0.0$ deg. Dawson (2013) runs `MCMAC` on the Bullet Cluster data and finds $\alpha = 50^{+23}_{-23}$ deg, however, by adding a prior using the X–ray shock information, he is able to constrain the angle to $\alpha = 24^{+14}_{-8}$ deg, which is closer to the plane of the sky and also decreases significantly the error bars on the estimated collision times.

For instance, if we assume that the merger is nearly on the plane of the sky and constrain the merging angle, α, from `MCMAC` to be between 0° and 45°, then the resulting values are $\alpha = 25^{+6}_{-6}$ deg, TSP0=$0.73^{+0.09}_{-0.09}$ and TSP1=$2.10^{+0.51}_{-0.30}$, which are still within the previous estimated values (within the errors) and have smaller error bars. However, the estimated TSP1 is still higher than any of the ones estimated from the simulations (see Table 2.7).

A similar system is the one studied by Dawson et al. (2012); DLSCL J0916.2+2951, a major merging at $z = 0.53$, with a projected distance of $1.0^{+0.11}_{-0.14}$ Mpc. Their dynamical analysis gives masses similar to that of our structures (when using $\sigma_v - M$ scaling relations), with the mass ratio between their northern and southern structures of $M_S/M_N = 1.11 \pm 0.81$. Using an analytical model, they were able to recover a merging angle $\alpha = 34^{+20}_{-14}$ degrees and a physical separation

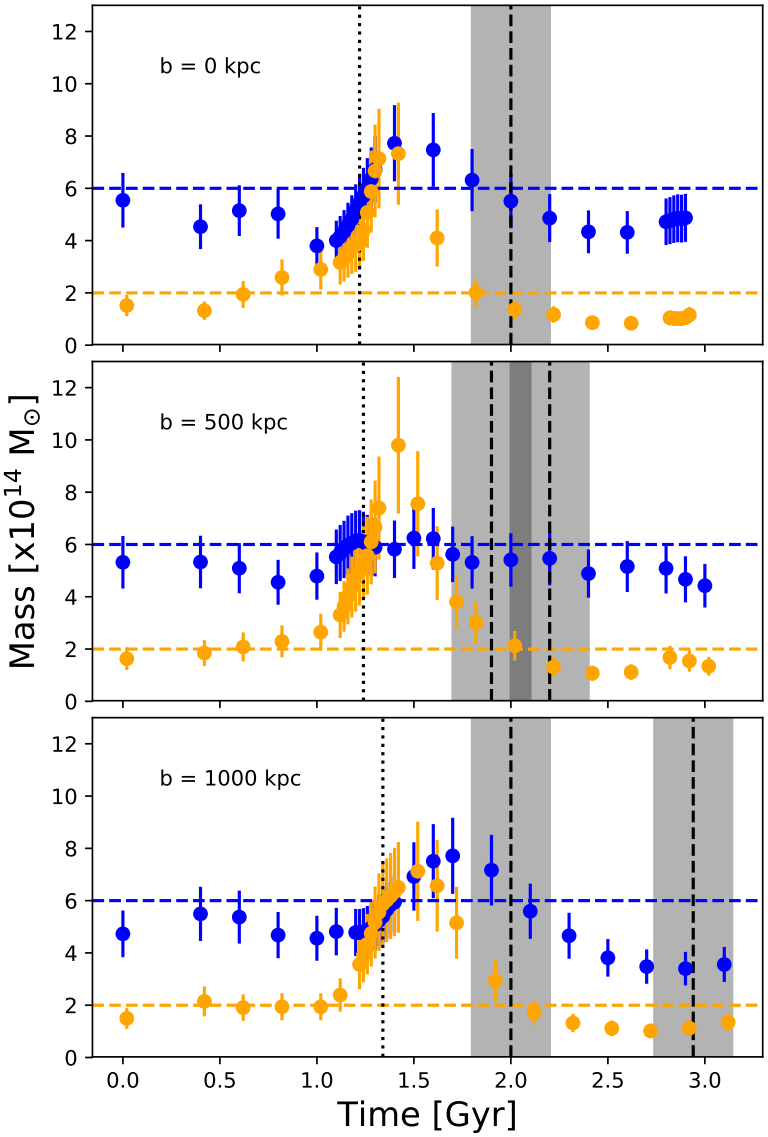

Figure 2.14: Velocity dispersion derived masses for the 1:3 mass ratio simulations used in this work, with different b. The x-axis is the time since the simulation started running, with the blue and orange dots corresponding to the main cluster and the secondary cluster, respectively. The blue and orange dashed lines represent the masses of 6×10^{14} and 2×10^{14} $M_\odot$, respectively. Black dotted lines mark the collision times estimated following §2.4.3. Vertical black dashed lines mark the estimated TSP0 and TSP1 shown in Table 2.7, with the gray area being the errors on this estimation.

of $d3D = 1.3^{+0.97}_{-0.18}$, both values in agreement with what we found. Furthermore, their time since collision is also similar to the one found for our outgoing system TSP= $0.7^{+0.2}_{-0.1}$, however they do not differentiate between an outgoing or incoming system.

Regarding 0307-6225C, the estimated velocity dispersion is very high ($\sigma_v = 1415$ km s^{-1}) and the density map shows that this region is not as dense as the other two, with no dominant

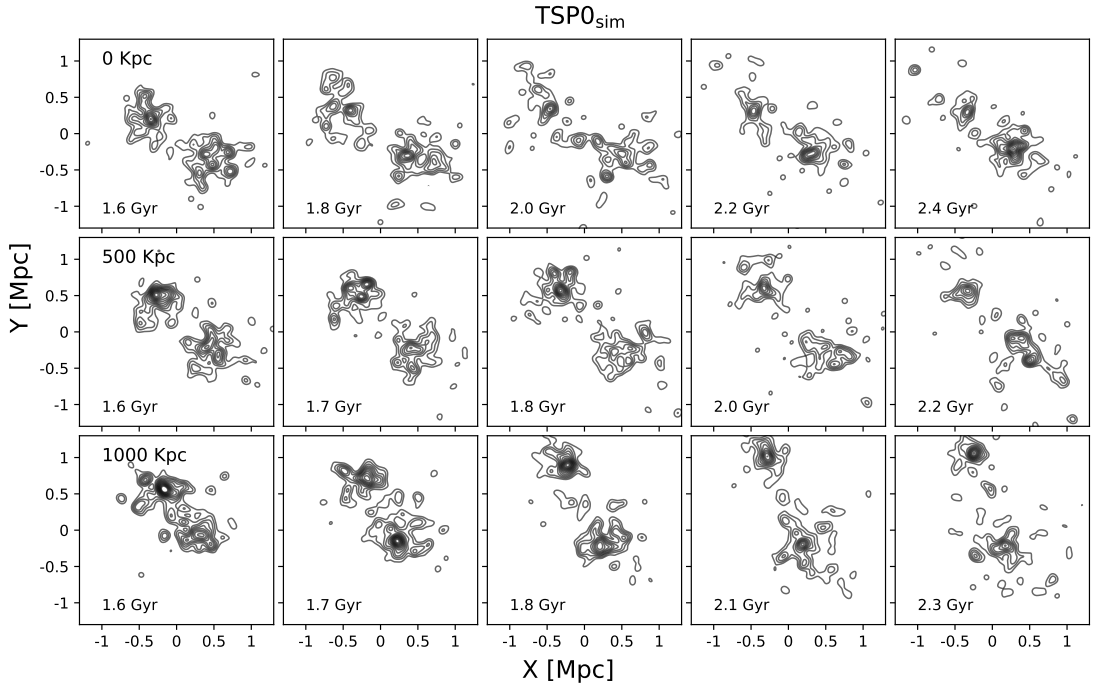

Figure 2.15: Density maps for the simulated 1:3 mass ratio cluster merger. Each row represents the time evolution around the TSP0 for the different impact parameters $b = 0, 500, 1000$ kpc shown at the top, middle and bottom rows, respectively. For each panel, the simulation time is written on the bottom left.

massive galaxy. To check whether it is common for a merging of two galaxy clusters, we take a look at how the density map varies in the 1:3 mass ratio simulations near the estimated TSP0. We show in Fig. 2.15, on each row, the density maps of the simulations with the corresponding time shown at the bottom left, and the impact parameter of the row at the top left of the first figure of each row. At different times, the density maps for the same impact parameter show to be rather irregular, with the in-between region changing from snapshot to snapshot. In particular, both $b = 0$ kpc and $b = 1000$ kpc show an overdense in-between area near the TSP0. However, this is not the case in other snapshots, so we cannot state with confidence that this is common for a merging cluster to show such a pronounced in-between overdense region.

2.5.1.3 Constraining the TSP with simulations

We compare the results derived by `MCMAC` with those estimated from a hydrodynamical simulation of two merging structures with a mass ratio of 1:3 (ZuHone, 2011; ZuHone et al., 2018). We chose this ratio since the X–ray morphologies of both the simulation and the system are a better match than the 1:1 mass ratio, where the X–ray intensity from the simulation is similar for the two structures (see Fig. 2.9 and 2.10), unlike our system, which have two distinctly different structures (see the orange contour in Fig. 2.1).

Using dark matter only simulations, Wittman (2019) looked for halos with similar configurations to those of observed merging clusters (such as the Bullet and Musket Ball clusters) and

compared the time since collisions to those derived by MCMAC and other hydrodynamical simulations, finding that with respect to the latter the derived merging angles and TSP are consistent. However, both the outgoing and incoming TSP and the angles are lower than those derived by MCMAC, attributing the differences to the MCMAC assumption of zero distance between the structures at the collision time.

Sarazin (2002) discuss that most merging systems should have a small impact parameter, of the order of a few kpc. Dawson et al. (2012) argues that, given the displayed gas morphology, the dissociative merging galaxy cluster DLSCL J0916.2+2951, has a small impact parameter. The argument is that simulations show that the morphology for mergers with small impact parameters, is elongated transverse to the merger direction (Schindler & Muller, 1993; Poole et al., 2006; Machado & Lima Neto, 2013). The X–ray morphology shown in this paper is similar to that from Dawson et al. (2012). It is also similar to that of Abell 3376 (Monteiro-Oliveira et al., 2017), a merging galaxy cluster which was simulated by Machado & Lima Neto (2013) with different impact parameters (b = 0, 150, 350 and 500 kpc), with their results suggesting that a model with $b < 150$ kpc is preferred. Given the similitude between SPT-CL J0307-6225 X–ray morphology and that of other systems such as Abell 3376 and DLSCL J0916.2+2951, then we suggest that the simulations with $b = 0$ kpc or $b = 500$ kpc are better representations of our system. This implies that the preferred scenario for this merging cluster is that of an outgoing system or a system very close to turnaround. This can also be seen when comparing the X–ray morphology of SPT-CL J0307-6225 with that of the 1:3 mass ratio simulations at the estimated $TSP0_{sim}$ and $TSP1_{sim}$, shown in Fig. 2.9 and Fig. 2.10, respectively, with the X–ray contours at $TSP0_{sim}$ being more similar than the ones at $TSP1_{sim}$ for b = 500, 1000 kpc.

2.5.1.4 Proposed merger scenario

We propose that the merger scenario that best describes the observations of 0307-6225 is that of a post-merger seen $0.96^{+0.31}_{-0.18}$ Gyr after collision. Combining the simulations with results from literature we constraint the impact parameter to be $b < 500$ kpc. Simulations also support a mass ratio closer to 1:3 than 1:1, given the X–ray morphology.

2.5.2 Galaxy population in a merging galaxy cluster

An interesting feature of our system is that 90% of the EL galaxies belong to any of the main substructures (Fig. 2.12). Stroe & Sobral (2021) found that, for merging galaxy clusters, 40% (80%) of EL galaxies are located within 1.5 Mpc (3 Mpc) of the cluster center. To study this behavior further and analyse if our EL galaxies favour a spatial position within the substructures, we compare their galaxy radial distribution to that of the central region. We combine the galaxy distributions of 0307-6225S and 0307-6225N by normalizing the clustercentric distances R by the virial radius, R_{200}, of each substructure (1.16 Mpc for 0307-6225S and 1.06 Mpc for 0307-6225N) and then estimate the fraction of EL and NEL galaxies within bins of R/R_{200}. In the case of the central region we use the SZ position as the center and average the R_{200} of the main substructures as the normalization radius (choosing only one of the radius does not affect the results). Fig. 2.16 shows the estimated fractions as a function of the clustercentric distance for $R < 0.5 \times R_{200}$, with

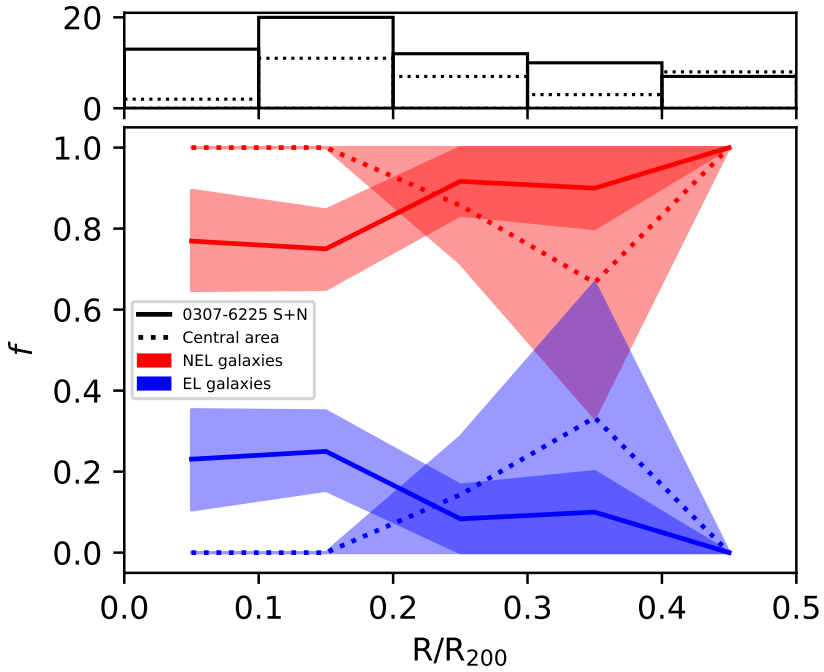

Figure 2.16: Fraction of EL (blue) and NEL (red) galaxies with respect to the distance to the centers of 0307-6225S and 0307-6225N (continuous lines) and the SZ center (dotted lines). To compute the errors we do 10000 Poisson realizations around the true number of NEL and EL galaxies within a radial bin, and re-estimate the fractions for each area. We then compute the 16th and 84th percentiles as the 1σ error regions, which are shown as shaded areas. The top panel shows the total number of galaxies per bin and per area (continuous lines for 0307-6225 S+N and dotted lines for the central area).

the total number of galaxies per bin shown in the upper panel. The fraction of EL galaxies towards the inner regions of the substructures (blue continuous line) is higher compared to that of the central area (blue dotted line), which is non-existent, at the 1σ level. Overall, EL galaxies are preferentially located at distances of $R < 0.2\times R_{200}$ from the substructures centers.

We will divide the discussion of the galaxy population by studying the differences between the two clumps, analysing the red EL galaxy population and also the population in the area in-between 0307-6225S and 0307-6225N. Following the work of Kelkar et al. (2020) we also study the EW(Hδ) vs D_n4000 plane in order to analyze the properties of the galaxy population. Kelkar et al. (2020) studied the galaxy population in the merging cluster Abell 3376 (A3376), a young post merger (~0.6 Gyr) cluster at $z \sim 0.046$ with clear merger shock features, analyzing the location of the galaxies, in particular of PSB galaxies. The D_n4000 index corresponds to the ratio between the flux redward and blueward the 4000 Å break, indicating the ages of the stellar population of the galaxies, which makes it an interesting measurement against the EW(Hδ) in absorption. Fig. 2.17 shows the EW(Hδ)-D_n4000 plane for our EL and blue NEL galaxies, where we estimate the D_n4000 index following Balogh et al. (1999). We will further discuss the positions within the plane of the different galaxy types in the following subsections.

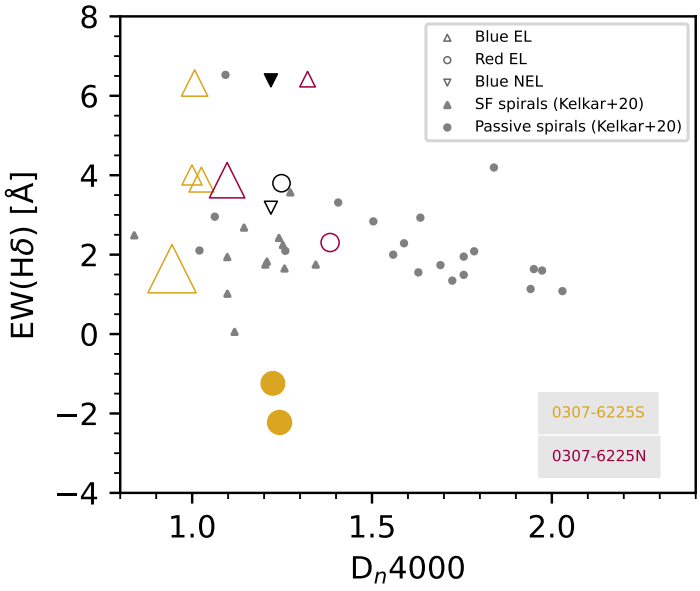

Figure 2.17: EW(Hδ) vs D$_n$4000 index for our EL and blue NEL galaxies, continuing the symbols of Fig. 2.12. Gray markers show SF (triangles) and passive (circles) spiral galaxies from Kelkar et al. (2020).

2.5.2.1 Comparison between the northern and southern sub-clusters

One interesting optical feature of 0307-6225S, is the two bright galaxies ($d_{\rm proj}$ = 41 kpc) at the center of its distribution (Fig. 2.13). A similar, but rather extreme case is that of the galaxy cluster Abell 3827 at z = 0.099, which shows evidence for a recent merger with four nearly equally bright galaxies within 10 kpc from the central region (Carrasco et al., 2010; Massey et al., 2015). Using GMOS data, Carrasco et al. (2010) found that the peculiar velocities of at least 3 of these galaxies are within ~300 km s^{-1} from the cluster redshift, with the remaining one having an offset of ~1000 km s^{-1}.

BCGs have low peculiar velocities in relaxed clusters, whereas for disturbed clusters it is expected that their peculiar velocity is 20-30% the velocity dispersion of the cluster (Yoshikawa et al., 2003; Ye et al., 2017). For 0307-6225S, one of the bright galaxies has a peculiar velocity of ~666 km s^{-1}, which is ~88% the velocity dispersion of this subcluster. This could be evidence of a past merging between 0307-6225S and another cluster previous to the merger with 0307-6225N. The AD test gives a Gaussian distribution, where the results do not change by applying a 3-σ iteration, which could indicate that the substructure is a post-merger.

We apply the Raouf et al. (2019) magnitude gap method to separate between relaxed and unrelaxed systems, to 0307-6225S and 0307-6225N independently. They use the magnitude difference between the first and second brigthest galaxy and select relaxed clusters as those with $\Delta M_{12} < 1.7$, whereas for unrelaxed clusters they use $\Delta M_{12} < 0.5$. We find that for 0307-6225S the magnitude difference is $\Delta M_{12} = 0.0152 < 0.5$, which supports the scenario that 0307-6225S suffered a previous merger prior to the one with 0307-6225N. Central galaxies take ≈1 Gyr to settle to the cluster centre during the post-merger phase (White, 1976; Bird, 1994), meaning

that this previous merger must have taken place over 1 Gyr before the observed merger between 0307-6225S and 0307-6225N. On the other hand, for 0307-6225N the value is $\Delta M_{12} \approx 1.8 > 1.7$, meaning 0307-6225N was a relaxed system prior to this merger.

Regarding the overall galaxy population, the fraction of EL galaxies in 0307-6225S (24%) is nearly two times that of 0307-6225N (∼13%), although consistent within 1σ. All EL galaxies from 0307-6225N have small peculiar velocities (within $1\sigma_v$), while for 0307-6225S 75% (50%) of the blue SF galaxies have peculiar velocities higher than $2\sigma_v$ ($3\sigma_v$), as seen in Fig. 2.12. These galaxies, which are bluer than the blue EL galaxies of 0307-6225N (Fig. 2.11), could be in the process of being accreted.

Fig. 2.17 shows that blue EL galaxies located in 0307-6225N tend to have older stellar populations than their blue counterparts from 0307-6225S. Apart from the PSB galaxy (black filled triangle), there are 2 other blue galaxies with similar measured EW(Hδ). Both of this galaxies might be dusty star forming galaxies (spectral type A+em, Balogh et al., 1999), with the one from 0307-6225S having the smallest peculiar velocity of the blue galaxies from this subcluster ($\approx -1400 kms^{-1}$). The recent infall of this galaxy might be the reason behind the truncated star formation, whereas for the blue galaxy from 0307-6225N, with an older stellar population and a peculiar velocity within $1\sigma_v$ (within the errors), the merger itself might be the reason.

Stroe et al. (2015) found that the increase of Hα emission of galaxies in the "Sausage" merging galaxy cluster, compared to galaxies in the "Toothbrush" merging galaxy cluster could be explained by their time since collision, with the "Toothbrush" cluster being more evolved (TSP∼2 Gyr, Brüggen et al., 2012) than the "Sausage" (TSP∼1Gyr, van Weeren et al., 2011). This timescales are similar to what we see from the merger of 0307-6225N and 0307-6225S (TSP=$0.96^{+0.31}_{-0.18}$) and the possible previous merger of 0307-6225S, which happened at least ≈1 Gyr prior to the collision with 0307-6225N. This previous merger could have exhausted the star formation of the galaxies of 0307-6225S, which might be the reason that there are no blue star forming galaxies towards the central region (within $1\sigma_v$) of 0307-6225S compared to 0307-6225N.

2.5.2.2 Red EL galaxies

Of particular interest are our EL galaxies located in the RCS. Out of the 4 red EL galaxies, 3 are located in the cores of the two main structures, with 2 of them classified as SSB. Most of the blue SF galaxies are best matched by a high-redshift star forming or late-type emission galaxy template, whereas most of the red SF galaxies are best matched with an early-type absorption galaxy template. Our red EL galaxies have older stellar populations than our blue EL galaxies (except for 1, Fig. 2.17), with the red EL galaxy from 0307-6225N having older stellar populations than those of 0307-6225S, which might be expected given that they are SSB.

Koyama et al. (2011) studied the region in and around the $z = 0.41$ rich cluster CL0939+4713 (A851) using Hα imaging to distinguish SF emission line galaxies. A851 is a dynamically young cluster with numerous groups at the outskirts. They found that the red Hα emitters are preferentially located in low-density environments, such as the groups and the outskirts, whereas in the core of the cluster they did not find red Hα emitters. Similar results were found by Einasto et al. (2018) for the galaxy cluster Abell 2142, with star forming galaxies (which includes red star-

forming galaxies) located at 1.5-2.0 h^{-1} Mpc from the cluster centre. Ma et al. (2010) studied the galaxy population of the merging galaxy cluster MACS J0025.4-1225 at $z = 0.586$. In the areas around the cluster cores (with a radius of 150 kpc) they find emission line galaxies corresponding to two spiral galaxies (one for each subcluster), plus some spiral galaxies without spectroscopic information, accounting for 14% of the total galaxies within the radius. Their Fig. 15 shows that they also have red EL galaxies, however they don't specify whether the 2 spiral galaxies within the cluster core are part of this population. Results from Ma et al. (2010), Koyama et al. (2011) and Einasto et al. (2018) indicate that red EL galaxies are not likely to be found within the cores of dense regions.

Sobral et al. (2016) studied the population of Hα emitters in the super-cluster Abell 851, finding that galaxies with higher dust extinctions to be preferentially located towards the densest environments. The results deviate from the expected extinctions given the masses of the galaxies. There is evidence for a population of RCS sequence galaxies with residual star formation in galaxy clusters as seen using ultra violet images. Crossett et al. (2014) found these galaxies to be red spirals located in low-density environments and towards the outskirts of massive clusters, concluding that they are either spirals with truncated star formation given their infall or high-mass spirals. Sheen et al. (2016) found that for four rich Abell clusters at $z \leq 0.1$, the fraction of red sequence galaxies with recent star formation that show signs of recent mergers is $\sim 30\%$, implying internal processes playing a significant role for the supply of cold gas to this galaxy population.

75% of our red EL galaxies do not have close neighbours which can supplement their gas reserves (Fig. 2.20). It is possible then that these objects accreted gas from the ICM, with the merger triggering then the SF. Given the peculiar velocity of the two SSB galaxy from our sample (which is classified as red), at least one of them was most likely part of the merging event. If, for example, merger shocks travelling through the ICM can trigger a starburst episode on galaxies with gas reservoirs for a few 100 Myr (Caldwell & Rose, 1997; Owers et al., 2012; Stroe et al., 2014, 2015), then these galaxies would make the outgoing scenario a better candidate than the incoming one. Another mechanism that can trigger a starburst of the gas is the rapid change of the tidal gravitational field due to the merger, which can drive gas to the inner part of galaxies (Bekki, 1999; Ferrari et al., 2003; Yoon et al., 2019).

Unfortunately, we do not see evidence of shocks in our X–ray data, likely due to it being shallow given the redshift. Shocks lasting 1-2 Gyr, are expected to generate in mergers of clumps with M$\geq 10^{13}$ M$_\odot$ with colliding velocities of 10^3 km s^{-1}, generating kinetic energies of over 10^{62} erg (Markevitch & Vikhlinin, 2007). Ha et al. (2018) found evidence for shocks using hydrodynamical simulations of merging galaxy clusters with mass ratio ~2, average virial masses similar to that of 0307-6225S and low impact parameters $b \leq 140$ kpc. They found that shocks are likely to be observed ~1 Gyr after the shock generation, at distances of 1-2 Mpc from the merger center, with mean mach numbers $M_S = 2 - 3$. Thus, we expect shocks to be have taken part in our system given the similar mass properties and the collision velocity we estimate with MCMAC (2300^{122}_{-96} km s^{-1}, Table 2.6).

2.5.2.3 Area in-between the main substructures

The central area, meaning 0307-6225C and other galaxies not associated to any substructure, is comprised of ~86% red passive galaxies, with the only EL galaxy belonging to the RCS. Moreover, the 2 blue galaxies are classified as a passive and a PSB. Ma et al. (2010) found a fraction of post-starburst galaxies in the major cluster merger MACS J0025.4-1225, on the region in-between the collision between the two merging components, where, given the timescales, the starburst episode of them occurred during first passage. Similarly to our blue galaxies in this region, they found that their colors are located between those of blue EL galaxies and red passive galaxies (Fig. 2.11).

Kelkar et al. (2020) divided the PSB population in three subsamples: bright, faint and blue. Although they don't find a trend for the first two, they find that blue PSB tend to be concentrated between the two BCGs, along the merger axis, although showing a wide variety of line-of-sight velocities. Fig. 2.6 shows a similar trend for the PSB (red filled triangle) and passive (red unfilled triangle) blue galaxies in the central region. However, the velocity of the PSB (Fig. 2.12) indicates that this might be the result of the infall in the cluster rather than an outcome of the merger. This does not seem to be the case for the blue passive galaxy, with a velocity of ~310 km s^{-1}.

Pranger et al. (2013) found a high fraction of NEL spiral galaxies towards the cluster core (< 1.2 Mpc) of the merging galaxy cluster Abell 3921 ($z = 0.093$). Their results are in agreement with the idea of passive spirals being preferentially located in high density environments in relaxed clusters (Bösch et al., 2013, e.g.), being an intermediate stage before developing to S0 galaxies (e.g. Vogt et al., 2004; Moran et al., 2007). Passive spirals are believed to be the results of ram pressure stripping during their infall onto galaxy clusters (e.g. Vogt et al., 2004), which correlates with the small velocity and EW(Hδ) of our blue passive galaxy. It is worth noting that our photometric data does not have the resolution to morphologically classify our galaxy population, meaning that some of our red passive galaxies might be passive spirals with colors similar to those of elliptical galaxies (e.g. Goto et al., 2003).

2.6 Summary and Conclusions

In this paper we use deep optical imaging and new MUSE spectroscopic data along with archival GMOS data to study the photometric and spectral properties of the merging cluster candidate SPT-CL J0307-6225, estimating redshifts for 69 new galaxy cluster members. We used the data to characterize (a) its merging history by means of a dynamical analysis and (b) its galaxy population by means of their spectroscopic and photometric properties.

With respect to the merging history, we were able to confirm the merging state of the cluster and conclude that:

- Using the galaxy surface density map of the RCS galaxies we can see a bi-modality in the galaxy distribution. However, the cluster does not show signs of substructures along the line-of-sight.

- We assign galaxy members to each substructure by means of the DBSCAN algorithm. We

name the two main substructures as 0307-6225N and 0307-6225S, referring to the northern and southern overdensities, respectively.

- For each substructure we measured the redshift, velocity dispersion and velocity-derived masses from scaling relations. We find a mass ratio of $M_S/M_N \approx 1.3$ and a velocity difference of $v_N - v_S = 342$ km s^{-1} between the northern and southern structures.
- To estimate the time since collision we use the MCMAC algorithm, which gave us the times for an outgoing and incoming system. By means of hydrodynamical simulations we constrained the most likely time to that of an outgoing system with TSP=$0.96^{+0.31}_{-0.18}$ Gyr.
- The outgoing configuration is also supported by the comparison between the observed and simulated X–ray morphologies. This comparison between the X–ray morphologies also provide a constraint on the masses, where a merger with a mass ratio of 1:3 seems more likely than that of a 1:1 mass merger.

With respect to the galaxy population, we find that:

- EL galaxies are located preferentially near the cluster cores (projected separations), where the average low peculiar velocities of red SF galaxies indicates that they were most likely accreted before the merger between 0307-6225N and 0307-6225S occurred.
- EL galaxies on 0307-6225N have smaller peculiar velocities and older stellar populations than those of 0307-6225S, where in the latter it appears that blue SF galaxies were either recently accreted or are in the process of being accreted.
- 0307-6225S shows two possible BCGs, which are very close in projected space. The magnitude and velocity differences between them are ~ 0 mag and $\sim$674 km s^{-1}, respectively, with one of them having a peculiar velocity close to 0 km s^{-1} with respect to 0307-6225S, while the other is close to the estimated $1\sigma_v$. However, the velocity distribution of the cluster shows no signs of being perturbed. This suggests that 0307-6225S could be the result of a previous merger which was at its last stage when the observed merger occurred.
- With respect to the in-between region, the galaxy population is comprised mostly of red galaxies, with the population of blue galaxies classified as passive or PSB, with colors close to the RCS.

In summary, our work supports a nearly face-on, in the plane of the sky, major merger scenario for SPT-CL J0307-6225. This interaction accelerates the quenching of galaxies as a result of a rapid enhancement of their star formation activity and the subsequent gas depletion. This is in line with literature findings indicating that the dynamical state of a cluster merger has a strong impact on galaxy population. Of particular importance is to differentiate dynamically young and old mergers. Comparisons between such systems will further increase our understanding on the connection between mergers and the quenching of star formation in galaxies. In future studies, we will replicate the analysis performed on SPT-CL J0307-6225, to a larger cluster sample, including the most disturbed cluster candidates on the SPT sample. These studies will be the basis for a comprehensive analysis of star formation in mergers with a wide dynamical range.

2.7 Appendix: Completeness of MUSE catalog

Since our aim is to look at the properties of the galaxy population, we need to first characterise a limiting magnitude to define that population. Fig. 2.2 shows that the population of spectroscopic RS galaxies stops at $i_{\rm auto} \approx 22.8$, with blue galaxies going as deep as $i_{\rm auto} \approx 23.3$. In order to find out the limiting magnitude we want to use, we compare our photometric catalog inside the cubes footprints within magnitude bins, checking the fraction of spectroscopically confirmed galaxies within each bin. This check allows us to (1) validate our method for selecting RCS members, which will become important when looking for substructures (see §2.3.3), and (2) to look for potential cluster members not found by MARZ.

In Fig. 2.18 we show the estimated completeness within different magnitude bins, where the lines are color coded according to the galaxy population. Continuous lines represent all the galaxies with spectroscopic information with MUSE, while dashed lines are only cluster members.

For the red galaxies, we have a completeness of 100% up to $m^* + 1$, with one galaxy at $i_{\rm auto} < m^*$ and z = 0.611 ($\Delta v = 5,940$ km s^{-1}), while at $m^* \leq i_{\rm auto} < m^* + 1$ we have two galaxies at $z = 0.612$ and $z = 0.716$ ($\Delta v = 6,130$ km s^{-1} and $\Delta v = 25,867$ km s^{-1}, respectively). The latter one showed similar properties to the galaxies that belong to the cluster; size, visual color and spatially close to the BCG. Fig. 2.3 shows the spectra of this galaxy in cyan. Its $r - i$ color index was also part of, towards the higher end, the rather generous width used for our RCS catalog. At $i_{\rm auto} \geq m^* + 2$, galaxies look like they belong to the cluster, but do not show strong spectral features with which we can estimate the redshift accurately. Blue galaxies show a similar trend as for red galaxies, with completeness of 100% up to $m^* + 1$, and over 80% at $i_{\rm auto} < m^* + 2$. However most of the blue galaxies, unlike red galaxies, do not belong to the cluster.

2.8 Appendix: Comparison to GMOS data

To estimate the redshifts of the 35 from the GMOS spectroscopic archival data we use the IRAF task FXCOR. For these estimations we use 4 template spectra from the IRAF package RVSAO; *eltemp* and *sptemp* that are composites of elliptical and spiral galaxies, respectively, produced with the FAST spectrograph for the Tillinghast Telescope (Fabricant et al., 1998); *habtemp0* produced with the HECTOSPEC spectrograph for the MMT as a composite of absorption line galaxies (Fabricant et al., 1998); and a synthetic galaxy template *syn4* from stellar spectra libraries constructed using stellar light ratios (Quintana et al., 2000). The redshifts are solved in the spectrum mode of FXCOR taking the r-value (Tonry & Davis, 1979) as the main reliability factor of the correlation following Quintana et al. (2000). They consider $r > 4$ as the limit for a reliable result, here we use the resulting velocity only if it follows that (a) at least 3 out of the 4 estimated redshifts from the templates agree with the heliocentric velocity within ± 100 km s^{-1} from the median and (b) at least 2 of those have $r > 5$. Finally, the radial heliocentric velocity of the galaxy and its error is calculated as the mean of the values from the "on-redshift" correlations.

Out of the 35 GMOS spectra, we have 12 galaxies with a common MUSE measurement, 10 belonging to the cluster. We use these 12 galaxies in common to compare the results given by

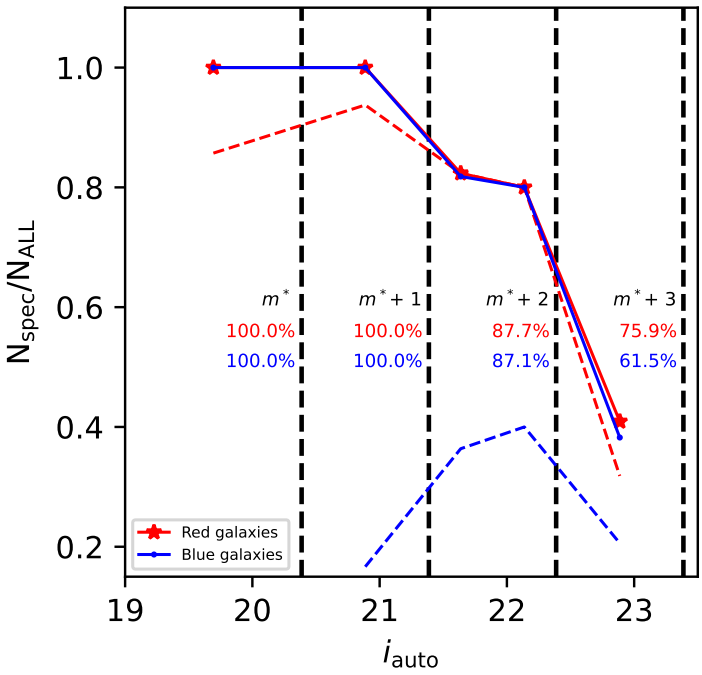

Figure 2.18: Ratio of the spectroscopically confirmed members with respect to the galaxies from our catalog (photometrically and spectroscopically selected) at different bins of magnitudes. Continuous lines show the completeness of all the MUSE catalog with measured redshifts, while dashed lines are only those galaxies identified as cluster members. Lines are color coded according to the galaxy population. Black dashed lines denote the limits for m^*, m^*+1, m^*+2 and m^*+3, with the percentages being the accumulated completeness for a given limit of the MUSE catalog.

FXCOR and MARZ, obtaining a mean difference of 60 ± 205 km s^{-1} on the heliocentric reference frame. Fig. 2.19 shows the estimated redshifts of these sources with the two different methods. Only one galaxy shows a velocity difference higher than 3σ. Excluding this galaxy from the analysis gives a mean velocity difference of 4 ± 96 km s^{-1}.

With respect to the redshift measurements presented in Bayliss et al. (2016), we find that the velocity difference within ± 5000 km s^{-1} from their redshift estimation of the cluster ($z_{cl} = 0.5801$) is of $|\Delta cz| \approx 300$ km s^{-1} with a big dispersion. Regarding potential cluster members, we select only galaxies where the redshifts reported by Bayliss et al. (2016) and the ones estimated using FXCOR have a difference smaller than 500 km s^{-1}, which at $z_{cl} = 0.5801$ corresponds to a difference of $\sim$0.1%. This eliminates 2 potential cluster members, one from each method. Meaning that we add 8 cluster members from the GMOS data in the final sample

2.9 Appendix: Catalog of spectroscopically confirmed objects

Table 2.8 shows the properties of the 139 objects with spectroscopic information from MUSE (117) or GMOS (22) within the field. The "Field" column is a combination of the instrument plus the number of the observed field. In the case of MUSE data this corresponds to the data cubes

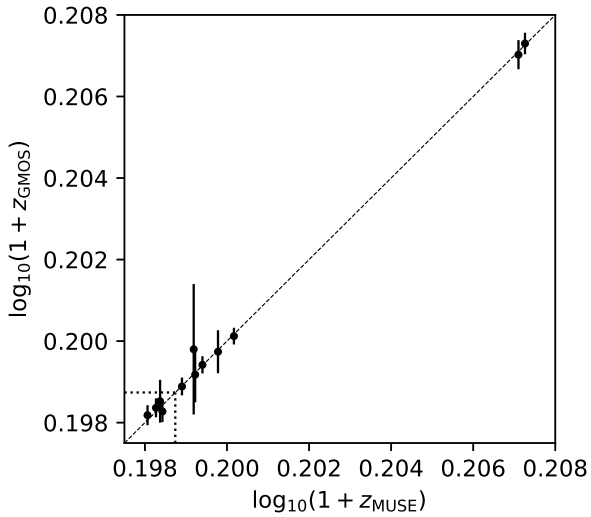

Figure 2.19: Redshift comparison between the GMOS spectra (FXCOR) and the MUSE spectra (MARZ) for the 12 galaxies in common. The dashed line shows the 1-to-1 relation while the dotted lines mark the redshift of the cluster at $z_{cl} = 0.5803$.

shown in Fig. 2.1 and Table 2.1, whereas in the case of GMOS, this corresponds to the first or second observed mask (see Bayliss et al., 2016). The ID column are the object's unique ID within the observed field. Redshifts for MUSE objects correspond to the ones derived using MARZ, while for GMOS they correspond to the ones derived using FXCOR. Magnitudes are the derived using SExtractor's MAG_AUTO parameter, while color indexes are derived using SExtractor's `mag_aper` parameter, with a fixed aperture of ~ 40 kpc at the cluster's redshift. The last column, Q, corresponds to the cluster membership, with 1 for galaxies within the ± 3000 km s^{-1} cut from the cluster's redshift, and 0 otherwise.

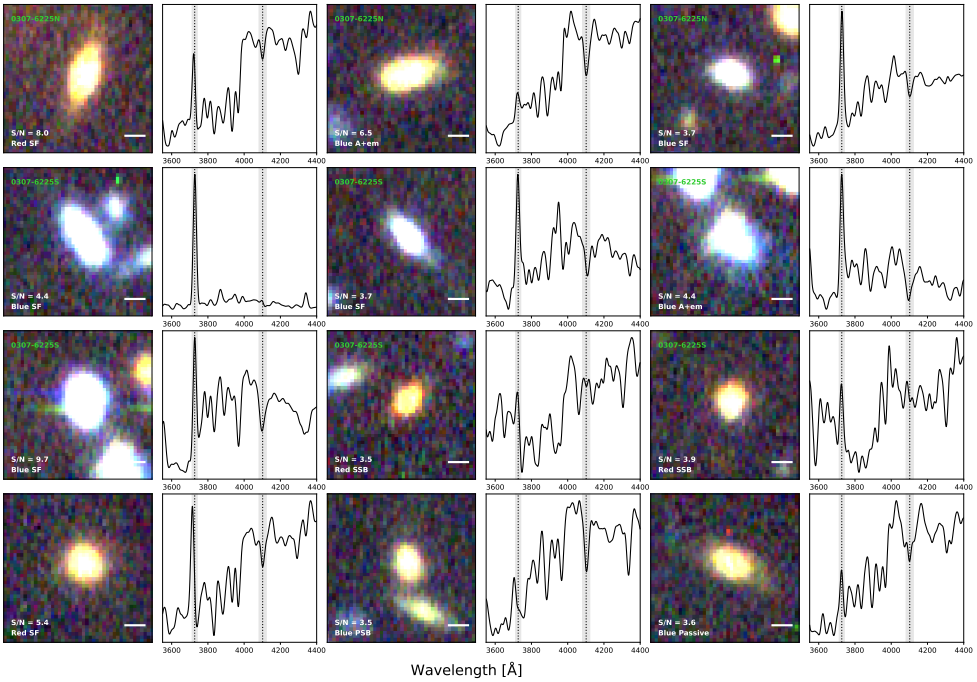

Figure 2.20: Pseudo-color crop images (box size of 7×7 arcsec^2) of the SF, A+em, SSB and PSB galaxies from our sample (plus one blue passive galaxy). On the bottom left of each image the spectral type of the galaxy is shown, with a white bar on the bottom right representing the scale size of 1 arcsec. Galaxies on the top and middle row belong to 0307-6225S and 0307-6225N, respectively, while galaxies on the bottom row are those that do not belong to any of the aforementioned. The doppler corrected spectra of each galaxy it's shown to the right, with the dotted lines showing the Hδ and [OII] λ3727 Å lines and the gray area marking the width of each line that we use to classify the galaxy (Balogh et al., 1999).

Table 2.8: Properties of the spectroscopically confirmed objects. The first and second columns are the sky coordinates of the objects. Columns (3) and (4) are the instrument (along with the corresponding field) and the object ID within the field. The heliocentric redshifts are listed in column (5). Columns (6) through (10) are the derived magnitudes and the $g-r$, $r-i$ color indexes (from aperture magnitudes). The last column corresponds to the cluster membership, where 1 means galaxies within the ± 3000 km s^{-1} cut from the cluster's redshift $z_{\rm cl} = 0.5803$.

R.A. (J2000)	Dec. (J2000)	Field	ID	$z_{\rm Helio}$	$g_{\rm auto}$ mag	$r_{\rm auto}$ mag	$i_{\rm auto}$ mag	$g-r$ mag	$r-i$ mag	Q
03:07:17.96	-62:27:12.19	MUSE-1	01	0.6116	22.992	21.859	21.039	1.092	0.812	0
03:07:13.92	-62:27:28.50	MUSE-1	02	0.3711	19.972	20.888	20.336	-4.737	0.567	0
03:07:16.80	-62:26:23.12	MUSE-1	04	0.5761	25.292	23.188	22.13	2.198	1.100	1
03:07:16.12	-62:26:25.34	MUSE-1	06	0.5855	22.298	20.752	19.94	1.572	0.808	1
03:07:16.02	-62:26:28.51	MUSE-1	07	0.5716	23.813	22.289	21.487	1.412	0.816	1
03:07:14.51	-62:26:26.89	MUSE-1	09	0.5761	24.473	22.600	21.651	1.884	0.961	1
03:07:14.27	-62:26:30.88	MUSE-1	13	0.5780	23.503	21.574	20.718	1.915	0.868	1
03:07:13.04	-62:26:33.31	MUSE-1	17	0.2153	21.541	20.511	20.163	1.012	0.359	0
03:07:14.69	-62:26:38.69	MUSE-1	28	0.6128	24.248	23.121	22.272	0.836	0.626	0
03:07:14.50	-62:26:39.97	MUSE-1	29	0.2405	22.284	21.324	20.911	0.935	0.414	0
03:07:15.11	-62:26:40.05	MUSE-1	30	0.5694	21.449	21.037	20.809	0.419	0.234	1
03:07:16.57	-62:26:41.39	MUSE-1	35^{S_2}	0.5827	22.270	20.261	19.329	1.878	0.950	1
03:07:14.87	-62:26:43.18	MUSE-1	38	0.5729	22.122	21.722	21.598	0.393	0.104	1
03:07:13.68	-62:26:43.71	MUSE-1	40	0.5806	24.252	22.775	21.914	1.474	0.897	1
03:07:18.09	-62:26:45.12	MUSE-1	42	0.5961	23.000	22.413	22.129	0.532	0.327	1
03:07:16.23	-62:26:47.06	MUSE-1	45	0.5813	24.609	22.317	21.333	2.235	0.979	1
03:07:16.75	-62:26:47.43	MUSE-1	46^{S_1}	0.5792	22.222	20.303	19.344	1.912	0.984	1
03:07:16.61	-62:26:49.61	MUSE-1	47	0.5740	24.553	22.549	21.600	2.008	0.950	1
03:07:19.06	-62:26:48.68	MUSE-1	50	0.5792	25.137	22.992	22.082	2.830	0.986	1
03:07:16.22	-62:26:50.30	MUSE-1	51	0.5786	24.482	22.556	21.596	2.091	0.967	1
03:07:20.30	-62:26:51.05	MUSE-1	54	0.3284	23.391	22.886	22.846	0.319	0.014	0
03:07:17.33	-62:26:52.89	MUSE-1	59	0.1599	23.182	22.511	22.081	0.752	0.473	0
03:07:16.57	-62:26:54.58	MUSE-1	60	0.0000	23.038	21.449	20.528	1.635	0.925	0
03:07:17.62	-62:26:55.52	MUSE-1	61	0.5830	25.192	23.469	22.601	2.023	0.855	1
03:07:14.52	-62:26:55.41	MUSE-1	63	0.5920	23.568	23.349	23.142	0.133	0.179	1
03:07:20.82	-62:26:56.46	MUSE-1	65	0.5778	24.663	22.861	21.961	1.445	0.882	1
03:07:14.26	-62:26:59.07	MUSE-1	66	0.3713	22.570	21.772	21.617	0.790	0.157	0
03:07:20.45	-62:26:58.98	MUSE-1	68	0.6110	23.229	21.268	20.307	1.965	0.962	0
03:07:18.42	-62:26:59.81	MUSE-1	69	0.5749	24.896	23.217	22.356	1.530	0.948	1
03:07:14.37	-62:27:03.47	MUSE-1	74	0.5879	24.141	23.651	23.327	0.493	0.225	1
03:07:15.77	-62:27:09.28	MUSE-1	80	0.5949	22.051	21.692	21.540	0.355	0.154	1
03:07:12.74	-62:27:10.74	MUSE-1	82	0.5766	23.951	21.885	20.898	2.124	0.982	1
03:07:14.87	-62:27:11.27	MUSE-1	84	0.5855	25.151	23.060	22.152	1.918	0.960	1
03:07:15.89	-62:27:22.14	MUSE-1	92	0.5807	24.375	22.546	21.520	1.649	1.044	1
03:07:15.65	-62:27:22.78	MUSE-1	93	0.5834	23.585	21.895	20.951	1.686	0.947	1
03:07:20.02	-62:25:55.68	MUSE-2	01	0.2151	19.335	18.233	17.874	1.115	0.366	0
03:07:17.20	-62:25:02.68	MUSE-2	04	0.5787	24.702	22.803	22.021	1.968	0.811	1
03:07:16.48	-62:25:03.86	MUSE-2	05	0.4989	22.090	20.726	20.156	1.367	0.568	0
03:07:17.02	-62:25:05.38	MUSE-2	08	0.5724	24.943	23.234	22.432	1.539	0.736	1
03:07:21.08	-62:25:13.45	MUSE-2	17	0.5850	24.040	23.281	22.917	0.740	0.186	1
03:07:16.80	-62:25:18.55	MUSE-2	21	0.5867	24.156	23.286	23.073	0.965	0.052	1
03:07:22.93	-62:25:18.19	MUSE-2	23	0.5856	25.186	23.491	22.706	1.767	0.733	1
03:07:21.73	-62:25:19.71	MUSE-2	25	0.5829	24.928	23.157	22.273	1.763	0.933	1
03:07:19.30	-62:25:26.50	MUSE-2	29	0.0001	20.574	18.993	17.758	1.570	1.243	0
03:07:15.16	-62:25:26.79	MUSE-2	37	0.2146	22.552	21.950	21.804	0.606	0.160	0
03:07:21.16	-62:25:31.01	MUSE-2	38	0.5894	22.802	21.098	20.166	1.695	0.937	1
03:07:21.88	-62:25:36.13	MUSE-2	48	0.5829	24.595	22.628	21.681	1.640	0.918	1
03:07:22.26	-62:25:37.21	MUSE-2	49	0.5749	23.628	21.943	20.991	1.698	0.956	1
03:07:17.23	-62:25:40.65	MUSE-2	51	0.5705	24.285	22.422	21.494	1.881	0.952	1
03:07:21.00	-62:25:36.43	MUSE-2	55	0.5922	24.272	22.430	21.593	1.927	0.846	1
03:07:17.81	-62:25:46.31	MUSE-2	61	0.5780	23.132	21.614	20.797	1.528	0.816	1
03:07:19.81	-62:25:45.89	MUSE-2	63	0.5910	24.260	23.520	23.285	0.959	-0.010	1
03:07:16.65	-62:25:48.66	MUSE-2	66	0.5899	25.097	23.280	22.409	2.206	0.842	1
03:07:19.27	-62:25:48.79	MUSE-2	67	0.5797	24.414	22.930	22.072	1.427	0.774	1
03:07:20.70	-62:25:50.35	MUSE-2	68	0.5786	21.944	20.670	19.892	1.313	0.837	1
03:07:17.89	-62:25:51.37	MUSE-2	70	0.2156	22.441	21.887	21.752	0.532	0.137	0
03:07:14.00	-62:25:53.24	MUSE-2	71	0.0002	20.385	19.029	18.514	1.343	0.526	0
03:07:16.65	-62:25:54.11	MUSE-2	78	0.5690	25.203	23.675	22.783	1.251	0.948	1
03:07:20.69	-62:25:53.83	MUSE-2	79	0.5797	23.168	22.083	21.708	1.079	0.406	1
03:07:17.81	-62:25:56.62	MUSE-2	85	0.5859	23.620	21.788	20.898	1.767	0.890	1
03:07:21.40	-62:25:58.09	MUSE-2	87	0.5795	24.844	22.919	22.016	2.217	0.872	1
03:07:22.17	-62:26:00.71	MUSE-2	90	0.5736	25.099	23.136	22.234	2.199	1.002	1
03:07:17.12	-62:26:01.83	MUSE-2	93	0.3696	22.866	22.028	21.840	0.855	0.201	0
03:07:18.83	-62:26:03.72	MUSE-2	97	0.2754	23.377	22.672	22.485	0.724	0.170	0
03:07:19.20	-62:24:20.28	MUSE-3	07	-0.0001	21.096	19.504	18.474	1.566	1.037	0
03:07:22.83	-62:24:18.61	MUSE-3	08	0.5830	23.713	21.908	21.033	1.815	0.894	1
03:07:26.09	-62:24:23.49	MUSE-3	16	0.3977	23.064	21.461	20.977	1.488	0.478	0
03:07:23.40	-62:24:27.73	MUSE-3	19	0.5790	23.336	21.591	20.792	1.780	0.829	1
03:07:23.14	-62:24:29.86	MUSE-3	23	0.5783	23.821	22.027	21.164	1.729	0.855	1
03:07:20.55	-62:24:32.88	MUSE-3	24	0.5720	24.798	23.100	22.272	1.597	0.825	1
03:07:26.25	-62:24:37.72	MUSE-3	30	0.8144	24.928	24.661	24.227	0.215	-0.075	0
03:07:21.46	-62:24:42.92	MUSE-3	34	0.5819	24.244	22.718	22.052	1.583	0.632	1
03:07:23.08	-62:24:45.50	MUSE-3	39	0.2843	25.394	23.964	23.418	1.922	0.450	0
03:07:23.09	-62:24:47.93	MUSE-3	42	0.5753	24.943	23.012	22.102	2.120	0.920	1
03:07:21.52	-62:24:50.91	MUSE-3	48	0.8614	23.911	23.659	23.156	0.201	0.591	0
03:07:20.51	-62:24:50.94	MUSE-3	49	0.5779	23.558	22.119	21.186	1.279	0.934	1
03:07:22.81	-62:24:54.49	MUSE-3	51	0.1385	23.274	22.694	22.561	0.576	0.106	0
03:07:22.94	-62:24:59.24	MUSE-3	58	0.5899	23.944	22.708	22.196	1.254	0.513	1
03:07:22.83	-62:25:01.47	MUSE-3	62	0.5810	24.242	23.165	22.819	0.987	0.403	1
03:07:25.25	-62:25:03.18	MUSE-3	64	0.5864	23.979	22.442	21.643	1.631	0.803	1
03:07:23.71	-62:25:06.47	MUSE-3	67	0.3701	21.651	20.579	20.236	1.094	0.346	0
03:07:24.01	-62:25:07.89	MUSE-3	68	0.7218	23.492	22.811	22.385	0.582	0.440	0
03:07:23.33	-62:25:06.18	MUSE-3	73	0.6039	23.649	22.947	22.548	0.607	0.489	0
03:07:22.16	-62:25:09.17	MUSE-3	76	0.5725	25.233	23.313	22.376	1.834	0.945	1
03:07:26.14	-62:23:18.43	MUSE-4	03	0.0003	21.154	19.698	19.118	1.429	0.588	0
03:07:26.83	-62:23:21.89	MUSE-4	05	-0.0001	21.896	20.397	19.85	1.475	0.566	0
03:07:24.84	-62:23:22.06	MUSE-4	06	0.8031	21.605	21.330	21.035	0.268	0.303	0
03:07:25.41	-62:23:25.29	MUSE-4	12	0.8032	23.058	22.731	22.206	0.307	0.551	0
03:07:21.89	-62:23:28.04	MUSE-4	13	0.5771	22.956	21.217	20.409	1.712	0.817	1
03:07:26.90	-62:23:28.12	MUSE-4	14	0.5860	24.507	22.764	21.955	1.719	0.821	1
03:07:30.15	-62:23:31.80	MUSE-4	15	0.5790	23.320	22.259	21.585	1.081	0.684	1
03:07:28.25	-62:23:36.05	MUSE-4	16	0.5867	24.186	22.344	21.405	1.873	0.945	1
03:07:26.92	-62:23:36.95	MUSE-4	18	0.5733	24.684	22.822	21.887	1.848	0.939	1
03:07:28.56	-62:23:37.56	MUSE-4	19	0.5837	24.066	22.388	21.656	1.716	0.774	1
03:07:23.54	-62:23:39.02	MUSE-4	21	0.5345	23.556	22.927	22.717	0.623	0.246	0
03:07:29.30	-62:23:41.94	MUSE-4	23	-0.0000	23.342	21.804	20.677	1.410	1.144	0
03:07:22.44	-62:23:43.38	MUSE-4	24	0.5803	24.648	22.691	21.821	1.540	0.800	1
03:07:26.13	-62:23:44.20	MUSE-4	25	0.5823	24.632	22.951	22.073	1.452	0.785	1
03:07:28.48	-62:23:44.96	MUSE-4	26	0.5841	23.618	22.073	21.463	1.525	0.615	1
03:07:22.96	-62:23:48.47	MUSE-4	28	0.1160	19.287	18.934	18.727	0.338	0.206	0
03:07:22.87	-62:23:57.32	MUSE-4	32	0.1160	22.113	21.097	20.812	0.954	0.275	0
03:07:26.77	-62:23:51.79	MUSE-4	33	0.5759	23.908	22.160	21.280	1.667	0.888	1
03:07:26.15	-62:23:52.93	MUSE-4	34	0.7777	24.493	23.999	23.557	0.657	0.337	0
03:07:25.74	-62:23:54.13	MUSE-4	35	0.5815	23.969	22.147	21.224	1.659	0.892	1
03:07:27.18	-62:23:54.43	MUSE-4	37	0.5846	23.364	22.609	22.225	0.663	0.417	1
03:07:22.48	-62:24:04.09	MUSE-4	40	0.3700	25.763	23.961	23.271	2.543	0.788	0
03:07:27.50	-62:23:59.31	MUSE-4	44	0.5779	23.718	23.318	23.221	0.445	-0.117	1
03:07:24.51	-62:24:00.96	MUSE-4	46	0.5830	25.350	23.431	22.477	1.936	1.023	1
03:07:23.86	-62:24:02.18	MUSE-4	47	0.5802	24.326	22.439	21.528	1.890	0.934	1
03:07:24.76	-62:24:02.30	MUSE-4	48	0.5780	25.216	23.322	22.370	1.979	1.044	1
03:07:26.41	-62:24:03.99	MUSE-4	49	0.3706	23.287	21.736	21.318	1.543	0.457	0
03:07:28.06	-62:24:03.86	MUSE-4	50	0.5819	24.889	22.947	22.140	1.978	0.730	1
03:07:28.16	-62:24:04.95	MUSE-4	51	0.5808	25.110	22.964	22.049	2.527	0.821	1
03:07:24.55	-62:24:06.62	MUSE-4	54	0.5821	24.272	22.389	21.518	1.816	0.868	1
03:07:24.18	-62:24:07.57	MUSE-4	56	0.5753	24.073	22.237	21.278	1.844	0.938	1
03:07:23.85	-62:24:10.12	MUSE-4	57^{N}	0.5809	21.834	19.980	19.014	1.893	0.993	1
03:07:25.10	-62:24:11.12	MUSE-4	60	0.5847	25.519	23.725	22.702	1.705	1.137	1
03:07:28.71	-62:24:00.09	GMOS-1	07	0.6728	23.967	23.583	22.932	0.380	0.674	0
03:07:19.96	-62:23:53.24	GMOS-1	09	0.6115	21.843	20.844	20.292	1.033	0.570	0
03:07:07.53	-62:24:35.80	GMOS-1	11	0.5837	22.640	21.539	20.904	1.102	0.650	1
03:07:10.56	-62:24:29.11	GMOS-1	12	0.5426	22.485	21.501	21.115	0.971	0.397	0
03:07:30.38	-62:25:23.19	GMOS-1	13	0.4893	24.082	22.990	22.305	1.063	0.708	0
03:07:29.24	-62:25:04.10	GMOS-1	14	0.5732	23.335	22.153	21.477	1.204	0.674	1
03:07:26.39	-62:25:37.16	GMOS-1	15	0.5815	24.048	22.468	21.734	1.567	0.729	1
03:07:27.26	-62:25:13.13	GMOS-1	17	0.5807	23.876	22.555	22.108	1.283	0.444	1
03:07:19.96	-62:24:50.10	GMOS-1	18	0.4669	22.338	22.203	21.867	0.175	0.342	0
03:07:30.26	-62:26:01.74	GMOS-1	23	0.4999	23.540	22.904	22.794	0.531	-0.004	0
03:07:22.09	-62:28:03.40	GMOS-1	30	0.5010	23.523	22.756	22.052	0.738	0.684	0
03:07:21.13	-62:27:49.79	GMOS-1	31	0.5709	22.413	21.213	20.553	1.165	0.652	1
03:06:59.11	-62:27:39.75	GMOS-1	33	0.5347	22.723	22.663	22.309	0.032	0.400	0
03:07:13.96	-62:28:30.21	GMOS-1	34	0.6023	23.281	22.520	21.927	0.702	0.599	0
03:07:06.69	-62:24:36.80	GMOS-2	09	0.4743	23.406	22.474	21.841	0.850	0.672	0
03:07:26.50	-62:25:18.88	GMOS-2	10	0.5804	23.213	21.235	20.273	2.007	0.961	1
03:07:08.71	-62:24:51.92	GMOS-2	12	0.5752	22.124	21.322	20.898	0.792	0.420	1
03:07:25.63	-62:25:43.17	GMOS-2	14	0.6392	23.659	21.941	20.965	1.618	0.967	0
03:07:03.68	-62:25:35.52	GMOS-2	15	0.8106	22.900	21.932	21.234	0.957	0.689	0
03:07:36.22	-62:25:54.68	GMOS-2	20	0.5810	22.772	21.325	20.757	1.381	0.563	1
03:07:09.78	-62:27:02.67	GMOS-2	23	0.6405	25.008	23.449	22.752	1.355	0.706	0
03:07:17.94	-62:27:54.13	GMOS-2	27	0.4043	24.333	22.794	22.088	1.444	0.640	0

[N] BCG of 0307-6225N
[S_1] First BCG of 0307-6225S
[S_2] Second BCG of 0307-6225S

CHAPTER 3

THE PSZ-MCMF CATALOGUE OF *PLANCK* CLUSTERS OVER THE DES REGION

D. Hernández-Lang, M. Klein, J. J. Mohr, S. Grandis, J.-B. Melin, P. Tarrío, M. Arnaud, G.W. Pratt, T. M. C. Abbott, et al., *Monthly Notices of the Royal Astronomical Society* 525, 24-43, 2023

We present the first systematic follow-up of *Planck* Sunyaev-Zeldovich effect (SZE) selected candidates down to signal-to-noise (S/N) of 3 over the 5000 deg^2 covered by the Dark Energy Survey. Using the `MCMF` cluster confirmation algorithm, we identify optical counterparts, determine photometric redshifts and richnesses and assign a parameter, $f_{\rm cont}$, that reflects the probability that each SZE-optical pairing represents a random superposition of physically unassociated systems rather than a real cluster. The new PSZ-MCMF cluster catalogue consists of 853 `MCMF` confirmed clusters and has a purity of 90%. We present the properties of subsamples of the PSZ-MCMF catalogue that have purities ranging from 90% to 97.5%, depending on the adopted $f_{\rm cont}$ threshold. Halo mass estimates M_{500}, redshifts, richnesses, and optical centers are presented for all PSZ-MCMF clusters. The PSZ-MCMF catalogue adds 589 previously unknown *Planck* identified clusters over the DES footprint and provides redshifts for an additional 50 previously published *Planck* selected clusters with S/N>4.5. Using the subsample with spectroscopic redshifts, we demonstrate excellent cluster photo-z performance with an RMS scatter in $\Delta z/(1+z)$ of 0.47%. Our `MCMF` based analysis allows us to infer the contamination fraction of the initial S/N>3 *Planck* selected candidate list, which is ∼50%. We present a method of estimating the completeness of the PSZ-MCMF cluster sample. In comparison to the previously published *Planck* cluster catalogues, this new S/N>3 `MCMF` confirmed cluster catalogue populates the lower mass regime at all redshifts and includes clusters up to z∼1.3.

3.1 Introduction

The intracluster medium (ICM) in galaxy clusters can be detected through what are now easily observed ICM signatures, providing a means to select cluster samples based on their ICM properties. At high temperatures of up to T $\sim 10^8$ K (for massive clusters), photons are emitted at X–ray wavelengths via thermal bremsstrahlung. Moreover, the ICM can leave an imprint on the cosmic microwave background (CMB). At mm-wavelengths, it is possible to study galaxy clusters via the thermal Sunyaev-Zeldovich effect (SZE; Sunyaev & Zeldovich, 1972), which is produced by inverse Compton scattering of CMB photons by hot electrons in the ICM.

Large X–ray selected galaxy cluster catalogues have been created using X-ray imaging data from the ROSAT All Sky Survey and the XMM-Newton telescope (e.g. Piffaretti et al., 2011; Klein et al., 2019; Finoguenov et al., 2020; Koulouridis et al., 2021) as well as the recently launched eROSITA mission (Brunner et al., 2022; Liu et al., 2022; Klein et al., 2022). The *Planck* mission mapped the whole sky between 2009 to 2013 in mm and infrared wavelengths, with the goal of studying CMB anisotropies. The latest cluster catalogue released by the *Planck* collaboration is the second *Planck* catalogue of Sunyaev-Zeldovich sources (PSZ2; Planck Collaboration et al., 2016a), containing over 1600 cluster candidates down to a signal-to-noise ratio (S/N) of 4.5, detected from the 29-month full-mission data. Other projects such as the South Pole Telescope (SPT; Carlstrom et al., 2011) and the Atacama Cosmology Telescope (ACT; Marriage et al., 2011) have also been used to create large SZE selected cluster catalogues (e.g. Bleem et al., 2015; Hilton et al., 2021; Bleem et al., 2020).

Although ICM-based cluster selection from an X-ray or SZE sky survey is efficient, the resulting candidate lists must be optically confirmed to extract galaxy based observables such as precise photometric redshifts (e.g. Staniszewski et al., 2009; High et al., 2010; Song et al., 2012b; Liu et al., 2015; Klein et al., 2019, 2022). The optical followup also allows for a cleaning or removal of the contaminants (falsely identified clusters) from ICM selected samples, because noise fluctuations in the ICM candidate lists do not have physically associated galaxy systems. It is possible for a noise fluctuation in the ICM candidate list to overlap by chance with a physically unassociated galaxy system. With the use of the Multi-Component Matched Filter followup technique (`MCMF`; Klein et al., 2018, 2019), it is possible to account for this random superposition possibility for each ICM cluster candidate and to deliver empirically estimated, precise and accurate measurements of the residual contamination in the final cluster catalogue.

To enable efficient optical followup and precise estimates of the purity of the final confirmed cluster catalogue, large and homogeneous photometric datasets are beneficial. The Dark Energy Survey (DES; Abbott et al., 2016) covers ~5000 deg^2 with deep, multiband imaging in g, r, i, z, Y bands with the DECam instrument (Flaugher et al., 2015). These imaging data are processed and calibrated using the DES data management system (Morganson et al., 2018), and to date two major data releases have taken place (Abbott et al., 2018, 2021).

Large, homogeneous multi-band imaging surveys also support the direct galaxy-based selection of cluster catalogues (e.g., Gladders et al., 2007; Rykoff et al., 2014; Maturi et al., 2019; Wen & Han, 2022). However, without a second cluster observable, as in the case of the ICM based selection followed up by optical confirmation, it is more challenging to empirically estimate or control the contamination of the final cluster catalogue. One can use statistical comparisons to

well understood ICM-based samples (see SPTxRM analyses in Grandis et al., 2020, 2021) to estimate the contamination (as well as the mass completeness modeling) or one can attempt to simulate the contamination of the cluster sample directly (e.g., Song et al., 2012a; Crocce et al., 2015; DeRose et al., 2019), in which case the contamination estimates are impacted by the level of realism of the simulations.

The utility of optically based cluster sample cleaning methods, like that available with the `MCMF` algorithm, becomes ever more central to the cluster catalogue creation as one considers lower signal to noise ICM signatures as cluster candidates, because these candidate samples are more contaminated with noise fluctuations. With an effective optically based cleaning method, it becomes possible to create dramatically larger confirmed cluster samples from a given X-ray or mm-wave survey, while still maintaining low levels of contamination (i.e., high sample purity). As an example, the X-ray cluster sample MARDY3 selected from ROSAT in combination with DES produced an increase of an order of magnitude in the number of ROSAT selected clusters over the DES area (Klein et al., 2019). Significant gains are currently being seen in the extraction of cluster samples from lower signal to noise candidate lists from the SPT-SZ 2500d and the SPTpol 500d survey (Klein et al. in prep, Bleem et al. in prep).

Leveraging the rich dataset provided by *Planck*, we have developed a new cluster candidate catalog that extends to lower signal-to-noise levels (S/N>3), enhancing the number of candidate clusters identified. However, extending the catalog to lower signal-to-noise levels leads to a higher number of spurious sources or noise fluctuations being classified as *Planck* detections, resulting in a decrease in the candidate catalog purity. To address this reduced purity, we utilize the DES dataset together with the `MCMF` cluster confirmation algorithm to confirm *Planck* clusters and to reject spurious sources.

In this analysis, we present the PSZ-MCMF[1] cluster catalog. To construct this catalogue, we extend the `MCMF` tool to deal with the larger positional uncertainties that come with *Planck* selected cluster candidates and then apply this tool to a *Planck* based candidate list down to S/N=3 using DES photometric data. In Section 3.2 we give a description of the DES and *Planck* data used. In Section 3.3 we describe the enhanced `MCMF` cluster confirmation method, while in Section 3.4 we report our findings. Finally, in Section 3.5 we summarise our findings and report our conclusions. Throughout this paper we adopt a flat ΛCDM cosmology with $\Omega_M = 0.3$ and $H_0 = 70$ km s^{-1} Mpc^{-1}.

3.2 Data

3.2.1 DES multi-band photometric data

In this work we use the DES Y3A2 GOLD photometric data, which is based on DES imaging data obtained from the first three years of the survey (Abbott et al., 2018). We employ g, r, i, z band photometry, which has 95% completeness limits of 23.72, 23.34, 22.78 and 22.25 mag, respectively. The YA32 GOLD catalogue has been optimized for cosmological studies with DES,

[1]PSZ stands for the *Planck* Sunyaev-Zeldovich cluster candidate list, whereas MCMF comes from the algorithm, which allows us to maximise the number of clusters from any given parent candidate list.

similar to the Y1A1 GOLD catalogue (Drlica-Wagner et al., 2018). Because we build upon the same `MCMF` cluster confirmation method applied in a ROSAT×DES analysis (Klein et al., 2019), we refer the reader to that source for further details of the filtering and handling of the optical multi-band data.

In summary, we make use of the single-object fitting photometry (SOF), which is based on the ngmix code (Sheldon, 2014). The photometry is performed by fitting a galaxy model for each source in each single epoch image of a given band at the same time, interpolating the point-spread functions (PSFs) at the location of each source. This fitting is done masking neighbouring sources. We make use of the star-galaxy separator included in the GOLD catalogs (Drlica-Wagner et al., 2018) and exclude unresolved objects with i <22.2 mag. We also make use of the masking provided by Y3A2 GOLD (similar to that described in Y1A1 GOLD, Drlica-Wagner et al., 2018) to exclude regions around bright stars.

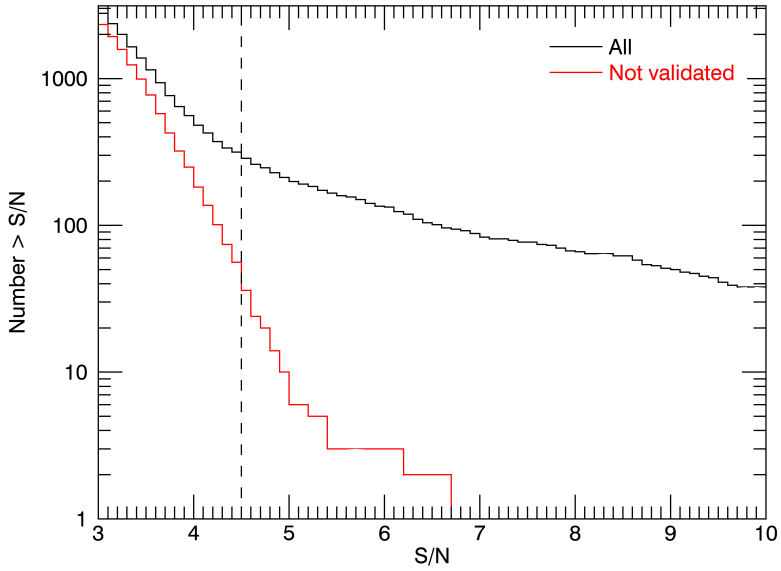

Figure 3.1: Cumulative number of *Planck* cluster candidates down to S/N=3. The solid black line represents the full sample of 3130 candidates. The solid red line represents the 2670 candidates that have not been validated in previous works via a simple cross-identification with known SZE and X-ray clusters (see text). The dashed line corresponds to the S/N=4.5 limit for the PSZ2 catalogue (Planck Collaboration et al., 2016a).

3.2.2 Planck SZE candidate list

We build a catalogue of *Planck* SZE sources with S/N>3 located within the DES footprint. The SZE catalogue is created using a matched multi-filter (MMF) approach (see for example Herranz et al., 2002; Melin et al., 2006), namely the MMF3 algorithm used and described in Planck

Collaboration et al. (2014b) and improved for the PSZ2 catalogue. The cluster detection is done using a combination of the *Planck* maps and assuming prior knowledge on the cluster profile. In this application of MMF3, we divide the sky into patches of $10^\circ \times 10^\circ$, generating 504 overlapping patches, and run the detection algorithm with two iterations; the first iteration detects the SZE signal and the second refines the SZE candidate position to allow for improved estimation of the S/N and other properties.

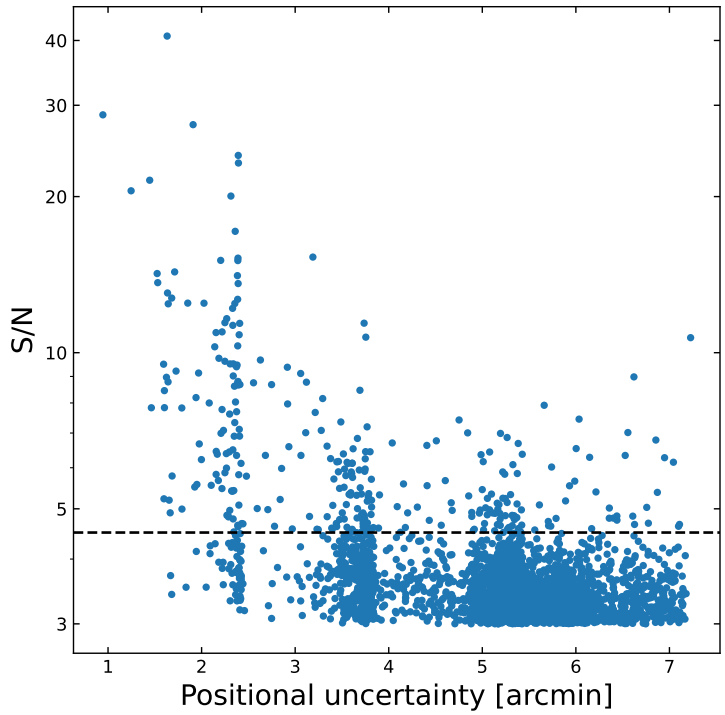

Figure 3.2: Positional uncertainty distribution in units of arcminutes versus *Planck* candidate S/N. The mean and median of the sources with S/N<4.5 are 5.26 arcmin and 5.36 arcmin, respectively. The black dashed line represents the threshold at S/N=4.5.

The filter works by combining the frequency maps from the *Planck* survey into a vector $\boldsymbol{Mx}$, where each component corresponds to a map at frequency ν_i with $i = 1, \ldots, N$ with N being the total number of maps. For *Planck*, we use the channel maps from 100 to 857 GHz, which correspond to the six highest-frequency maps.

For each cluster candidate at a given central position $\boldsymbol{x_0}$, the algorithm fits:

$$\boldsymbol{M_\nu x} = y_0 \, \boldsymbol{j_\nu} \, T_{\theta_c}(\boldsymbol{x} - \boldsymbol{x_0}) + \boldsymbol{n_\nu x} \tag{3.1}$$

where y_0 is the central value at position $\boldsymbol{x_0}$ and $\boldsymbol{n_\nu x}$ corresponds to the noise vector, which is the sum of the other emission components in the map that do not correspond to the cluster SZE (such as, e.g., primordial CMB anisotropies and diffuse galactic emission). The frequency dependence of the SZE is represented by $\boldsymbol{j_\nu}$. The spatial profile is defined as T_{θ_c}, with θ_c as the core radius. The assumed profile is chosen to be the universal pressure profile (Arnaud et al., 2010).

The filter is then employed to minimize the total variance estimate $\sigma^2_{\theta_C}$ on y_0 for each detected candidate, which yields an estimate $\hat{y}_0$. The S/N is then defined as $\hat{y}_0/\sigma_{\theta_C}$.

From this analysis we get the positions and associated uncertainties of the SZE sources plus the S/N and the SZE flux. At S/N>3, we get a total of 3130 *Planck* SZE sources (i.e. cluster candidates). Fig. 3.1 shows the cumulative number of cluster candidates (black) and unvalidated cluster candidates (red) for each S/N bin within the DES footprint. A candidate is considered to be validated if 1) it is less than 5 arcmin from a confirmed cluster (with known redshift) of the Meta-Catalog of SZ detected clusters (MCSZ) of the M2C database[2], or 2) it is less than 10 arcmin and less than θ_{500} from a confirmed cluster in the Meta-Catalog of X-Ray Detected Clusters of Galaxies (MCXC, Piffaretti et al., 2011). From the full sample of 3,130 candidates, 460 have been validated in this way (with 414 matching MCSZ clusters, and 46 matching MCXC only), while the remaining 2,670 are non-validated candidates but may nevertheless be real galaxy clusters.

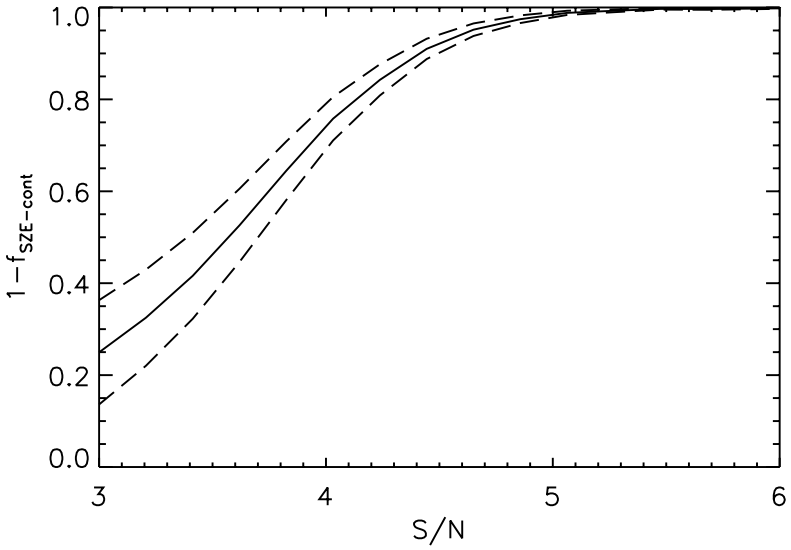

Figure 3.3: Purity as a function of signal-to-noise threshold of the cluster candidate list, estimated on *Planck* simulations. Dashed lines show the uncertainty of the estimated purity. The purity decreases from ∼1 at S/N>6 to ∼0.25 at S/N>3.

Fig. 3.2 contains the S/N versus the positional uncertainties of the *Planck* sources, where the black dashed line represents a S/N=4.5. The apparent structure of the positional uncertainty is due to the pixelization of the *Planck* maps. The detection algorithm filters the maps and finds the pixel which maximizes the S/N. The position assigned for a detection corresponds to the pixel center. The positional uncertainty is also computed on a pixelized grid.

We estimate the contamination of the *Planck* SZE candidate list using simulations. We use the *Planck* Sky Model (version 1.6.3; Delabrouille et al., 2013), to produce realistic all-sky mock observations. The simulations contain primary cosmic microwave background anisotropies,

[2] https://www.galaxyclusterdb.eu/m2c/

galactic components (synchrotron, thermal dust, free-free, spinning dust), extra-galactic radio and infrared point sources, and kinetic and thermal SZE. Each frequency map is convolved with the corresponding beam, and the instrumental noise consistent with the full mission is added. We run the thermal SZE detection algorithm down to S/N=3, and we match the candidate list with the input cluster catalogue adopting a 5 arcmin matching radius. We perform the matching after removing regions of the sky with high dust emission, leaving 75% of the sky available, and we only use input clusters with a measured Compton parameter Y in a circle of radius $5 \times R_{500}$[3], Y_{5R500}, above 2×10^{-4} arcmin^2. We adopt the SZE flux-mass relation

$$E^{-2/3}(z) D_A^2(z) Y_{5R500} = A \left[\frac{M_{500}}{3 \times 10^{14} h_{70}^{-1} M_\odot} \right]^{5/3} \tag{3.2}$$

with $A = 2.59 \times 10^{-5}\ h_{70}^{-1}$ Mpc^2 (see equation B.3 in Arnaud et al., 2010). $E(z) = H(z)/H_0$ is the Hubble parameter normalized to its present value and $D_A(z)$ is the angular diameter distance. The Compton parameter Y_{5R500} is given in steradians. We estimate the purity of the sample as the number of real clusters divided by the number of detected clusters. This ratio is computed for various S/N thresholds. The result is shown in Fig. 3.3. The uncertainty in purity is considered to be the difference between the best estimate and the lower limit of the purity (Fig. 11 and Fig. 12 in Planck Collaboration et al., 2016a, respectively) of the PSZ2 catalog, for the union 65% case. We fit this difference as a function of the contamination with a power law in the range S/N=4.5-20. We extrapolate this down to S/N=3. From here on, we refer to this contamination as the initial contamination: $f_{\rm SZE-cont}$. At high S/N threshold (S/N>6), the purity, $1 - f_{\rm SZE-cont}$, is close to unity. Reducing the S/N threshold to 4.5 leads to a purity close to 0.9, which is consistent with previous estimates (Planck Collaboration et al., 2016a). When reducing the threshold to S/N=3, we measure a purity of ∼25% corresponding to a contamination $f_{\rm SZE-cont} = 0.75$ in the simulations.

3.3 Cluster confirmation method

To identify optical counterparts and estimate photometric redshifts we use a modified version of the `MCMF` cluster confirmation algorithm on the *Planck* candidate list and DES-Y3 photometric catalogues. For each potential cluster, the radial position and the galaxy color weightings are summed over all cluster galaxy candidates to estimate the excess number of galaxies, or richness (λ), with respect to the background. Klein et al. (2019) contains further details of `MCMF` weights and the counterpart identification method.

We expect only a fraction 1-$f_{\rm SZE-cont}$ of the *Planck* candidates to be real clusters, with a large fraction ($f_{\rm SZE-cont} = 0.75$) corresponding to contaminants (we return to the value of $f_{\rm SZE-cont}$ in Section 3.4.4.2). Most of these contaminants have no associated optical system, but some will happen to lie on the sky near a physically unassociated optical system or a projection of unassociated galaxies along the line of sight. We refer to these contaminants as "random

[3] R_{500} is defined as the radius within which the density is 500 times the critical density

superpositions". The MCMF method has been designed to enable us to remove these contaminants from the *Planck* candidate list. To estimate the likelihood of a "random superposition" (e.g, a spurious *Planck* candidate being associated with one of the two cases above), we run MCMF at random positions in the portion of the sky survey that lies away from the candidates. With this information we can reconstruct the frequency and redshift distribution of optical systems, and this allows us to estimate the probability that each candidate is a contaminant (see details in Section 3.3.2.2).

3.3.1 Cluster confirmation with MCMF

In the MCMF method the sky coordinates of the cluster candidates are used to search the multi-band photometric catalogues with an associated galaxy red sequence (RS) model, to estimate galaxy richness λ as a function of redshift along the line of sight to each candidate. The weighted richnesses are estimated within a default aperture of R_{500} centered at the candidate sky position (Klein et al., 2018, 2019). The weights include both a radial and a color component, with the radial filter following a projected Navarro, Frenk and White profile (NFW; Navarro et al., 1996, 1997), giving higher weights to galaxies closer to the center. The color filter uses the RS models and is tuned to give higher weights to cluster red sequence galaxies. These RS models are calibrated using over 2,500 clusters and groups with spectroscopic redshifts from the literature, including: the SPT-SZ cluster catalogue (Bleem et al., 2015), the redMaPPer Y1 catalogue (only for clusters with spectroscopic redshifts, McClintock et al., 2019), and the 2RXS X-ray sources cross-matched with the MCXC cluster catalogue (Piffaretti et al., 2011). These richnesses are estimated for each redshift bin with steps of $\Delta z = 0.005$. The richness as a function of redshift is then searched to find richness peaks; the three strongest λ peaks, each with a different photometric redshift, are recorded for each candidate.

The mean positional uncertainty of the *Planck* sources is ∼5.3 arcmin, which, adopting the cosmology from Section 3.1, translates into an uncertainty of ∼0.6 Mpc and ∼1.9 Mpc at $z = 0.1$ and $z = 0.5$, respectively. Given the large positional uncertainty of the *Planck* candidates, the SZE position of a cluster could in some cases be offset by several times R_{500}. These large positional uncertainties enhance the probability of a spurious *Planck* candidate being paired to a physically unassociated optical system. To address this large positional uncertainty, we run the MCMF algorithm twice. The first run adopts the positions from the *Planck* candidate catalogue, and carries out a search for possible optical counterparts within an aperture that is 3 times the positional uncertainty of the candidate, corresponding to a mean aperture of ∼15.9 arcmin.

This first run gives us up to three possible optical counterparts for each *Planck* candidate, with the corresponding photometric redshift, optical center and λ for each. For all potential counterparts, the RS galaxy density maps are used to identify the peak richness, which is adopted as the optical center. In the top row of Fig. 3.4 we show the richness distribution in redshift (estimated in this first run) of two different *Planck* candidates, at $z_{\rm MCMF} \approx 0.24$ (left) and $z_{\rm MCMF} \approx 0.88$ (right), with their corresponding pseudo-color images shown on the bottom row.

All potential counterparts identified in the first run are then used for a second MCMF run with the goal of identifying the most likely optical counterpart for each *Planck* candidate and refining the estimation of the photometric redshift and richness. We proceed with the second run of MCMF

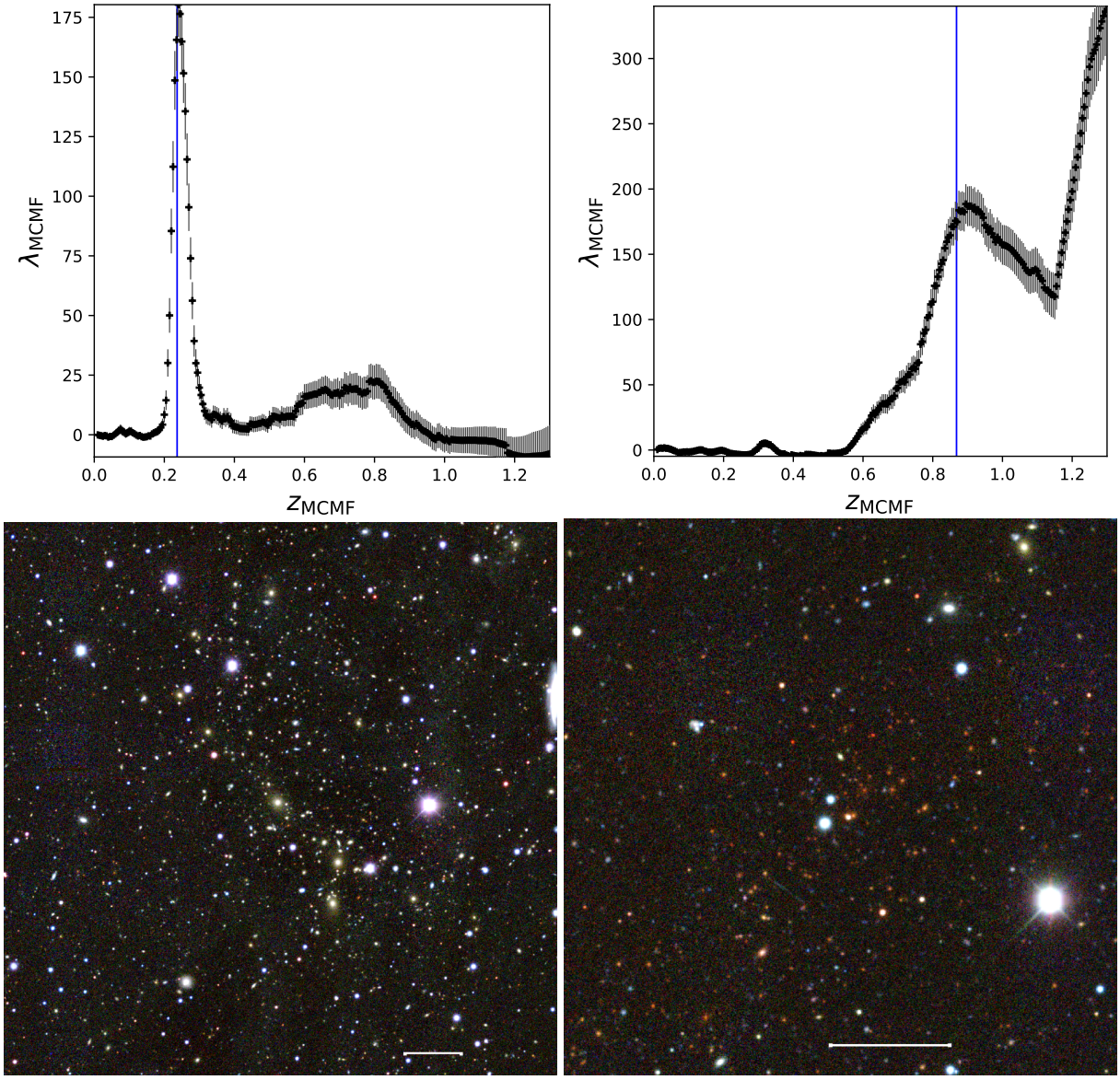

Figure 3.4: Example *Planck* cluster candidates with IDs PSZ-SN3 J2135+0124 ($z_{\mathrm{MCMF}} = 0.24$, left) and PSZ-SN3 J0102-4915 ($z_{\mathrm{MCMF}} = 0.87$, right). *Above:* Richness as a function of redshift for each candidate. The blue line marks the most likely redshift of the candidate. *Below:* DES pseudo-color images at the cluster positions. The white bar at the bottom denotes a scale of 1 arcmin. North is up and east is to the left.

using the optical counterpart positions as the input, but now using R_{500} as the aperture within which to search for counterparts. R_{500} is derived using a NFW profile and the *Planck* candidate mass estimation, M_{500}, at the redshift of each potential counterpart. For each candidate, redshift-dependant masses are estimated using the SZE mass proxy (for details see Section 7.2.2 of Planck Collaboration et al., 2014b). The *Planck* flux measured with the matched filter is degenerate with

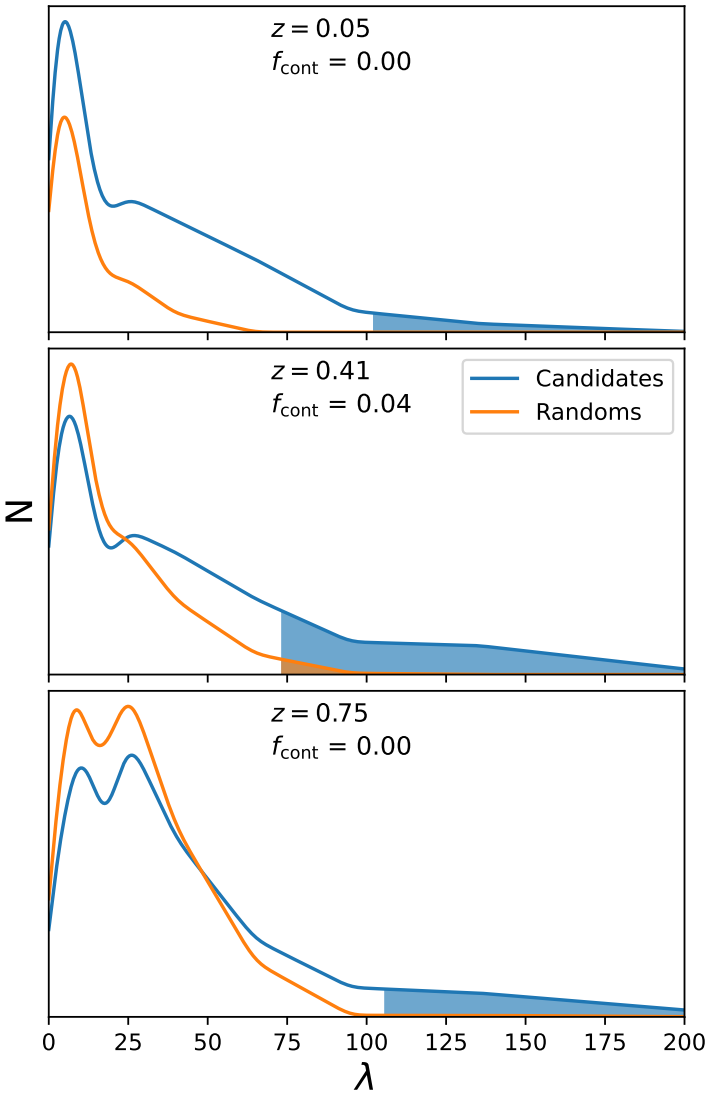

Figure 3.5: Examples of normalized richness distributions for random lines of sight (orange) and for *Planck* cluster candidates (blue) for all sources within an estimated $\delta z < 0.05$ of 3 *Planck* candidates shown from top to bottom at $z = 0.05, 0.41, 0.75$. For each of the sources, the area under the curves where the richness is equal to or greater than that of the *Planck* candidate is shaded. These shaded regions correspond to the numerator (orange) and denominator (blue) of equation (3.4).

the assumed size. We break this size-flux degeneracy using the flux-mass relation given by (see also equation 5 in Planck Collaboration et al., 2014b)

$$E^{-2/3}(z)\left[\frac{D_{\rm A}^2(z)Y_{500}}{10^{-4}{\rm Mpc}^2}\right] = 10^{-0.19}h_{70}^{-0.21}\left[\frac{M_{500}}{6\times10^{14}M_\odot}\right]^{1.79}, \tag{3.3}$$

where $E = H(z)/H_0$ and $H(z)$ is the Hubble parameter, and $D_{\rm A}$ is the angular diameter distance. This second run also gives us up to three redshift peaks for each source, but we select the richness

peak whose redshift is the closest to the output redshift from the first run.

In summary, we obtain the positions and the redshifts of up to three potential optical counterparts with the first MCMF run, and in the second run we obtain the final redshifts and richnesses of each of these optical counterparts. The information from the second run allows us to select the most probable counterpart in most cases, with some candidates having more than one probable counterpart, as discussed below.

3.3.2 Quantifying probability of random superpositions

As already noted, with MCMF we leverage the richness distributions along random lines of sight in the survey as a basis for assigning a probability that each potential optical counterpart of a *Planck* selected candidate is a random superposition (e.g., it is not physically associated with the *Planck* candidate). We describe this process below.

3.3.2.1 Richness distributions from random lines of sight

A catalogue along random lines of sight is generated from the original *Planck* catalogue, where for each candidate position we generate a random position on the sky, with a minimum radius of approximately 3 times the mean positional uncertainty (5.5 arcmin). We also impose the condition that the random position has to be at least $\sim 3\times5.5$ arcmin away from any of the *Planck* candidates. We analyze the catalogue of random positions using MCMF in the same manner as for the data, except that, for the NFW profile used in the second run, the mass information needed to estimate the R_{500} is randomly selected from any of the *Planck* candidates (removing the candidate from which the random was generated).

To have sufficient statistics we select two random positions for each *Planck* candidate, so we have approximately two times as many random lines of sight as *Planck* candidates. Given the large positional uncertainties in the *Planck* candidate catalogue, optical counterparts of random lines of sight might be assigned to an optical counterpart of a *Planck* candidate. To account for this, we remove from our random lines of sight catalogue those positions that 1) have $\lambda \geq 30$ (e.g., lines of sight with massive clusters), and 2) are within 3 Mpc of any *Planck* source from our final, confirmed catalogue and have $|z_{\rm Planck} - z_{\rm random}| < 0.1$. Also, once the second set of random lines of sight has been analysed, we remove those positions that lie within 3 arcmin from any random source position from the first set to avoid double counting the same optical structures.

3.3.2.2 Estimating the random superposition probability $f_{\rm cont}$

With the random lines of sight we can use the $f_{\rm cont}$ estimator presented in Klein et al. (2019), which is proportional to the probability of individual *Planck* candidates being random superpositions of physically unassociated structures (Klein et al., 2022). By imposing an $f_{\rm cont}$ threshold on our final cluster catalog, we are able to quantify (and therefore also control) the contamination fraction. To estimate $f_{\rm cont}$ for each *Planck* candidate, we integrate the normalized richness distributions along random lines of sight $f_{\rm rand}(\lambda, z)$, within multiple redshift bins, that have $\lambda \geq \lambda_{\rm src}$, where

$\lambda_{\rm src}$ is the richness of the *Planck* candidate. We do the same for the richness distribution of the *Planck* candidates $f_{\rm obs}(\lambda, z)$ and then we estimate $f_{\rm cont}$ as the ratio

$$f_{\rm cont}(\lambda_{\rm src}, z) = \frac{\int_{\lambda_{\rm src}}^{\infty} d\lambda\, f_{\rm rand}(\lambda, z)}{\int_{\lambda_{\rm src}}^{\infty} d\lambda\, f_{\rm obs}(\lambda, z)}. \tag{3.4}$$

In Fig. 3.5 we show three examples of *Planck* candidates with the estimated $f_{\rm cont}$. The blue and orange lines are the interpolated richness distributions of *Planck* candidates and of random lines of sight, respectively, at the redshift of the best optical counterpart. The orange (blue) shaded area shows the integral in the numerator (denominator) in equation (3.4), starting at the richness $\lambda_{\rm src}$ of the *Planck* candidate.

In simple terms, a constant value of $f_{\rm cont}$ can be translated to a redshift-varying richness value $\lambda(z)$. Thus, selecting candidates with a value of $f_{\rm cont}$ lower than some threshold, is similar to requiring the final cluster sample to have a minimum richness that can vary with redshift ($\lambda_{\rm min}(z)$), above which the catalogue has a fixed level of contamination. We refer to this threshold as $f_{\rm cont}^{\rm max}$, which yields a catalogue contamination estimated as $f_{\rm cont}^{\rm max} \times$ (initial contamination), independent of redshift. Because the initial contamination of the *Planck* selected sample is $f_{\rm SZE-cont}$ and the final contamination of the cluster sample selected to have $f_{\rm SZE-cont} < f_{\rm cont}^{\rm max}$ is $f_{\rm cont}^{\rm max} \times f_{\rm SZE-cont}$, one can think of the $f_{\rm cont}^{\rm max}$ selection threshold as the fraction of the contamination in the original candidate sample that ends up being included in the final confirmed cluster sample. Thus, through selecting an $f_{\rm cont}$ threshold one can control the level of contamination in the final confirmed cluster catalogue.

3.4 Results

In Section 3.4.1 we present PSZ-MCMF, the confirmed cluster catalogue extracted from the *Planck* candidate list after an analysis of the DES optical followup information using the `MCMF` algorithm. We then discuss in more detail the mass estimates (Section 3.4.2), the cross-comparison with other ICM selected cluster catalogues (Section 3.4.3) and the catalogue contamination and incompleteness (Section 3.4.4).

3.4.1 Creating the PSZ-MCMF cluster catalogue

As mentioned above, the `MCMF` algorithm allows us to identify up to three different richness peaks, corresponding to different possible optical counterparts, for each of the 3130 *Planck* candidates. To generate a final cluster catalogue, we select the most likely optical counterpart for each of the 3130 *Planck* candidates by choosing the counterpart that has the lowest probability $f_{\rm cont}$ of being a random superposition (i.e., of being a contaminant rather than a real cluster).

With `MCMF` we identify optical counterparts for 2,938 of the 3130 *Planck* candidates, whereas for the remaining 192 *Planck* candidates no counterpart is found (see Section 3.4.1.2 for details). Of the 2,938 candidates with optical counterparts, 2,913 have unique counterparts, while the remaining 25 share their counterpart with another candidate that is closer to that counterpart (see

Section 3.4.1.3 for details). Finally, we consider a candidate to be confirmed when its optical counterpart has $f_{\rm cont}$ below the threshold value $f_{\rm cont}^{\rm max} = 0.3$. This results in 1092 confirmed *Planck* clusters. Of these confirmed clusters, 120 have two prominent redshift peaks with $f_{\rm cont}$ below the threshold value $f_{\rm cont}^{\rm max}$, and are considered to be candidates with multiple optical counterparts.

The top panel of Fig. 3.6 shows the redshift distribution for different values of the threshold $f_{\rm cont}^{\rm max}$, while the bottom panel shows the richness as a function of the redshift for the best optical counterpart of the *Planck* candidates in this final catalogue. Small dots represent sources with an estimated $f_{\rm cont}\geq 0.3$, while bigger dots are color coded as green, black, blue or red according to whether $0.2\leq f_{\rm cont}<0.3$, $0.1\leq f_{\rm cont}< 0.2$, $0.05\leq f_{\rm cont}<0.1$ or $f_{\rm cont}<0.05$, respectively.

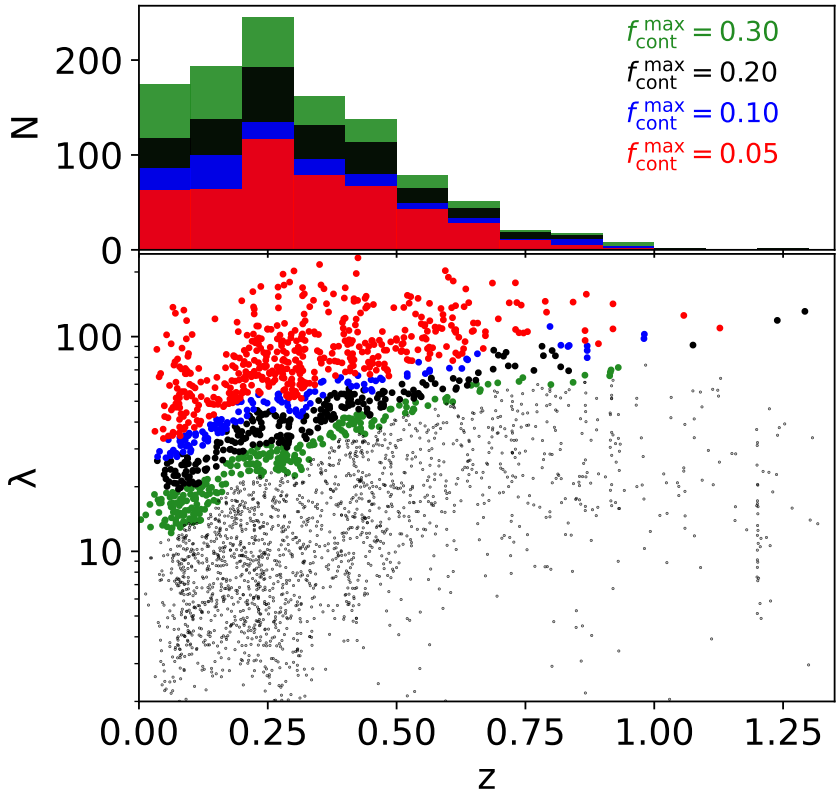

Figure 3.6: *top:* Redshift distribution of the 2913 *Planck* candidates. The green, black, blue and red histograms show the distributions of candidates with $f_{\rm cont}$ below $f_{\rm cont}^{\rm max} = 0.3, 0.2, 0.1$ and 0.05, respectively. *bottom:* Richness versus redshift for the best optical counterpart for each *Planck* candidate. Pairs with a probability of being random superpositions (contamination) $f_{\rm cont} > 0.3$ are shown as small black dots. Bigger green, black, blue and red dots represent counterparts with $0.2 \leq f_{\rm cont} < 0.3$, $0.1 \leq f_{\rm cont} < 0.2$, $0.05 \leq f_{\rm cont} < 0.1$ and $f_{\rm cont} < 0.05$, respectively, corresponding to subsamples with decreasing contamination.

In Table 3.1 we show the number of cluster candidates with $f_{\rm cont}$ below different values of the threshold $f_{\rm cont}^{\rm max}$, and different *Planck* candidate S/N thresholds. With this analysis we are adding 589 (828) clusters to the *Planck* cluster sample at $f_{\rm cont}^{\rm max} = 0.2$ (0.3) when going from the *Planck* S/N>4.5 to S/N>3.

Table 3.1: Number of confirmed *Planck* clusters with $f_{\rm cont}^{\rm max} = 0.3, 0.2, 0.1$ and 0.05 presented by row. Results are split by S/N. The second and third columns, for each S/N subsample, show the purity of the sample (Section 3.4.4.1) and the completeness (Section 3.4.4.2). The PSZ-MCMF sample presented in this paper corresponds to clusters with more restrictive MCMF cleaning in the case of the low signal to noise sample than in the higher signal to noise sample. These subsamples (see discussion in Section 3.4.1.4) are listed in bold face.

$f_{\rm cont}^{\rm max}$	S/N>3			S/N>4.5		
	$N_{\rm cl}$	Purity	Comp.	$N_{\rm cl}$	Purity	Comp.
0.3	1092	0.847	0.648	**264**	**0.974**	**0.990**
0.2	**842**	**0.898**	**0.530**	253	0.983	0.957
0.1	604	0.949	0.402	236	0.992	0.900
0.05	479	0.975	0.327	213	0.996	0.816

3.4.1.1 Candidates with a second optical counterpart

If the cluster candidate has two prominent redshift peaks with $f_{\rm cont}<f_{\rm cont}^{\rm max}$=0.3, where either (1) the redshift offset ($\delta z = (z_1 - z_2)/(1 + z_1)$) is greater than 2% or (2) the on-sky separation is greater than 10 arcmin, then we classify this candidate as a one with multiple optical systems, because a second optical counterpart with $f_{\rm cont}$<0.3 is an indication that the probability of being a chance superposition is lower than $f_{\rm cont}^{\rm max} \times f_{\rm SZE-cont}$. We give the redshifts, sky-positions, richnesses and other values for this second optical counterpart in the full cluster catalogue. In the case that both counterparts have the same $f_{\rm cont}$, we select the one that is closer to the *Planck* candidate position. In Appendix 3.6 we discuss a specific example.

3.4.1.2 Candidates with no optical counterpart

Out of the 3130 *Planck* candidates, there are 192 for which the MCMF analysis delivers no optical counterpart– not even with a high $f_{\rm cont}$. Most of these candidates (all but 26) are located near the edges of the DES footprint, suggesting that with more complete optical data many of these candidates could be associated with an optical counterpart. The 26 candidates that lie away from the DES survey edge show either a bright star or bright low-z galaxy near the *Planck* position or a lack of photometric information in one or more DES bands. Regions of the sky with these characteristics are masked by MCMF and this is the likely reason that no optical counterpart is identified for those candidates.

3.4.1.3 Candidates sharing the same optical counterpart

Given the rather generous search aperture used in the first run of MCMF, it is possible that some *Planck* candidates lying near one another on the sky share the same optical counterpart. There are 41 candidates, at $f_{\rm cont}$<0.3, that share 20 optical counterparts. The criteria we use to identify these 41 candidates is similar to the one used above to identify candidates with more than one possible optical counterpart. If the distance between the optical counterparts for the two *Planck* candidates is less than 10 arcmin and the redshift offset satisfies $|\delta z| \leq 0.02$, then we consider the

two candidates to be sharing the same optical counterpart. In Appendix 3.7 we discuss a specific example.

To account for such cases, we add a column to our catalogue that refers to which *Planck* candidate is the most likely SZE counterpart by using the distance between the SZE and the optical centers. The *Planck* candidate with the smallest projected distance from the optical center normalized by the positional uncertainty of the *Planck* candidate is considered to be the most likely SZE source.

3.4.1.4 Final PSZ-MCMF sample

With considerations of this last class we end up with 2913 *Planck* candidates, which are the closest to their respective optical counterparts. Table 3.1 contains the numbers of confirmed clusters, the purity (Section 3.4.4.1) and the completeness (Section 3.4.4.2) for different selection thresholds in $f_{\rm cont}$ and S/N. Given how the catalogue contamination of *Planck* candidates depends strongly on the S/N threshold (see Fig. 3.3), we decide to use two different values of $f_{\rm cont}^{\rm max}$ for the low S/N (S/N>3) and high S/N (S/N>4.5) samples. The low S/N sample will be defined as clusters with S/N>3 that meet the $f_{\rm cont}^{\rm max}$=0.2 threshold (second row of the S/N>3 sample in Table 3.1), whereas the high S/N sample will be defined as clusters with S/N>4.5 that meet the $f_{\rm cont}^{\rm max}$=0.3 threshold (first row of the S/N>4.5 sample in Table 3.1). The combination of these two samples corresponds to the PSZ-MCMF cluster sample, with a total of 853 clusters.

As previously noted in Section 3.3.2.2, the contamination fraction of the confirmed cluster sample is $f_{\rm cont}^{\rm max} \times f_{\rm SZE-cont}$ and depends on the $f_{\rm cont}$ selection threshold applied. The full PSZ-MCMF cluster catalogue will be made available online at the VizieR archive[4]. Table 3.2 contains a random subsample of the PSZ-MCMF catalogue with a subset of the columns.

In much of the discussion that follows we focus on the PSZ-MCMF cluster catalog; however, we will define two subsamples that will be used in specific cases: the low S/N sample and the high S/N sample. The low S/N sample ($f_{\rm cont}^{\rm max}$ = 0.2 and S/N>3), consists of 842 clusters with a ∼90% purity and 53% completeness. The high S/N sample ($f_{\rm cont}^{\rm max}$=0.3 and S/N>4.5) consists of 264 clusters with a ∼97% purity and 99% completeness. Other sample selections could be made, and the basic properties of twelve samples are presented in Table 3.1.

3.4.1.5 Comparison with spectroscopic redshifts

Starting with the ∼2,500 clusters and groups with spectroscopic redshifts used to calibrate the RS models of `MCMF`, we cross-match the cluster positions with the optical coordinates of each of our *Planck* candidates, selecting as matches those that lie within an angular distance of 3 arcmin. We choose to match with the optical counterpart positions, because they provide a more accurate sky position than the *Planck* SZE positions, which have a typical uncertainty of 5 arcmin. We use this cross-matched sample of clusters with spectroscopic redshifts to refine the red-sequence models of the `MCMF` algorithm (Klein et al., 2019).

We find 181 clusters in common with the PSZ-MCMF cluster catalogue, including a $z = 1.1$ cluster (SPT-CL J2106-5844). Of this sample, 18 clusters have another `MCMF` richness peak with

[4]http://vizier.u-strasbg.fr/

$f_{\rm cont}$ below the threshold value $f_{\rm cont}^{\rm max}$=0.2. Of these 18 candidates, the primary richness peak (lowest $f_{\rm cont}$) in 16 shows good agreement with the corrsponding spectroscopic redshift $z_{\rm spec}$, while for the remaining two the secondary peak lies at the $z_{\rm spec}$. Of the full cross-matched sample, there are two sources that have no secondary peak and exhibit a large redshift offset in the primary richness peak. We discuss these two cases in Appendix 3.8.1.

To characterise the redshift offset, we fit a Gaussian to the distribution of $\Delta z = (z_{\rm spec} - z_{\rm MCMF})$ / $(1+z_{\rm spec})$ of the 181 clusters, finding that the standard deviation is σ=0.00468 (indicating a typical `MCMF` redshift uncertainty of 0.47%), with a mean offset $\mu = -0.00005$ (indicating no `MCMF` redshift bias). This is consistent with the previously reported results from applications of the `MCMF` algorithm (Klein et al., 2019).

3.4.2 Estimating PSZ-MCMF cluster masses

Each *Planck* candidate comes with a function $M_{500}^{\rm Pl}(Y_{500}, z)$ that allows an initial mass estimate using the redshift and the SZE signal Y_{500} of the candidate (see equation 3.3). Therefore, for each of the 853 PSZ-MCMF clusters, we use the final photometric redshift from our `MCMF` analysis to estimate a mass.

It is important to note that candidates with multiple optical counterparts may have a biased SZE signature Y_{500} due to contributions from both physical systems, which would impact the estimated $M_{500}^{\rm Pl}$. However, because we do not have enough information to be able to separate the SZE emission coming for each component of the multiple counterparts, we adopt masses that are derived from the redshift of the first ranked richness peak. These masses are biased as discussed further below, and we therefore present a different mass esimate M_{500} in the final PSZ-MCMF catalogue (see the example Table 3.2).

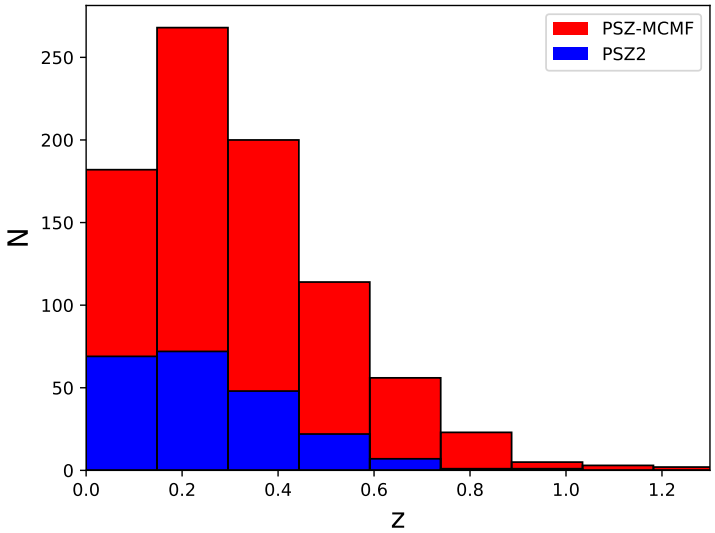

Figure 3.7: Redshift distribution of the PSZ-MCMF clusters (red) and the PSZ2 clusters within the DES region (blue). The new PSZ-MCMF catalogue presented here is significantly larger and extends to higher redshift.

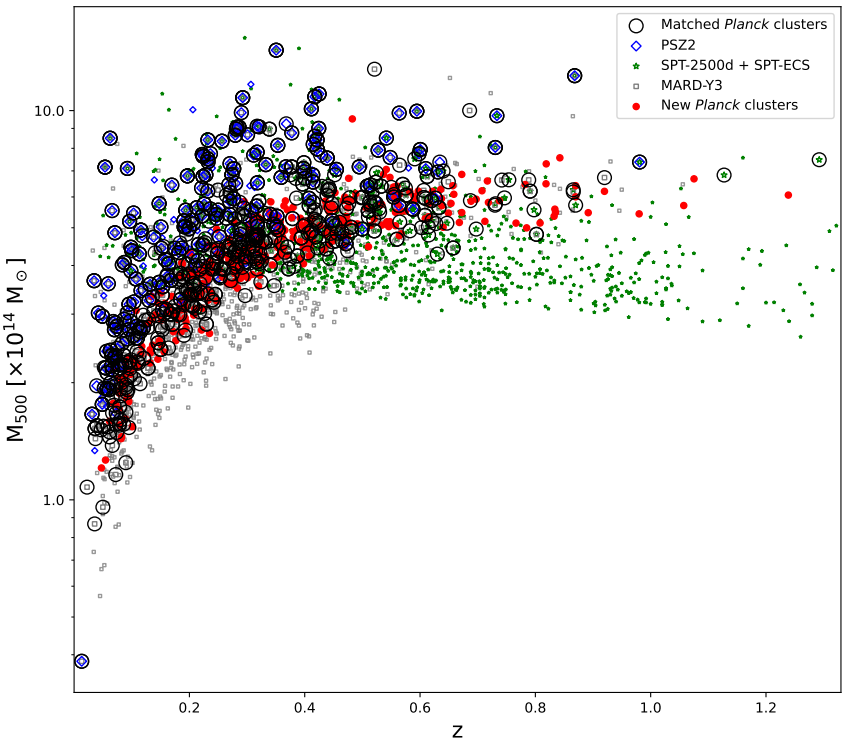

Figure 3.8: Mass versus redshift for the different cluster samples MARD-Y3, PSZ2, SPT and the PSZ-MCMF cluster catalog. SPT, PSZ2 and MARD-Y3 clusters are shown as green stars, blue diamonds or gray squares, respectively. New PSZ-MCMF clusters identified in this analysis (no match to PSZ2, SPT or MARD-Y3) are shown with red dots whereas clusters that match with at least one of the other catalogues appear as black circles. In the case of matches, masses and redshifts are those of our PSZ-MCMF catalog.

We expect a mass shift between the PSZ-MCMF cluster sample and both SPT and MARD-Y3, that is largely due to the hydrostatic mass bias that has not been accounted for in the *Planck* estimated masses (see, e.g., von der Linden et al., 2014; Hoekstra et al., 2015; Planck Collaboration et al., 2020a; Melin et al., 2021). In contrast, the SPT and MARD-Y3 masses are calibrated to weak lensing mass measurements (Bocquet et al., 2019), and should not be impacted by hydrostatic mass bias. We therefore apply a systematic bias correction to the *Planck* masses to bring all samples onto a common mass baseline represented by M_{500}.

To be able to compare our masses with different surveys accurately, we use cross-matched clusters and estimate the median mass ratio between the SPT/MARD-Y3 and the *Planck* mass estimates (see Section 3.4.3.2 for details), finding a median of $M_{500}^{\rm Pl}/M_{500} \approx 0.8$. This value is in agreement with both weak lensing (von der Linden et al., 2014; Hoekstra et al., 2015) and CMB lensing (Planck Collaboration et al., 2020a) analyses of *Planck* clusters. Therefore, we correct the masses of the PSZ-MCMF clusters identified in our current analysis by this factor. Because the previously published PSZ2 catalogue has masses that are calculated in a manner similar to the $M_{500}^{\rm Pl}$ described above, we correct PSZ2 masses also using a correction of $(1-b) = 0.8$. However, we note a further shift of $M_{500}^{\rm Pl}/M_{500}^{\rm PSZ2} \approx 0.95$ with respect to our corrected masses, and so we further correct the PSZ2 masses for the final comparison.

It should be noted that the mass bias of *Planck* clusters is still an ongoing topic. In summary, the masses we present in the following sections and the final cluster catalogue Table 3.2 are denoted as M_{500} and are rescaled to be consistent with results from a range of weak lensing calibration analyses. These masses are larger than the *Planck* masses $M_{500}^{\rm Pl}$ by a factor $1/0.8 = 1.25$.

3.4.3 Comparison to other ICM selected cluster catalogues

To check how the PSZ-MCMF cluster sample compares to others, we select three cluster catalogues that have been selected using ICM signatures and that lie within the DES footprint: MARD-Y3 (Klein et al., 2019), SPT-2500d (Bocquet et al., 2019) along with SPT-ECS (Bleem et al., 2020) and PSZ2. MARD-Y3 is an X–ray selected cluster catalogue confirmed with DES Y3 photometric data, using the same tools as for the *Planck* analysis presented here. This MARD-Y3 catalogue has 2,900 clusters with $f_{\rm cont} < 0.2$. On the other hand, both the SPT and PSZ2 cluster catalogues are based on SZE selection. For SPT we select sources with a redshift measurement (photometric or spectroscopic), giving a total of 964 clusters. It is worth noting that PSZ2 is an all sky survey, and for the comparison we select sources that lie within the DES survey region and have a redshift measurement (226 clusters).

3.4.3.1 Comparison to PSZ2 catalogue

We compare the estimated redshifts of our 2,938 candidates with optical counterparts (no $f_{\rm cont}^{\rm max}$ applied) with those from the PSZ2 catalogue (Planck Collaboration et al., 2016a), because the two catalogues should contain a similar number of clusters at S/N>4.5, with small variations expected due to the different algorithms used to detect clusters. There are 1,094 PSZ2 clusters with a measured redshift, and, out of those, 226 lie within the DES footprint. We match these 226 clusters with sources from our catalogue that have good photometric redshift estimations and S/N≥4.5, using a matching radius of 3 arcmin. In this case we do the matching using both the *Planck* SZE position and the optical positions.

We find 217 matching sources, but one of those matches does not correspond to the closest cluster in our catalogue so we exclude it and use the 216 remaining sources. Of the 9 PSZ2 sources for which we find no match, 7 have missing photometric information in one or more DES bands. The remaining two clusters with IDs PSZ2 G074.08-54.68 and PSZ2 G280.76-52.30, are further discussed in Appendix 3.9.

Of this matched sample of 216 systems, 207 (214) systems have $f_{\rm cont} < 0.2$ (0.3) and redshifts that are in good agreement with ours. The cases of disagreement are discussed in detail in Appendix 3.8.2. By comparing the 214 matching clusters with $f_{\rm cont} < 0.3$ to the numbers shown on the Table 3.1 (264 at S/N>4.5), it becomes apparent that the analysis we describe here has led to photometric redshifts and optical counterparts for 50 PSZ2 clusters that previously had no redshift information. Fig. 3.7 shows the redshift distribution of our cluster catalogue (red histogram) and of the PSZ2 catalogue within DES (blue histogram).

3.4.3.2 PSZ-MCMF mass-redshift distribution

We compare the mass-redshift distribution of PSZ-MCMF clusters with that of MARD-Y3, SPT and PSZ2. Our first step in cross-matching is to select clusters that are the closest to their respective optical counterpart (853 clusters). Then the cross-match comparison is done by using both a positional match within 3 arcmin from the *Planck* positions or from the optical positions. We also add a redshift constraint, where only candidates with a redshift offset $\delta z < 0.02$ (using only the first peak) are considered. This gives a total of 500, 187 and 233 matches with MARD-Y3, PSZ2 and SPT (2500d + ECS), respectively. In total, then, 329 PSZ-MCMF clusters are not matched to any of the three published catalogues.

In Fig. 3.8 we show the mass versus redshift distribution for the different cluster samples. The SPT, PSZ2 and MARD-Y3 samples are shown as green stars, blue diamonds or gray squares, respectively. PSZ-MCMF clusters are shown with red dots if they are unmatched to clusters in SPT, PSZ2 or MARD-Y3 and as black circles if they are matched. The red systems are the previously unknown SZE selected clusters in the DES region. In the case of matches to previously published samples, we adopt the mass and redshift estimates from the PSZ-MCMF sample to ensure the points lie on top of one another. Fig. 3.8 contains more than 10 massive clusters ($M_{500} \gtrsim 10^{15}$ $M_\odot$ and $z < 0.5$) with no matches to the PSZ-MCMF cluster sample. Visual inspection shows that those systems were slightly outside the DES footprint or within masked regions within the general DES footprint.

For MARD-Y3, we clean the unmatched sources by selecting those without multiple X–ray sources to avoid double counting clusters, and also exclude clusters with strong AGN contamination as indicated by their AGN exclusion filter (see section 4.2.1 in Klein et al., 2019). Also, following their mass versus redshift distribution, we use a threshold of $f_{\rm cont} < 0.05$ and also remove sources with a second counterpart with $f_{\rm cont} < 0.05$.

The mass-redshift distribution of our *Planck* sample is similar to that of the MARD-Y3 X–ray selected sample, which finds more lower mass systems at lower redshifts. In contrast, the SPT sample mass-redshift distribution exhibits only a slight redshift trend (Bleem et al., 2015), but it lacks the lower mass systems seen at low redshift in the *Planck* and MARD-Y3 samples. For the *Planck* selection, it is the multi-frequency mapping that enables the separation of the thermal SZE from the contaminating CMB primary temperature anisotropy, and this enables the detection of low redshift and low mass systems in a way that resembles the flux limited selection in the MARD-Y3 catalogue. SPT, on the other hand, has coverage over a narrow range of frequency and cannot as effectively separate the thermal SZE and the primary CMB anisotropies. The SPT cluster extraction is therefore restricted to a smaller range of angular scales, which is well matched to cluster virial regions at $z \gtrsim 0.3$, but at lower redshifts an ever smaller fraction of the SZE signature is obtained, making it ineffective at detecting the low mass and low redshift systems seen in the *Planck* and MARD-Y3 samples. At $z \lesssim 0.6$, MARD-Y3 selects lower mass clusters than we are able to with our *Planck* sample, but at higher redshifts both catalogues follow similar distributions. When comparing with PSZ2, our new *Planck* catalogue contains lower mass clusters at all redshifts, which is expected given that we are pushing to lower S/N with our *Planck* catalogue. Our *Planck* sample also contains the first $z > 1$ *Planck* selected clusters.

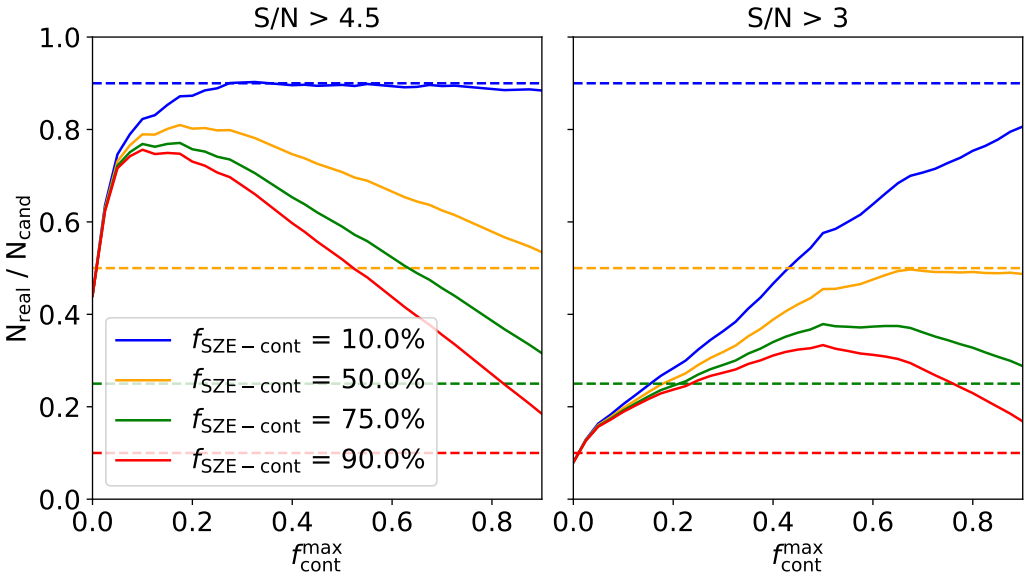

Figure 3.9: Ratio of the estimated number of real clusters $N_{\rm real}$ to the total number of candidate clusters $N_{\rm cand}$ in the *Planck* sample as a function of the $f_{\rm cont}$ threshold value applied. The solid lines show different curves from equation (3.5) with four different values of the contamination $f_{\rm SZE-cont}$ of the initial *Planck* candidate list. The dashed lines show $1 - f_{\rm SZE-cont}$, with colors encoding different initial contamination levels. The analysis indicates an initial contamination of 10% in the S/N>4.5 (left) and 50% in the S/N>3 (right) *Planck* candidate samples.

3.4.4 PSZ-MCMF contamination and incompleteness

An application of the *Planck* based cluster finding algorithm to mock data suggests that at S/N>3 we should expect about 75% of the candidates to be contamination (noise fluctuations; see Section 3.2.2). In this section we explore that expectation using information from the `MCMF` followup. Moreover, as one subjects the confirmed PSZ-MCMF sample to more restrictive $f_{\rm cont}$ selection thresholds (i.e., smaller values), one is removing not only chance superpositions (contaminants) from the sample, but also some real clusters. In the following subsections we also explore the incompleteness introduced by the $f_{\rm cont}$ selection.

3.4.4.1 Estimating contamination

With the `MCMF` analysis results in hand, we can now estimate the true contamination fraction of the initial candidate list by analysing the number of real cluster candidates $N_{\rm real}$ from the number of selected clusters $N_{\rm cl}$ as a function of the $f_{\rm cont}$ threshold $f_{\rm cont}^{\rm max}$ and input *Planck* candidate catalogue contamination $f_{\rm SZE-cont}$. The number of real clusters is estimated as

$$N_{\rm real}(f_{\rm cont}^{\rm max}) = N_{\rm cl}(f_{\rm cont} < f_{\rm cont}^{\rm max}) \left[1 - f_{\rm cont}^{\rm max} f_{\rm SZE-cont}\right] \tag{3.5}$$

where $N_{\rm cl}(f_{\rm cont} < f_{\rm cont}^{\rm max})$ is the total number of confirmed *Planck* candidates with $f_{\rm cont} < f_{\rm cont}^{\rm max}$ and $\left[1 - f_{\rm cont}^{\rm max} f_{\rm SZE-cont}\right]$ represents the fraction of real clusters in a sample of `MCMF` confirmed clusters. As discussed in Section 3.3.2.2, $f_{\rm cont}$ is defined in a cumulative manner and the final

contamination of an $f_{\rm cont} < f_{\rm cont}^{\rm max}$ selected sample is the product $f_{\rm cont}^{\rm max} f_{\rm SZE-cont}$ where $f_{\rm SZE-cont}$ is the contamination fraction of the original *Planck* candidate list, and $f_{\rm cont}^{\rm max}$ is the fraction of this contamination that makes it into the final confirmed cluster sample.

In this way, we can estimate $N_{\rm real}$ for a number of values of $f_{\rm cont}^{\rm max}$ and $f_{\rm SZE-cont}$. Under the assumption that the $f_{\rm cont}$ selection restricts contamination as expected, we can then solve for the input candidate list contamination $f_{\rm SZE-cont}$, which again was estimated through *Planck* sky simulations to be ∼0.75. The catalogue contamination should give a constant ratio of $N_{\rm real}/N_{\rm cand} = 1 - f_{\rm SZE-cont}$ at higher $f_{\rm cont}^{\rm max}$ where this $f_{\rm cont}$ selection becomes unimportant.

It is instructive to start with a less contaminated sample similar to PSZ2 by taking into account only *Planck* candidates with S/N>4.5 (284 candidates). In Fig. 3.9 we plot the ratio of the number of estimated real clusters $N_{\rm real}$ to the total number of *Planck* candidates as a function of the $f_{\rm cont}$ threshold value $f_{\rm cont}^{\rm max}$ used to select the sample. Each solid curve represents the estimated number of real clusters $N_{\rm real}$, color coded according to the assumed *Planck* candidate sample contamination $f_{\rm SZE-cont}$. The horizontal dashed lines show $1 - f_{\rm SZE-cont}$, which is showing the fraction of *Planck* candidates that are expected to be real clusters and therefore could be confirmed using `MCMF` . We would expect that for threshold values $f_{\rm cont}^{\rm max}$ approaching 1, where the `MCMF` selection is having no impact, that the fraction plotted in the figure would reach the value $1 - f_{\rm SZE-cont}$.

The input contamination that best describes the high S/N sample is $f_{\rm SZE-cont} = 8.5\%$, where at $f_{\rm cont} < 0.3$ the fraction of confirmed candidates has reached the maximum possible within the *Planck* candidate list. A further relaxing of the $f_{\rm cont}$ threshold has essentially no impact on the number of real clusters $N_{\rm real}$; it just adds contaminants to the list of `MCMF` confirmed clusters $N_{\rm cl}$ at just the rate that matches the expected increase in contamination described in equation (3.5). This contamination is in line with the ∼ 91% reliability estimated for the PSZ2 cluster cosmology sample (see Fig. 11 in Planck Collaboration et al., 2016a).

Note the behavior of the blue line at $f_{\rm cont}$ values < 0.3. The confirmed ratio falls away from 90%, indicating the onset of significant incompleteness in the `MCMF` selected sample. This is an indication that as one uses $f_{\rm cont}$ to produce cluster samples with lower and lower contamination fractions, one is also losing real systems and thereby increasing incompleteness. We discuss this further in the next subsection (Section 3.4.4.2).

For the more contaminated S/N>3 *Planck* candidate sample (2913 candidates) the results are shown in the bottom panel of Fig. 3.9. When the $f_{\rm cont}$ threshold is 0.2, the estimated number of real clusters $N_{\rm real}$ is roughly 25% of the total number of *Planck* candidates, which implies a 75% contamination. However, unlike the S/N>4.5, the curve does not flatten until $f_{\rm cont} \geq 0.65$, and only for initial contamination values $f_{\rm SZE-cont} = 50\%$. This later flattening reflects the low mass range (and therefore lower richness range) of the S/N>3 candidate list. Additionally, our analysis indicates that the initial contamination of the *Planck* S/N>3 candidate list is ≈51% rather than the estimated 75% from *Planck* mock sky experiments. We explore these differences further in Appendix 3.10.

Finally, using this 51% initial contamination (yellow lines), we expect to lose 286 clusters when going from an $f_{\rm cont}$ threshold of <0.2 (∼90% purity) to <0.05 (∼97.5% purity). Indeed, any $f_{\rm cont}$ threshold below 0.6 will remove real *Planck* selected clusters from the `MCMF` confirmed sample, but including these systems comes at the cost of higher contamination (purity drops

to ∼70%). The purity for different thresholds of $f_{\rm cont}^{\rm max}$ is listed in Table 3.1 for the two *Planck* candidate S/N ranges.

Given how the PSZ-MCMF cluster catalog is constructed (the combination of the low and high S/N subsamples), the final purity is estimated to be ∼90%.

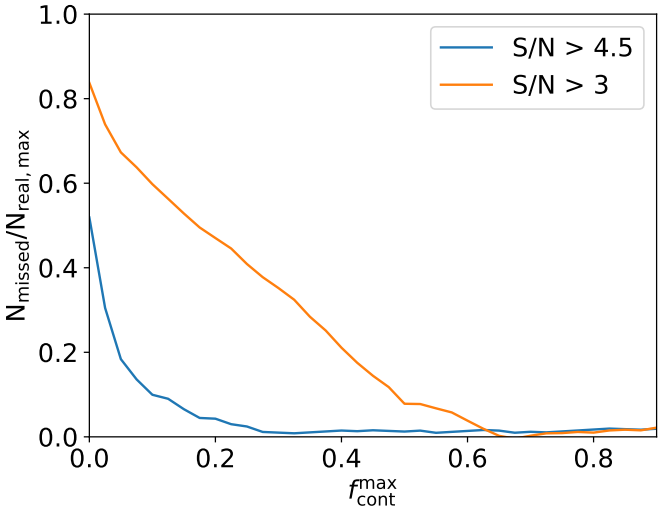

Figure 3.10: Estimate of the fractional incompleteness versus the $f_{\rm cont}$ threshold $f_{\rm cont}^{\rm max}$ for the S/N>4.5 ($f_{\rm SZE-cont}$ = 8.5%) and S/N>3 ($f_{\rm SZE-cont}$ = 51%) confirmed cluster samples, shown with blue and orange lines, respectively. The contamination of the resulting cluster catalogue is given by $f_{\rm SZE-cont} f_{\rm cont}^{\rm max}$ (see Section 3.4.4.1 and Table 3.1).

3.4.4.2 Estimating incompleteness

From this analysis, we can estimate the number of missed clusters $N_{\rm missed}$ or equivalently the fractional incompleteness for a given $f_{\rm cont}$ threshold. First, we estimate the maximum number of real clusters in the full sample as $N_{\rm real,max} = N_{\rm cand}(1 - f_{\rm SZE-cont}) = 1,427$ (for the S/N>3 sample), where in this case $N_{\rm cand}$ is the full *Planck* candidate list. Then, we estimate the number of missed clusters using the total number of expected real clusters minus the number of real SZE selected clusters at a particular $f_{\rm cont}$ threshold value:

$$N_{\rm missed}(f_{\rm cont} < f_{\rm cont}^{\rm max}) = N_{\rm real,max} - N_{\rm real}(f_{\rm cont}^{\rm max}) \tag{3.6}$$

where $N_{\rm real}(f_{\rm cont}^{\rm max})$ is defined as in equation (3.5). In Fig. 3.10 we show the ratio of missed clusters over the expected maximum number of real clusters for the samples at S/N>3 (orange line) and S/N>4.5 (blue line). An $f_{\rm cont}$ threshold of 0.2 in the S/N>4.5 *Planck* sample would be missing slightly over 3% of the real clusters, while at S/N>3 and the same threshold 0.2, we expect to miss ∼ 47% of the real clusters. With an $f_{\rm cont}$ threshold of 0.05 we miss ∼ 70% of the real clusters. The completeness for different selection thresholds $f_{\rm cont}^{\rm max}$ is shown in Table 3.1, for the two *Planck* candidate S/N ranges. We estimate the completeness of the PSZ-MCMF cluster catalog to be ∼54%.

The higher incompleteness for the lower S/N sample is expected, because as discussed in Section 3.4.2 this sample pushes to lower masses and therefore lower richnesses than the S/N>4.5 sample. At lower richness, real clusters cannot be as effectively differentiated from the typical background richness distribution (see random line of sight discussion in Section 3.3.2). In this low mass regime, along with the large positional uncertainties, the cost of creating a higher purity *Planck* sample is the introduction of high incompleteness.

3.5 Summary & Conclusions

In this analysis we create the PSZ-MCMF cluster catalogue by applying the `MCMF` cluster confirmation algorithm to DES photometric data and an SZE selected cluster candidate list extracted down to S/N=3 from *Planck* sky maps. In contrast to previous analyses employing the `MCMF` algorithm, the low angular resolution of *Planck* together with the low S/N threshold result in much larger positional uncertainties of the SZE selected candidates. To overcome this challenge we apply the `MCMF` algorithm twice, first using the *Planck* candidate coordinates to define a search region with an aperture that is 3 times the *Planck* candidate positional uncertainty, and then second using the positions of the optical counterparts found in the first run, with an aperture based on an estimate of the halo radius $R_{500}(z)$ that employs the mass constraints from the *Planck* dataset.

We control the contamination of the final, confirmed sample by measuring the parameter $f_{\rm cont}$ for each *Planck* candidate. As discussed in Section 3.3.2.2, the value of this parameter is proportional to the probability that the *Planck* candidate and its optical counterpart are a chance superposition of physically unassociated systems rather than a real cluster of galaxies. About 10% of the *Planck* candidates exhibit multiple potential optical counterparts. In such cases we select the most likely optical counterpart by choosing the one with the lowest $f_{\rm cont}$ value (lowest chance of being contamination).

Our analysis of the PSZ-MCMF sample indicates that the initial contamination fraction of the *Planck* S/N>4.5 candidate list is $f_{\rm SZE-cont}$ ~9% and the S/N>3 candidate list is $f_{\rm SZE-cont}$ ~50%. The optical followup with `MCMF` allows us to reduce this contamination substantially to the product $f_{\rm cont}^{\rm max} \times f_{\rm SZE-cont}$, where $f_{\rm cont}^{\rm max}$ is the maximum allowed $f_{\rm cont}$ value in a particular subsample.

Table 3.2 contains the full PSZ-MCMF sample of 853 confirmed clusters, defined using an $f_{\rm cont}$ threshold of 0.3 for S/N>4.5 candidates and an $f_{\rm cont}$ threshold of 0.2 for S/N>3 candidates. Table 3.1 contains the number of clusters, the purity and the completeness of this cl catalogue (in bold face) together with other subsamples constructed using smaller $f_{\rm cont}$ thresholds of 0.2, 0.1 and 0.05 for both *Planck* S/N ranges. Whereas the full catalogue contains 853 clusters with a purity of 90% and completeness of 54%, the subsample with $f_{\rm cont}$ <0.2 (<0.1) contains 842 (604) clusters with purity and completeness of 90% (95%) and 53% (40%), respectively.

Furthermore, the cl cluster sample at S/N>3 excludes 47% of the real clusters when applying a limiting value at $f_{\rm cont} < 0.2$, while the same threshold on the S/N>4.5 sample excludes around 4%. We attribute the higher incompleteness of the confirmed low S/N sample to the fact that these systems have lower masses and richnesses. The lower richnesses for the real clusters in this regime are simply more difficult to separate from the characteristic richness variations along

random lines of sight in the DES survey. The relatively large positional uncertainties of the *Planck* candidates makes this effect even stronger.

Users are encouraged to select subsamples of the cl sample with lower contamination, depending on their particular scientific application. The PSZ-MCMF catalogue adds 828 previously unknown *Planck* identified clusters at S/N>3, and it delivers redshifts for 50 previously published S/N>4.5 *Planck* clusters.

For each of the confirmed clusters we derive photometric redshifts. By comparing the PSZ-MCMF cluster sample with spectroscopic redshifts from the literature, we find a mean redshift offset $< 10^{-4}$ and an RMS scatter of 0.47%. With these redshifts together with the *Planck* mass constraints, we estimate halo masses for all confirmed clusters. These original *Planck* based mass estimates contain no correction for hydrostatic mass bias, and so these are rescaled by the factor $1/0.8 = 1.25$ to make them consistent with the weak lensing derived SPT cluster masses (Bocquet et al., 2019). Optical positions, redshifts and halo masses M_{500} are provided for each confirmed cluster in Table 3.2.

We crossmatch the PSZ-MCMF cluster catalogue to different SZE and X-ray selected cluster catalogues within the DES footprint. We find that the PSZ-MCMF mass distribution with redshift is similar to that of the X–ray selected MARD-Y3 cluster catalogue. However, at redshifts lower than $z < 0.5$ the PSZ-MCMF catalogue does not contain the lower mass systems that the X-ray selected MARD-Y3 catalogue contains. When comparing with the previous *Planck* SZE source catalogue PSZ2, we have optical counterparts for most of the systems that lie within the DES footprint, finding in general good agreement with their previously reported redshifts. Compared to the higher S/N PSZ2 sample, we find that most of our new lower S/N PSZ-MCMF systems lie at lower masses at all redshifts and extend to higher redshift, as expected. Probing to lower masses allows for the confirmation of the first $z > 1$ *Planck* identified galaxy clusters. Crudely scaling these results to the full extragalactic sky (~ 30000 deg^2) implies that the *Planck* full sky candidate list confirmed using `MCMF` applied to DES like multi-band optical data would yield a sample of ~6000 clusters, which is ~ 6 times the number of clusters in the PSZ2 all-sky cluster catalogue with redshift information.

3.6 Appendix: Multiple optical counterparts

In Fig. 3.11 we show an example of the *Planck* candidate PSZ-SN3 J0605-3519, which is classified as a candidate with multiple optical counterparts. The upper figure shows the richness as a function of redshift, which shows two prominent peaks at $z_{\rm MCMF} = 0.15$ and $z_{\rm MCMF} = 0.52$. The lower image contains the pseudo-color image from *gri* DES cutouts. White and red contours are derived from the RS galaxy density map for galaxies at $z_{\rm MCMF} = 0.15$ and $z_{\rm MCMF} = 0.52$, respectively. The richness for these two counterparts are $\lambda_{\rm MCMF} = 84$ and $\lambda_{\rm MCMF} = 156$ for the white and red contours, for the two optical candidates at $z_{\rm MCMF} = 0.15$ and $z_{\rm MCMF} = 0.52$ respectively. For this candidate, the estimated $f_{\rm cont}$ of both redshift peaks is 0, indicating a vanishing small probability that either one is a random superposition. We choose the one at $z_{\rm MCMF} = 0.15$ as the "preferred" counterpart because it lies nearer to the *Planck* candidate position. The reported spectroscopic redshift for this cluster comes from the REFLEX cluster

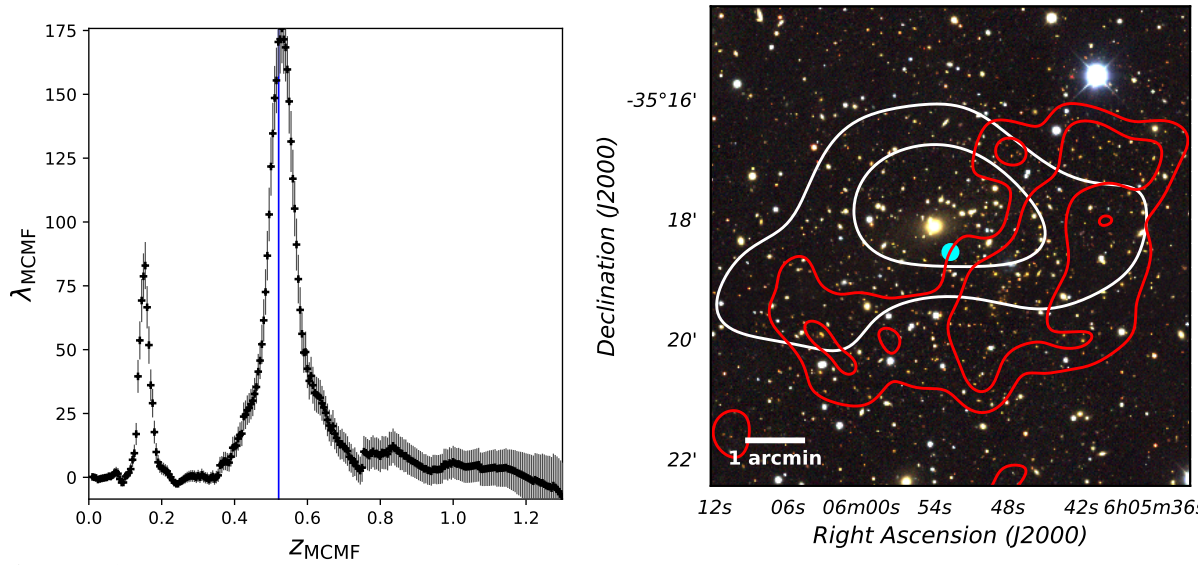

Figure 3.11: *left*: Richness as a function of redshift for the *Planck* candidate PSZ-SN3 J0605-3519. Two richness peaks at $z_{\rm MCMF} = 0.15$ and $z_{\rm MCMF} = 0.52$ are evident. *right*: pseudo-color image from DES *g*, *r*, *i* cutouts around the *Planck* candidate coordinates, marked by a cyan dot. Contours are from the galaxy density maps of the counterpart at $z_{\rm MCMF} = 0.15$ (white) and at $z_{\rm MCMF} = 0.52$ (red).

catalogue, with $z_{\rm spec} = 0.1392$ for cluster RXCJ0605.8-3518 (Böhringer et al., 2004).

3.7 Appendix: Shared optical counterpart

In Fig. 3.12 we show an example of two *Planck* candidates (PSZ-SN3 J2248-4430 with $f_{\rm cont} = 0.00$ and PSZ-SN3 J2248-4436 with $f_{\rm cont} = 0.18$) sharing the same optical counterpart, where the *Planck* positions are marked with dots. The optical center of the preferred counterpart for each candidate is marked with a cross of the same color. White contours are the RS galaxy density map from the first `MCMF` run, where the optical centers are determined. Both redshifts point toward a cluster at $z_{\rm MCMF} = 0.35$, but it is pretty clear that the two *Planck* candidates have resolved to the same optical counterpart. Interestingly, this optical system also corresponds to a South Pole Telescope (SPT) cluster, namely SPT-CL J2248-4431, with a spectroscopic redshift of $z_{\rm spec} = 0.351$ (Bocquet et al., 2019).

To resolve such cases, we select the *Planck* candidate with the smallest projected distance from the optical center normalized by the positional uncertainty of the *Planck* candidate. We add a column to the catalogue that identifies which *Planck* candidate is the most likely SZE counterpart, $\rm flag_{closest}$, with a value of 0 for candidates pointing to a unique optical counterpart and 1 for candidates which share the optical counterpart with another candidate but are selected as the most likely SZE counterpart. We visually inspected each of the 41 ($f_{\rm cont} < 0.3$) cases, looking not only at the separation, but also at the S/N of the candidates, and the estimated $f_{\rm cont}$ and λ. The method described above correctly identifies the most likely candidate for a counterpart in 18 out of 20 cases for candidates at $f_{\rm cont} < 0.3$. For the remaining two, we manually select the

most likely SZE source. The final PSZ-MCMF cluster catalogue contains 853 clusters, which are the most likely SZE counterparts of their respective optical counterpart.

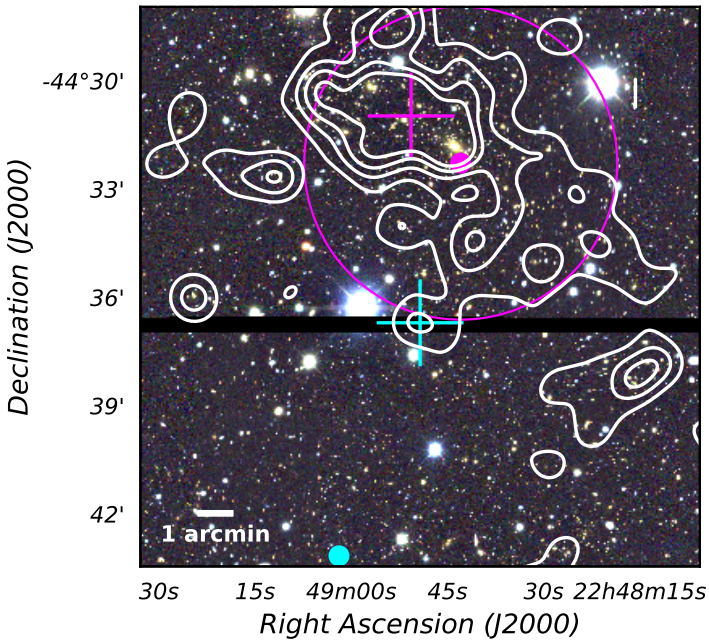

Figure 3.12: Pseudo-color image from DES *g*, *r*, *i* cutouts around the coordinates of *Planck* candidates PSZ-SN3 J2248-4430 and PSZ-SN3 J2248-4436, marked by magenta and cyan dots, respectively. White contours are from the galaxy density maps of the counterpart at $z_{\mathrm{MCMF}} = 0.35$. Crosses mark the position of the optical counterparts associated with each of the *Planck* sources, color coded according to the *Planck* source.

3.8 Appendix: Redshift comparisons

3.8.1 Spectroscopic redshifts

As discussed in Section 3.4.1.5, the full cross-matched sample contains two sources that have no significant secondary peak and exhibit a large redshift offset with respect to $z_{\mathtt{MCMF}}$. We inspect the DES images of these two clusters, namely PSZ-SN3 J2145-0142 ($z_{\mathtt{MCMF}} = 0.36$ and $f_{\mathrm{cont}} = 0.0$) and PSZ-SN3 J2347-0009 ($z_{\mathtt{MCMF}} = 0.26$ and $f_{\mathrm{cont}} = 0.02$), where the separation between the spectroscopic and optical counterparts are ∼150 and ∼180 arcseconds, respectively, and find that in both cases the spectroscopic redshift points towards a different structure. In the case of PSZ-SN3 J2145-0142, the spectroscopic redshift seems to be associated with a single galaxy. Fig. 3.13 shows the richness as a function of redshift for both PSZ-SN3 J2145-0142 (left) and PSZ-SN3 J2347-0009 (right), with the spec-z marked with blue dotted lines. In the case of PSZ-SN3 J2347-0009, the measured f_{cont} for the structure at $z \approx 0.53$ is greater than our $f_{\mathrm{cont}}^{\mathrm{max}} = 0.3$ threshold, indicating that this is not a significant richness peak.

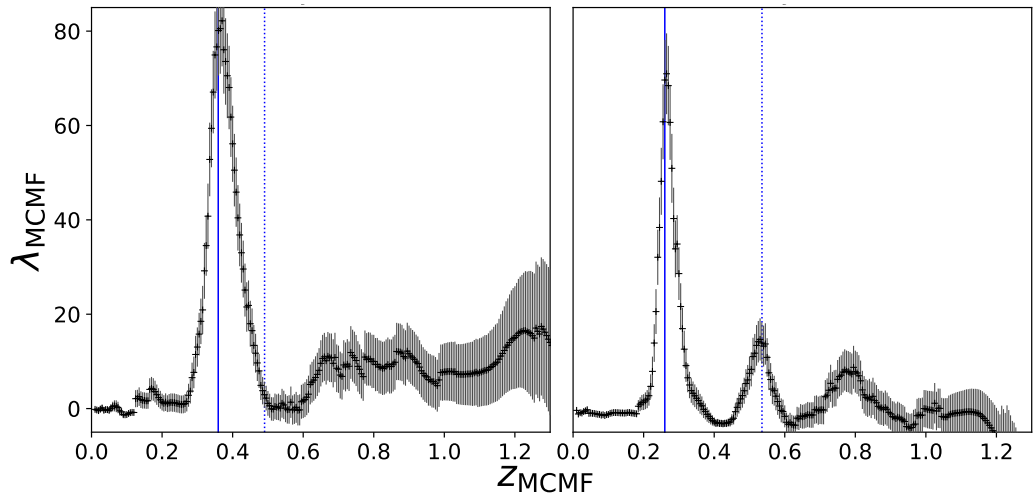

Figure 3.13: Richness as a function of redshift for PSZ-SN3 J2145-0142 (left) and PSZ-SN3 J2347-0009 (right). The PSZ-MCMF redshift is shown with a blue continuous line, while the spec-z is shown with a blue dotted line for both cases.

3.8.2 PSZ2 redshifts

In Fig. 3.14 we show the comparison of PSZ2 redshifts to the `MCMF` for the 216 matching systems. On the x-axis, we show the photometric redshift from `MCMF`, while redshifts from the PSZ2 catalogue are shown on the y-axis. Each source is color-coded according to their $f_{\rm cont}$ estimation. Continuous (dotted) lines show the enclosed area where $\delta z = |(z_{\rm MCMF} - z_{\rm PSZ2}) \times (1 + z_{\rm PSZ2})^{-1}| \leq 0.02$ (0.05). In case of multiple prominent redshift peaks with $f_{\rm cont} < 0.2$, we choose to plot only the redshift peak with the smaller δz for each match.

Fig. 3.14 shows that, although most of the estimated `MCMF` redshifts have Δz offsets at 2% level or less in comparison to the PSZ2 catalogue, there are some clusters with a higher offset or with $f_{\rm cont} \geq 0.2$. Out of the 216 matching clusters, 207 have $f_{\rm cont} < 0.2$, and 197 (205) have a redshift offset, with respect to the first redshift peak, lower than 2% (5%). If we consider also structures with a second peak, we get 201 (209) matches with an offset lower than 2% (5%). To further study the reasons for these catalogue discrepancies, we separate between high $f_{\rm cont}$ (> 0.2) and high δz (> 0.05).

First, out of the 9 clusters with $f_{\rm cont} \geq 0.2$, 8 have redshift offsets $\delta z < 0.02$, with 7 of them having $0.2 \lesssim f_{\rm cont} \lesssim 0.3$. DES images with artifacts such as missing bands can impact the `MCMF` estimation of the photometric redshifts or the cluster centres. The `MCMF` algorithm includes a masking of regions with artifacts when generating the galaxy density maps, thus avoiding the region entirely. Bright saturated stars can also bias the estimations of the richness and centers depending on where they are located. Thus, `MCMF` also masks areas with bright saturated stars for the estimation of the different parameters.

Out of the 11 matches with $\delta z > 0.05$, 4 have a second significant richness peak that is in agreement with the reported redshift from the PSZ2 catalogue. Of the remaining 7, 1 has a masked area due to a bright star. For the others, the correct counterpart (and therefore redshift) is a matter of debate. For one of the systems, the `MCMF` analysis finds a peak at the PSZ2 redshift,

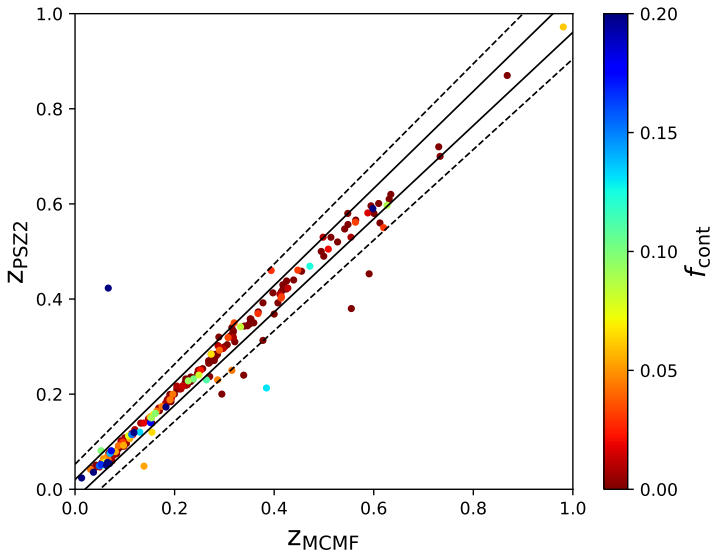

Figure 3.14: Comparison of `MCMF` photometric redshifts and those listed in the PSZ2 catalogue for the 216 matching clusters. Each cluster is color coded by the estimated $f_{\rm cont}$, saturated at $f_{\rm cont} = 0.2$. The solid and dashed lines enclose the areas where $|(z_{\rm MCMF} - z_{\rm PSZ2}) \times (1 + z_{\rm PSZ2})^{-1}| \leq 0.02, 0.05$, respectively.

although the estimated $f_{\rm cont}$ is 0.31, indicating that this counterpart has a much higher probability of being contamination as compared to the primary richness peak with $f_{\rm cont} = 0.05$.

3.9 Appendix: PSZ2 comparison examples

There are two PSZ2 clusters for which we do not find a match (see Section 3.4.3.1) in our list of optical counterparts: PSZ2 G074.08-54.68 and PSZ2 G280.76-52.30. PSZ2 G074.08-54.68 is a cluster at $z_{\rm PSZ2} = 0.305$, with $M_{500} = 5.40 \times 10^{14} {\rm M}_{\odot}$ and S/N=6.1, which is within the DES footprint and that shows a prominent optical counterpart at (R.A., Dec.)= 347.04601, -1.92133, with the redshift coming from the REFLEX catalogue (ID: RXC J2308.3-0155). The area around this cluster is not masked due to bright stars or missing DES data. Nevertheless, this cluster is not in our *Planck* SZE candidate catalogue. The PSZ2 cluster catalogue is a combination of three detection methods; PowellSnakes, MMF1 and MMF3, with the latter being the one used in this work. PSZ2 G074.08-54.68 is detected by the PowellSnakes algorithm, but not by MMF1 or MMF3. This could be due to the PSZ2 cluster being close to another cluster, PSZ2 G073.82-54.92, which might have been detected first, masking part (or all) of the flux of PSZ2 G074.08-54.68.

PSZ2 G280.76-52.30 is at $z_{\rm PSZ2} = 0.59$, with $M_{500} = 4.88 \times 10^{14}$ M$_{\odot}$ and S/N=4.5, and it has the closest *Planck* SZE position from our catalogue at 3.4 arcmin, with the optical position of that candidate having an offset of 5.8 arcmin to the PSZ2 G280.76-52.30 source. Thus, it lies just outside our 3 arcmin matching radius. The PSZ2 redshift comes from the SPT catalogue, with the SPT ID of this cluster being SPT-CL J0240-5952 (Bocquet et al., 2019). From the perspective of our analysis, the *Planck* candidate (PSZ-SN3 J0240-5945) has $z_{\rm MCMF} = 0.41$, with $\lambda_{\rm MCMF} = 79$

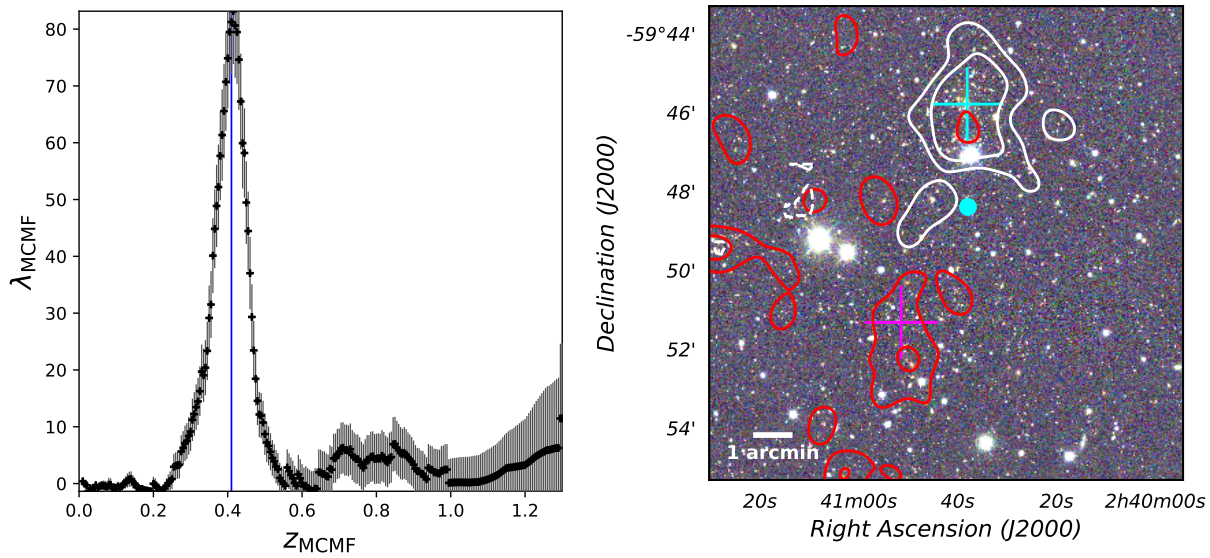

Figure 3.15: *left*: Richness as a function of redshift for the *Planck* source PSZ-SN3 J0240-5945, with the best fit peak at $z_{\rm MCMF} = 0.41$. *right*: pseudo-color image from DES *g*, *r*, *i* bands near the *Planck* candidate coordinates, which are marked with a cyan dot. Contours are from the galaxy density maps of the counterpart at $z_{\rm MCMF} = 0.41$ (white) and at $z_{\rm MCMF} = 0.605$ (red). The cyan cross marks the position of the optical counterpart found using `MCMF`, and the magenta cross marks the position of PSZ2 G280.76-52.30.

and an estimated $f_{\rm cont} = 0.03$. The second redshift peak that we find is at $z_{\rm MCMF} = 0.605$, which is closer to the PSZ2 redshift. In Fig. 3.15 we show (top) the richness as a function of redshift, while below we show the *gri* DES pseudo-color image. Overlayed are the density contours at $z_{\rm MCMF} = 0.41$ (white) and $z_{\rm MCMF} = 0.605$ (red). The cyan cross shows the optical position found by `MCMF`, and the magenta cross shows the position of PSZ2. For the peak at $z_{\rm MCMF} = 0.605$ with $\lambda_{\rm MCMF} = 50$, we estimate $f_{\rm cont} = 0.28$, which means that we consider this to be a candidate with a second optical counterpart (requires $f_{\rm cont} < 0.3$). It is worth noting that, by using the same cross-match aperture, we find a match with the SPT-2500d catalogue (Bocquet et al., 2019), SPT-CL J0240-5946, whose reported redshift is $z_{\rm SPT} = 0.4$.

3.10 Appendix: Further exploration of the *Planck* candidate list contamination

To investigate the difference between the observed contamination of ~51% and the 75% contamination estimated from the *Planck* sky simulations (see Section 3.2.2) we compare the detection threshold $Y^{\rm min}_{5R500} = 2 \times 10^{-4}$ arcmin^2 to the observed Y_{5R500} distribution of our candidates. We will do this in three steps: First we will estimate an observed mass by means of the $\lambda_{\rm MCMF}$-M_{500} relation derived in Section 3.10.1. Secondly, we will determine the excess distribution of candidates with respect to the random lines-of-sight in different redshift ranges for the S/N>3 sample, which should give an estimate of the number of real clusters within this redshift range. Finally, we will use the derived parameters from the scaling relation and we will map from $\lambda_{\rm MCMF}$ to

Y_{5R500} on our excess clusters, using the $M_{500} - Y_{5R500}$ relation from equation (3.2). With this, we can estimate the ratio of excess candidates with $Y_{5R500} \geq Y^{\rm min}_{5R500}$ with respect to the total number of excess candidates, which would give us an indication of how many real systems we expect to lose when applying this limiting value.

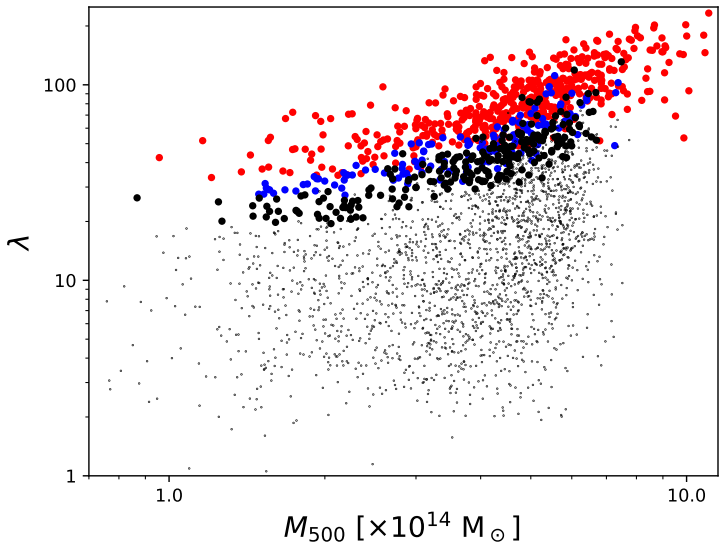

Figure 3.16: Richness versus mass for the *Planck* cluster candidates. Small black dots are candidates with $f_{\rm cont} \geq 0.2$. Bigger black, blue and red points represent counterparts with $0.1 \leq f_{\rm cont} < 0.2$, $0.05 \leq f_{\rm cont} < 0.1$ and $f_{\rm cont} < 0.05$, respectively.

3.10.1 Richness–mass relation

Previous analyses have shown that the number of galaxies in a cluster (or richness) is approximately linearly proportional to the cluster mass (Lin et al., 2004; Gladders et al., 2007; Rozo et al., 2009a,b; Klein et al., 2019), with some intrinsic scatter ($\sigma_{\rm int} \approx 25\%$, Rozo & Rykoff, 2014). Fig. 3.16 shows how the derived masses behave with the estimated `MCMF` richness of the candidates. Colors red, blue and black represent the different $f_{\rm cont}$ selection thresholds following Fig. 3.6. Although at $f_{\rm cont} > 0.2$ the cloud of points does not seem follow any particular relation, the more reliable clusters with $f_{\rm cont} < 0.2$ exhibit a roughly linear trend at $M_{500} \gtrsim 2.5 \times 10^{14}$ M$_\odot$. The trend is stronger at lower $f_{\rm cont}$, where the contamination of the cluster sample is lowest.

We fit a $\lambda_{\rm MCMF} - M_{500}$ relation to our data but only for the high S/N sample (S/N>4.5 and $f^{\rm max}_{\rm cont} = 0.3$), which, assuming a catalogue contamination of $f_{\rm SZE-cont} \approx 8.5\%$ (Section 3.2.2), means a purity of ~97.4%. For the fitting, we follow a similar procedure as the one described in Klein et al. (2022), where the distribution of richnesses $\lambda_{\rm MCMF}$ is assumed to follow a log-normal distribution which depends on the mass M_{500} and redshift z, so that

$$P(\ln\lambda|M_{500}, z) = \mathcal{N}(\ln\lambda; \langle\ln\lambda\rangle\,(M_{500}, z), \sigma^2(M_{500}, z)), \tag{3.7}$$

with mean

$$\langle \ln \lambda \rangle \, (M_{500}, z) = \ln \lambda_0 + \alpha_0 + \alpha_M \ln \left(\frac{M_{500}}{M_0} \right) + \alpha_z \ln \left(\frac{1+z}{1+z_0} \right) \tag{3.8}$$

and variance

$$\sigma^2 (M_{500}, z) = \exp(\ln \zeta(z) - \langle \ln \lambda \rangle) + \exp(s), \tag{3.9}$$

where λ_0, M_0 and z_0 are pivots and (α_0, α_M, α_z and s) are constrained by the likelihood analysis. For the pivots we use the median values of the richness, mass and redshift of the S/N>4.5 sample. The $\zeta(z)$ parameter corresponds to the richness correction factor used on `MCMF` (see equation 7 from Klein et al., 2019). This first term on the variance captures the Poisson noise on the measured richness, while the second term represents the intrinsic variance within the cluster population. We refer the reader to Appendix A of Klein et al. (2022) for further details on the likelihood analysis. We find best fit values for the parameters of $\alpha_0 = -0.004 \pm 0.023$, $\alpha_M = 0.961 \pm 0.071$, $\alpha_z = 0.095 \pm 0.252$ and $s = -2.151 \pm 0.101$.

3.10.2 On the difference between 51% and 75% initial contamination

For the distribution of excess candidates, we use the PSZ-MCMF cluster catalog to define the redshift ranges using the 25%, 50% and 75% percentiles, corresponding to ranges of $0 < z \leq 0.18$, $0.18 < z \leq 0.29$, $0.29 < z \leq 0.44$ and $0.44 < z \leq 1.32$. For each of these ranges we look at the $\lambda_{\rm MCMF}$ distribution of *Planck* candidates and that of the random lines-of-sight, rescaling the latter to fit the candidates distribution at low $\lambda_{\rm MCMF}$. We then subtract the number of scaled random lines-of-sight to the number of *Planck* candidates for each $\lambda_{\rm MCMF}$ bin within a redshift range. We refer to this as the distribution of excess candidates, which maps the distribution of real clusters down to low λ without accounting for catalogue purity, unlike when a $f_{\rm cont}$ threshold $f_{\rm cont}^{\rm max}$ is applied.

Finally we transform this $\lambda_{\rm MCMF}$ distribution into a M_{500} distribution, and then into a Y_{5R500} distribution. Fig. 3.17 shows the different steps on the estimation of the excess and the final transformations. The top panel shows the distributions of candidates and scaled random lines-of-sight for the $0 < z \leq 0.18$ redshift range as purple and orange lines, respectively, with the excess candidates, labeled as "residual", shown in red. The bottom shows the distribution of the excess candidates for the different redshift bins in terms of Y_{5R500}. The vertical gray line marks the $Y_{5R500}^{\rm min}$. Depending on how we scale the randoms to fit the low richness regime of the candidates, the ratio of the total excess candidates (summed at all redshifts) to that of the total number of candidates varies between 55-65%. Of those, ~50% are below the $Y_{5R500}^{\rm min}$ threshold, regardless of the normalisation, meaning that we would expect to lose half of the real sources by applying this threshold.

We note that we do not probe the $\lambda_{\rm MCMF} - M_{500}$ relation at $\lambda < 10$ (see Fig. 3.16), which, depending on the redshift, could translate to $M_{500} < 10^{14}$ $M_\odot$. However, this does not affect our analysis because, as can be seen in Fig. 3.6, the minimum λ in our sample is $\lambda_{\rm min} \sim 20$ at $z \approx 0.03$, which corresponds to a mass $M_{500} > 10^{14}$ $M_\odot$ using our scaling relation.

The different arrows on the bottom panel show, for each redshift range, the richness of a source with $f_{\rm cont} = 0.2$, $\lambda_{\rm MCMF,min}$, translated into a $Y_{5R500,\rm min}$, color codded according to the redshift

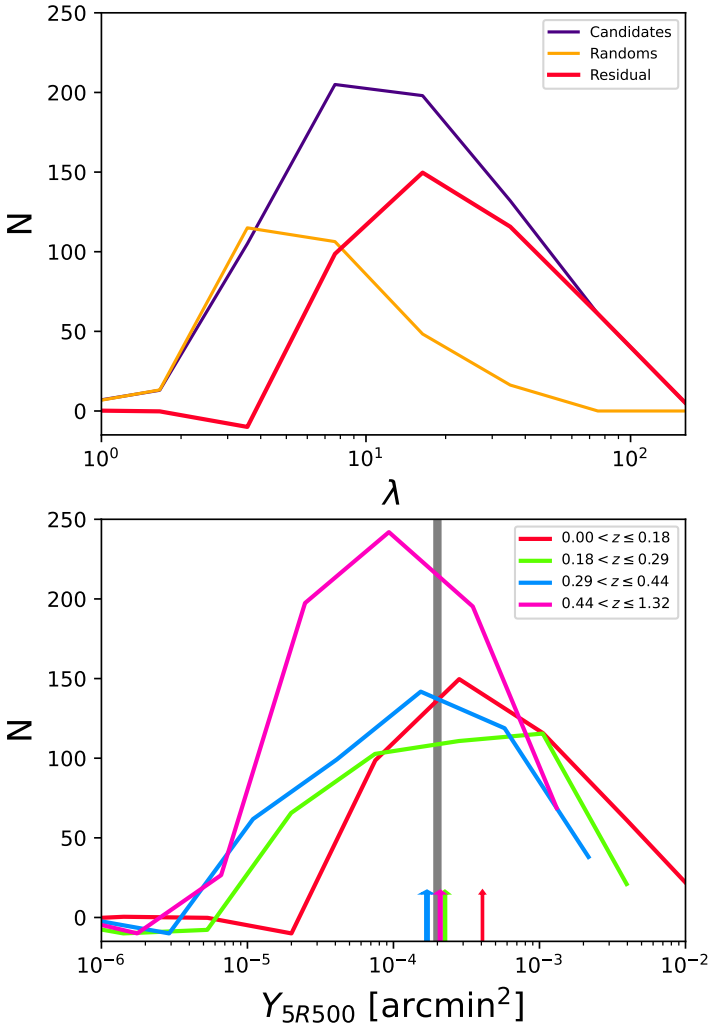

Figure 3.17: *top:* Richness distributions in the redshift range $0 < z_{\rm MCMF} \leq 0.18$ of candidates, rescaled randoms and excess candidates shown as purple, orange and red lines, respectively. *bottom:* Y_{5R500} distributions of excess candidates for different redshift bins. Y_{5R500} is determined from $\lambda_{\rm MCMF}$ using scaling relations. Colored arrows correspond to the richness derived Y_{5R500} a candidate would need to have for us to consider it a real cluster in our sample, estimated using the median redshift of all the candidates for each redshift bin. The gray line marks the $Y^{\rm min}_{5R500}$ used to estimate the purity in Section 3.2.2.

range. Each $\lambda_{\rm MCMF,min}(z)$ is estimated using the median redshift of each redshift range. These arrows can be interpreted as the selection thresholds that are applied when selecting candidates with $f_{\rm cont} < 0.2$, showing good agreement with $Y^{\rm min}_{5R500}$ at all redshifts. This can also be seen in Fig. 3.9, where at $f^{\rm max}_{\rm cont} = 0.2$, the number of clusters over the number of candidates is ∼25%, similar to the value expected using $Y^{\rm min}_{5R500}$ (Section 3.2.2).

We note that groups and clusters corresponding to the difference between 51% and 75% have $0.2 < f_{\rm cont} < 0.65$. They correspond to small black dots in Fig. 3.16 and are thus subject to a strong selection bias. For these systems, the measured SZE signal is dominated by a positive

noise fluctuation on top of an actual small SZ signal from the cluster. The conversion of the measured SZE signal to the mass thus provides strongly overestimated values. However, we are already excluding most of these systems in the final catalogue where we only add 11 S/N>4.5 clusters with $0.2 \leq f_{\rm cont} < 0.3$.

Table 3.2: Thirty entries of the new PSZ-MCMF cluster catalogue, limited to the most important columns of the catalogue. The full cluster catalogue containing 853 clusters will be available online via VizieR server at CDS (http://vizier.u-strasbg.fr) and as supplementary material, including additional columns. The first five columns show the cluster name, the *Planck* position, the S/N and the positional uncertainty of the candidates. The next columns are `MCMF` derived quantities for the most likely optical counterpart (lowest $f_{\rm cont}$). These are: (6 and 7) the optical position of the cluster, (8) the distance between the optical and the *Planck* positions normalised by the positional uncertainty, (9) the redshift and (10) the richness. Column (11) is the estimated probability that the source is a random superposition and column (12) contains the mass estimate M_{500}. Columns (13–15) contain flags that are described in the text.

Name	R.A.$_{Planck}$	Dec.$_{Planck}$	S/N	$\sigma_{\rm pos}$	R.A.*	Dec.*	$\sigma^*_{\rm SZ-Op}$	$z^*_{\texttt{MCMF}}$	$\lambda^*_{\texttt{MCMF}}$	$f^*_{\rm cont}$	M_{500}	FLAG	FLAG	FLAG
	deg.	deg.		arcmin	deg.	deg.					$\times10^{14}$ M$_\odot$	COSMO	CLEAN	QNEURAL
PSZ-SN3 J0543-1857	85.9328	-18.9048	3.8526	5.5398	85.9314	-18.9595	0.5922	0.6568	83.1978	0.0530	6.4115	1	1	1
PSZ-SN3 J0550-2233	87.6461	-22.6607	3.4805	6.9070	87.6388	-22.5545	0.9247	0.0801	34.5207	0.0476	1.9277	1	1	1
PSZ-SN3 J0548-2152	87.1619	-21.9176	4.0710	6.7710	87.0182	-21.8678	1.2612	0.0862	42.4760	0.0197	2.3530	1	1	1
PSZ-SN3 J0558-2629	89.7651	-26.5085	4.2204	6.1337	89.7495	-26.4901	0.2257	0.2681	84.8770	0.0000	4.4636	1	1	1
PSZ-SN3 J0555-2637	88.8111	-26.6496	3.3028	5.2495	88.8133	-26.6195	0.3446	0.2766	71.8997	0.0000	3.9173	1	1	1
PSZ-SN3 J0554-2556	88.7011	-25.9919	3.0569	5.5877	88.6952	-25.9386	0.5750	0.2714	54.1370	0.0505	3.6558	1	1	1
PSZ-SN3 J0616-3948	94.1341	-39.8328	8.7490	2.5549	94.1227	-39.8033	0.7235	0.1506	41.0926	0.0545	5.0250	1	1	1
PSZ-SN3 J0609-3700	92.4479	-36.9542	3.9032	4.9092	92.4358	-37.0106	0.6994	0.2685	72.3089	0.0000	4.1736	1	1	1
PSZ-SN3 J0638-5358	99.7055	-53.9805	13.5941	2.3905	99.6977	-53.9745	0.1891	0.2320	92.2980	0.0000	8.4066	1	1	1
PSZ-SN3 J2118+0033	319.7380	0.5439	4.8582	2.2707	319.7118	0.5598	0.8107	0.2686	97.6065	0.0000	5.4214	1	1	1
PSZ-SN3 J2119+0120	319.9540	1.3562	3.9086	4.4458	319.9744	1.3381	0.3680	0.1237	31.1533	0.0955	2.8914	1	1	1
PSZ-SN3 J0514-1951	78.6623	-19.9231	4.1396	4.9319	78.6868	-19.8534	0.8936	0.1357	48.7245	0.0156	3.3206	1	1	1
PSZ-SN3 J0502-1813	75.7216	-18.1993	3.4844	5.6044	75.6171	-18.2214	1.0885	0.5749	60.3850	0.1406	5.5214	1	1	1
PSZ-SN3 J0548-2530	87.1631	-25.4926	6.3700	5.4270	87.1715	-25.5048	0.1593	0.0311	36.2221	0.0431	1.6597	1	1	1
PSZ-SN3 J0528-2942	82.0841	-29.7250	5.7007	3.8145	82.0811	-29.7101	0.2386	0.1583	40.3784	0.0691	4.1431	1	1	1
PSZ-SN3 J0538-2038	84.5847	-20.6454	5.4451	4.9996	84.5856	-20.6370	0.1010	0.0871	37.0128	0.0421	2.9113	1	1	1
PSZ-SN3 J0520-2625	80.1255	-26.4429	4.9942	1.7900	80.1125	-26.4237	0.7537	0.2790	77.4486	0.0000	5.2686	1	1	1
PSZ-SN3 J0516-2237	79.2379	-22.6223	4.9108	4.2022	79.2386	-22.6249	0.0390	0.2949	98.6374	0.0000	5.6110	1	1	1
PSZ-SN3 J0529-2253	82.4507	-22.8641	4.4810	5.8471	82.4575	-22.8849	0.2227	0.1776	57.6785	0.0221	3.7982	1	1	1
PSZ-SN3 J0516-2521	79.1042	-25.3215	4.4406	6.2807	79.0631	-25.3624	0.5279	0.2808	56.1890	0.0331	4.6955	1	1	1
PSZ-SN3 J0519-2056	80.0775	-21.0020	3.8548	5.1985	79.9872	-20.9448	1.1760	0.3039	77.7563	0.0000	4.9633	1	1	1
PSZ-SN3 J0540-2127	85.1922	-21.4732	3.5656	5.4369	85.2095	-21.4615	0.2195	0.5242	77.4497	0.0461	5.7749	1	1	1
PSZ-SN3 J0521-2754	80.3559	-27.9377	3.9751	3.6068	80.3628	-27.9156	0.3821	0.3150	98.6217	0.0000	4.7847	1	1	1
PSZ-SN3 J0545-2556	86.3885	-25.9053	4.0783	5.5571	86.3658	-25.9377	0.4132	0.0469	37.0815	0.0371	1.5132	1	1	1
PSZ-SN3 J0533-2823	83.5642	-28.3836	3.7296	5.2059	83.4791	-28.3934	0.8706	0.1916	49.4947	0.0542	3.1730	1	1	1
PSZ-SN3 J0530-2226	82.6648	-22.4653	4.5761	3.3990	82.6643	-22.4453	0.3528	0.1680	71.5937	0.0067	3.6625	1	1	1
PSZ-SN3 J0546-2410	86.6542	-24.1719	3.1268	5.3599	86.5557	-24.1669	1.0076	0.3247	46.8889	0.0797	4.1055	1	1	1
PSZ-SN3 J0530-2556	82.6693	-25.7861	3.1967	5.1885	82.7240	-25.9483	1.9600	0.1926	49.8588	0.0535	2.9369	1	1	1
PSZ-SN3 J0605-3519	91.4696	-35.3091	15.2181	2.3889	91.4492	-35.3206	0.5087	0.5210	155.6621	0.0000	12.7754	1	1	1
PSZ-SN3 J0553-3342	88.3485	-33.7020	13.0358	1.6357	88.3436	-33.7078	0.2613	0.4174	146.0235	0.0000	10.8541	1	1	1

* In the case of a second prominent optical counterpart (with $f_{\rm cont} < 0.3$) at a different $z_{\texttt{MCMF}}$, we provide an entry for that counterpart as well.
† Masses derived in Section 3.4.2 are divided by 0.8 to correct for the estimated hydrostatic mass bias.

CHAPTER 4

RASS-MCMF: A FULL-SKY X-RAY SELECTED GALAXY CLUSTER CATALOG

Matthias Klein, Daniel Hernández-Lang, Joseph J Mohr, Sebastian Bocquet, and Aditya Singh, *Monthly Notices of the Royal Astronomical Society* 526, 3757-3778, 2023

We present the RASS-MCMF catalog of 8,449 X-ray selected galaxy clusters over 25,000 deg^2 of extragalactic sky. The accumulation of deep, multiband optical imaging data, the development of the Multi-Component Matched Filter cluster confirmation algorithm (MCMF), and the release of the DESI Legacy Survey DR10 catalog makes it possible– for the first time, more than 30 years after the launch of the ROSAT X-ray satellite– to identify the majority of the galaxy clusters detected in the second ROSAT All-Sky-Survey (RASS) source catalog (2RXS). The resulting 90% pure RASS-MCMF catalog is the largest ICM-selected cluster sample to date. RASS-MCMF probes a large dynamic range in cluster mass spanning from galaxy groups to the most massive clusters. The cluster redshift distribution peaks at $z \sim 0.1$ and extends to redshifts $z \sim 1$. Out to $z \sim 0.4$, the RASS-MCMF sample contains more clusters per redshift interval (dN/dz) than any other ICM-selected sample. In addition to the main sample, we present two subsamples with 6,912 and 5,506 clusters, exhibiting 95% and 99% purity, respectively. We forecast the utility of the sample for a cluster cosmological study, using realistic mock catalogs that incorporate most observational effects, including the X-ray exposure time and background variations, the existence likelihood selection and the impact of the optical cleaning with the algorithm MCMF. Using realistic priors on the observable–mass relation parameters from a DES-based weak lensing analysis, we estimate the constraining power of the RASS-MCMF×DES sample to be of 0.026, 0.033 and 0.15 (1σ) on the parameters Ω_m, σ_8 and w, respectively.

4.1 Introduction

Selecting galaxy clusters through their intracluster medium (ICM) signatures– either X-ray emission (e.g., Sarazin, 1988) or the thermal Sunyaev-Zel'dovich effect (SZE; Sunyaev & Zeldovich, 1972)– is an efficient way to create cluster samples that can be employed for cosmological analyses (e.g., Vikhlinin et al., 1998; Mantz et al., 2010; Planck Collaboration et al., 2016b). X-ray and SZE signatures are dominated by processes in the hot and dense cluster virial regions, which ensures that the distribution of clusters in observable and redshift– the so-called halo observable function (HOF)– can be related to the underlying halo mass function (HMF) through observable–mass relations. An accurate mapping from HOF to HMF is crucial for carrying out cosmological studies using the abundance of galaxy clusters (e.g., Hu, 2003; Majumdar & Mohr, 2004; Lima & Hu, 2005).

In contrast to ICM selection, selecting galaxy clusters through their passive galaxy populations– the so-called red sequence methods (Gladders & Yee, 2000; Rykoff et al., 2014)– relies upon a cluster signature that traces not only the dense cluster virial regions but also the low density regions outside cluster and group halos. Neither photometric redshifts, galaxy colors nor spectroscopic redshifts can be used to identify whether galaxies along the line of sight toward the cluster lie within the cluster virial region or in the surrounding region that extends 10 to 20 Mpc behind and in front of the cluster (e.g., Song et al., 2012a; Saro et al., 2013). This additional "contrast" challenge complicates the interpretation of the number of cluster galaxies– the richness λ– and its relationship to halo mass, and it may also weaken the required one-to-one relationship between optically selected clusters and collapsed halos, making it more difficult to use optically selected cluster abundance to study cosmology. Methods are being developed to overcome these challenges and have been employed to deliver cosmological constraints (e.g., Costanzi et al., 2019; Murata et al., 2019; Abbott et al., 2020; Lesci et al., 2022).

ICM-selected cluster samples have to be followed up optically to determine the cluster redshifts. With overlapping deep, multi-band surveys (e.g., KiDS, DES and HSC-SSP; de Jong et al., 2013; Flaugher et al., 2015; Aihara et al., 2018) it is possible to do much more. One can use the richness of the optical counterpart of an ICM selected cluster to exclude those cluster candidates with low significance optical counterparts, because they are likely contamination (Klein et al., 2018). The "contrast" challenge mentioned above has no impact on this process. Thus, the multi-band survey data allow one to make the most of an X-ray or SZE survey, because one can include ICM-selected counterparts with lower ICM detection significance without increasing the contamination fraction of the final cluster sample. A benefit of this approach is that lower-mass clusters are included at all redshifts, and the maximum redshift probed by the sample is increased.

Optical followup based on the passive galaxy population has been shown to be robust for clusters and for low redshift high mass groups. For systems with $M_{500} \geq 3 \times 10^{14} M_{\odot}$, purely ICM-selected cluster samples from, e.g., SPT (Carlstrom et al., 2011), exhibit dominant passive galaxy populations out to redshifts $z \sim 1$ (Hennig et al., 2017), and deep Spitzer and HST studies of five of the highest redshift SZE selected clusters from SPT at $1.4 < z < 1.72$ show higher passive fractions than the field at comparable redshift, indicating environmental quenching efficiencies in the range of 0.5 to 0.8 (Strazzullo et al., 2019). Moreover, in the recent eFEDS X-ray study with eROSITA (Predehl et al., 2021), extended sources with masses $M_{500} > 5 \times 10^{13} M_{\odot}$ exhibit

passive dominated galaxy counterparts over the redshifts range that they are sampled (Klein et al., 2022).

In addition to redshift estimation, data from deep, multiband surveys like KiDS, DES and HSC-SSP enable one to use the weak gravitational lensing shear and photometric redshift measurements of background galaxies to directly constrain the halo mass distribution of the clusters. In general, these survey weak lensing datasets are created with the goal of carrying out cosmic shear studies and are therefore more homogeneous with better understood systematics than the pointed cluster weak lensing datasets that have been employed for cluster cosmology in the past.

Larger, high-purity ICM-selected cluster samples extending over a broader mass and redshift range together with weak-lensing mass information on the full sample enable more accurate and precise cosmological studies. Thus, the combination of X-ray or SZE cluster surveys with deep and homogeneous multi-band optical survey data offers the promise to produce the most constraining cluster cosmological studies to date.

Initial examples of this approach are now emerging. The optical followup of ACT and SPT-selected cluster candidates already heavily relies on survey data from DES (Bleem et al., 2020; Hilton et al., 2021). New and dramatically larger ICM-selected cluster samples have been produced using the Multi-Component Matched Filter (MCMF) technique. An analysis of the ROSAT All-Sky X-ray Survey (RASS) together with the DES led to a sample of ∼2,000 X-ray selected clusters extending to redshift $z \sim 1$ with a surface density that is an order of magnitude higher than that of past RASS cluster catalogs over the same sky region (Klein et al., 2019). An analysis of the Planck SZE selected cluster candidate sample to lower signal to noise (Planck Collaboration et al., 2014b) in combination with DES led to a factor of four increase in the number of confirmed clusters over the DES region and extended the maximum redshift of the Planck cluster sample to $z \sim 1$ (Hernández-Lang et al., 2022). A similar analysis of the SPT-SZ and SPTpol 500d SZE surveys in combination with DES has also led to a significant increase in the mass range of that SZE selected sample (Klein et al., 2024a, Bleem et al.,in prep.).

At present no DES weak-lensing informed cosmological analysis of these enhanced samples has been presented, but several are underway. However, in the case of the eROSITA eFEDS pilot X-ray survey, the cosmological analysis of the cluster sample has been carried out in combination with the weak-lensing dataset of the HSC-SSP (Chiu et al., 2023).

These efforts provide evidence of the benefits of combining ICM-selected samples with large solid angle, deep multi-band optical surveys. In this paper we employ the MCMF tool in combination with the latest reanalysis of the all-sky X-ray survey RASS (2RXS; Boller et al., 2016) and the latest release of the optical and IR multi-band Legacy Survey DR10 (Dey et al., 2019, in prep.) to produce a new X-ray selected cluster sample called RASS-MCMF. Our multi-wavelength analysis extends over the bulk of the extragalactic sky, a region of over 25,000 deg^2, and yields a high-purity sample of over 8,000 X-ray selected galaxy clusters with redshifts, richnesses, optical centers and X-ray fluxes.

In Section 4.2 we present the data used in this analysis. Section 4.3 contains a summary of the methods used for cluster catalog construction, and Section 4.4 presents the RASS-MCMF cluster sample. The results of cross-comparison to other cluster samples is presented in Section 4.5, while Section 4.6 contains a cosmological forecast that highlights the usefulness of the sample. Conclusions are presented in Section 4.7. Throughout this paper we assume a flat ΛCDM model

with $\Omega_m = 0.3$ and $H_0 = 70\ \mathrm{km\,s^{-1}\,Mpc^{-1}}$ unless otherwise stated..

4.2 Data

In the following subsections we describe the X-ray and multi-band optical and IR datasets used for this analysis.

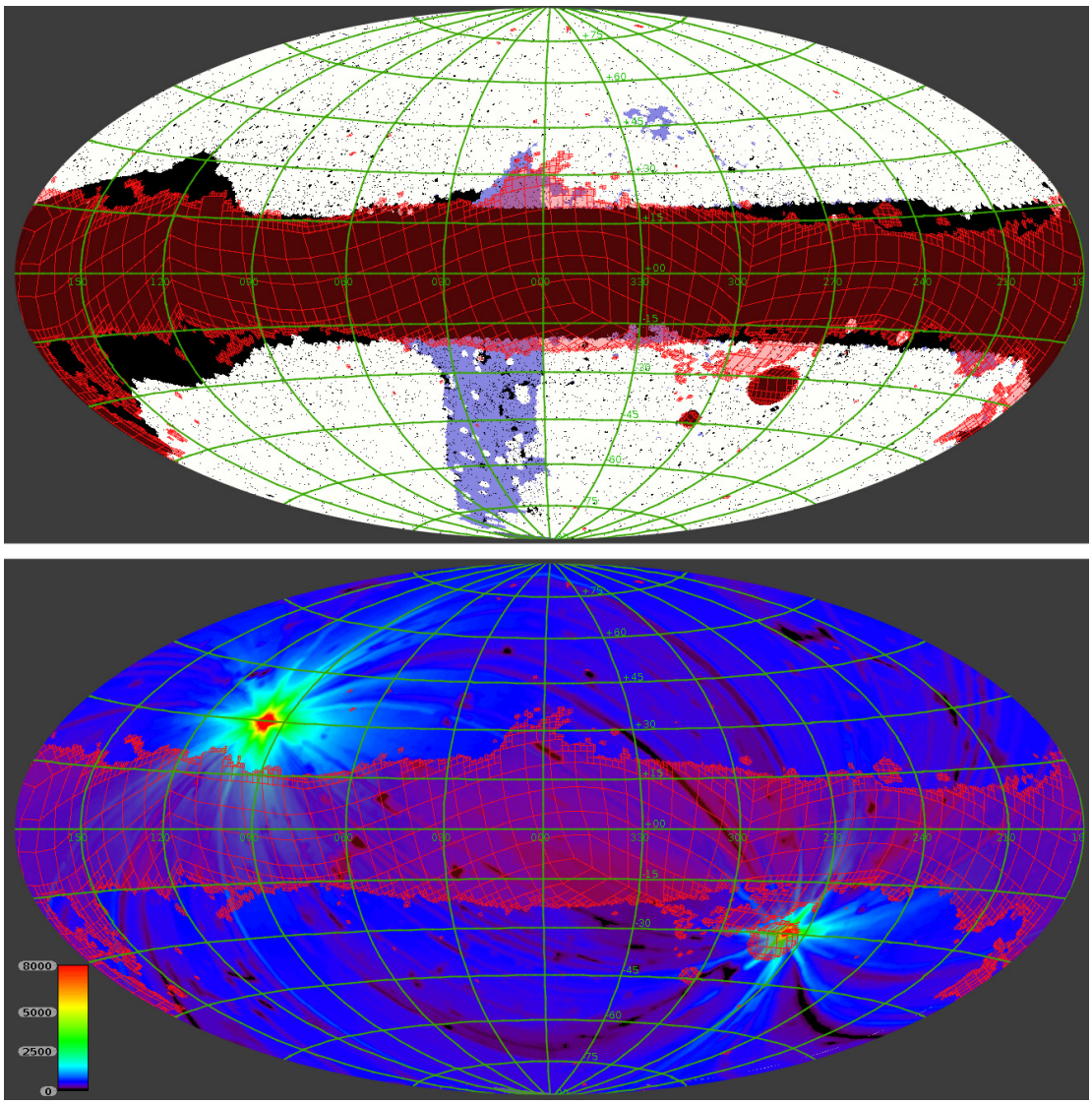

Figure 4.1: Top: The DESI Legacy Survey DR10 showing extragalactic coverage in grz (white) and gi (blue). Both datasets are supplemented with WISE $w1$ and $w2$ photometry. Bottom: The ROSAT All-Sky Survey exposure map from the 2RXS source catalog is drawn. Highlighted in red are sky regions impacted by high galactic NH column densities ($N_H > 10^{21}$ cm^{-2}) or high stellar densities and therefore less suited for cluster search.

4.2.1 The Second ROSAT All-Sky Survey Source catalog

The ROSAT All-Sky Survey (RASS) was performed more than thirty years ago with the ROSAT satellite (Trümper, 1993) from June 1990 to August 1991. It was the first all-sky imaging survey in X-rays, resulting in an increase in known X-ray sources by a factor of ∼100 (Voges et al., 1999). The second ROSAT All-Sky-Survey source catalog (2RXS; Boller et al., 2016) builds upon previous work (1RXS Voges et al., 1999, 2000) and uses the more recent RASS-3 processed photon event files together with an improved source detection algorithm resulting in a catalog of 135,000 X-ray sources. Dedicated simulations were performed to estimate the contamination by spurious sources as a function of existence likelihood (EXI_ML). The released 2RXS catalog is expected to include ∼30% spurious sources and contains all sources with existence likelihood EXI_ML≥6.5. We expect from previous studies (e.g., Table 2 in Hasinger (1996) and also Klein et al. (2019)) that $\sim 10-15\%$ of the X-ray sources in extragalactic regions are likely galaxy clusters. This suggest that 2RXS might include approximately 10, 000 groups and clusters. Given the typically low signal to noise of 2RXS sources together with the poor RASS angular resolution of ∼4 arcmin (Boese, 2000), a clean and complete selection of these 2RXS detected clusters is not possible from ROSAT data alone. Figure 4.1 shows the sky coverage of the RASS data used in this analysis. The primary 2RXS inputs for the analysis that follows are the source positions and X-ray count rates.

4.2.2 DESI Legacy Survey DR10

The DESI Legacy Imaging Surveys (LS; Dey et al., 2019) up to data release 8, was a combination of four imaging surveys, the 9,000 deg^2 grz-band DECam (Flaugher et al., 2015) based DECaLS survey, the 5,000 deg^2 BASS and MzLS surveys providing photometry in g, r and z-band, respectively, and the WISE and NEOWISE surveys in the mid-IR at 3.4μm and 4.6μm. With the subsequent data releases other DECam based imaging has been included. The most recent data release, DR10, includes imaging data from the Dark Energy Survey as well as from various other survey programs such as BLISS and the DeROSITAS survey. While the BLISS program focuses on imaging the complete DECam-observable sky, the focus of the DeROSITAS survey was to obtain imaging data to enable cluster identification for the portion of the eROSITA X-ray survey that lies within the western galactic hemisphere (the so-called German portion of the eROSITA sky). Therefore a special focus was put on image quality and depth; experience gained from previous studies of 2RXS sources over DES (Klein et al., 2019) helped define the DeROSITAS survey parameters and observing plan.

The recent Legacy DR10 (Dey et al., in prep.) is the addition of DECam imaging data from the i-band, mostly coming from the DeROSITAS and DES surveys. The imaging depth depends on sky position, given the differing requirements of the various surveys. The 5,000 deg^2 BASS and MzLS surveys, which we call Legacy Survey North, exhibits a typical 5σ point source depth of ∼24.3, ∼23.8 and ∼23.4 mag in the grz bands, respectively. The DECam based surveys show typically a double peaked imaging depth distribution [1], the shallower peak is mostly associated with the DECaLS survey with depths of ∼24.8, ∼24.2 and ∼23.3 mag (grz) and the deeper peak

[1] See https://www.legacysurvey.org/dr10/description/

with the DES survey with depths of ∼25.3, ∼25.0 and ∼23.9 mag. In all cases we consider the imaging depth to be sufficient for the identification of clusters in 2RXS.

The combination of the optical data with the most recent WISE data allows for improved redshift constraints and increased redshift range due to the strong dependency of the $z - w1$ or $i - w1$ color on redshift. In the top panel of Figure 4.1 we show the coverage of the DESI Legacy Survey DR10, split into two regions, those containing grz optical imaging (white) and those with gi imaging with no z band (blue). The union of both regions is employed for galaxy cluster identification in this analysis. While the area covered with gi imaging is as large as ∼17,000 deg^2 we only consider cluster measurements using the gi dataset within the blue region, which corresponds to ∼1,700 deg^2. Outside this region we employ the grz imaging for cluster studies. The combined footprint is $\sim 25,000$ deg^2, covering the majority of the extragalactic sky outside a Galactic latitude of ±17 deg.

Because the Legacy Survey North was conducted with different instruments than that of the rest of the survey, slight differences between colors are expected. We calibrate observed galaxy colors and magnitudes from the Legacy Survey North to the Legacy Survey South system (see, e.g., Duncan, 2022). Furthermore, we first treat the cluster analysis of the Legacy Survey North region independently, and only merge it with the DECam based dataset after cross checks have been performed.

4.3 Method

To identify the subset of 2RXS sources that are galaxy clusters detected due to their ICM X-ray signatures, we use the multi-component matched filter algorithm (MCMF; Klein et al., 2018, 2019). We search for optical counterparts of 2RXS sources in the Legacy Survey, measuring redshifts and richnesses for each so that we can evaluate the probability that each is an X-ray selected galaxy cluster. This allows us to select a high-purity sample of candidate galaxy clusters from 2RXS or indeed to select several samples with different sizes and sample purities. Thereafter, we combine multiple 2RXS detections of the same galaxy clusters, and finally we apply an additional point source rejection method, producing the final cluster catalog.

4.3.1 Selecting a high-purity cluster sample

MCMF was created for the identification of true clusters in ICM selected candidate catalogs. It was first applied to the same 2RXS catalog used in this work (Klein et al., 2018, 2019) to create a galaxy cluster catalog over the DES area. Later it was also used to identify the first clusters identified by eROSITA over the eFEDS footprint (Klein et al., 2022) and to identify SZE selected clusters from the Planck survey (Hernández-Lang et al., 2022). We therefore present only a brief description in this work and refer the reader to the aforementioned publications for additional details.

The MCMF identification of clusters is based on the red sequence (RS) of cluster galaxies (Gladders & Yee, 2000) and the weighted number– called richness λ– of excess RS galaxies within a certain magnitude and radial range around the X-ray position. The weights include a

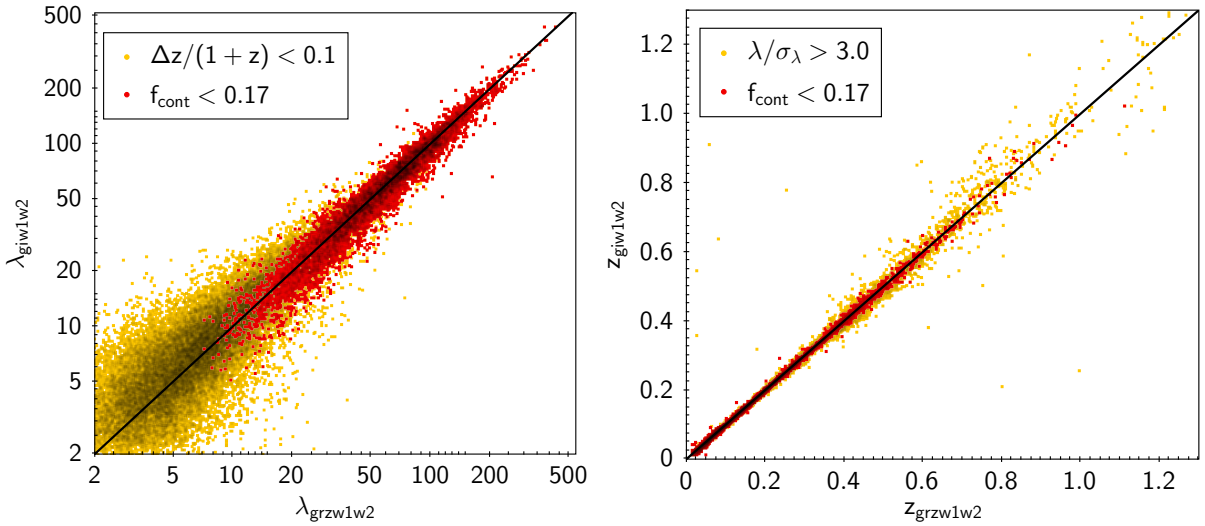

Figure 4.2: Comparison of richnesses λ and redshifts z when using grz,$w1$,$w2$ to those for the same clusters when using gi,$w1$,$w2$ bands. In the left plot we show in yellow all sources with similar redshifts in both runs and in red the subset of clusters entering the final cluster catalog. The right plot shows all sources with richnesses greater than three times the richness uncertainty (yellow) and clusters making the selection into the final cluster catalog. The plots are heavily saturated containing $> 33,000$ clusters on the left and $> 11,000$ on the right.

radial filter following a projected NFW profile (Navarro et al., 1996) and a color-magnitude filter tuned to select red sequence galaxies. The color-magnitude filter has the form,

$$w_i(z) = \frac{\prod_{j=1}^{N} G\left(c_{i,j} - \langle c(f,z)\rangle_j, \sigma_{c_{i,j}}(f,z)\right)}{N(\sigma_{c_{i,1}}(f,z), \sigma_{c_{i,2}}(f,z), \sigma_{c_{i,3}}(f,z))}. \tag{4.1}$$

Here $G\left(c_{i,j} - \langle c(f,z)\rangle_j, \sigma_{c_{i,j}}(f,z)\right)$ is the value of the normalized Gaussian function at a color offset between observed color j and predicted RS color given observed reference band magnitude f of source i and assumed redshift z.

In the current analysis, we run MCMF in two different set of bands. The grz-mode, uses the combinations c_j of $g-r$, $r-z$, $z-w1$ and $w1-w2$. In this mode we use the z-band magnitude as the magnitude reference band f. To maximize the footprint we also run MCMF in the gi-mode which uses $g-i$, $i-w1$ and $w1-w2$. Here we use the i-band for the reference band f. The standard deviation of the Gaussian weight function is the combination of intrinsic and measured scatter as $\sigma_{c_{i,j}}(f,z) = \sqrt{\sigma^2_{\mathrm{mcor}}(f,z) + \sigma^2_{\mathrm{meas,i}}}$.

Given the additional redshift information provided by adding Mid-IR WISE data to the optical set of colors we expand the redshift range compared to previous runs from $z = 1.3$ to $z = 1.5$ and for each sample we scan through 300 redshift bins calculating the richness. For each redshift bin an aperture corresponding approximately to r_{500} is estimated, based on the X-ray count rate, the redshift of the bin and an observable-mass scaling relation.

Following previous MCMF analyses (Klein et al., 2018, 2019), we convert the observed X-ray count rate to an estimate of the X-ray luminosity assuming an APEC plasma model (Smith et al., 2001) of fixed metallicity (0.4 solar) and temperature (5 keV). The conversion factor to luminosity is then derived for the Galactic neutral hydrogen column density and redshift of each source, given the ROSAT instrumental response. We adopt the same luminosity-mass scaling relation as in the previous MCMF work over the DES footprint,

$$L_{500,0.5-2.0\ \mathrm{keV}} = A_\mathrm{X} \left(\frac{M_{500}}{M_\mathrm{piv}}\right)^{B_\mathrm{X}} \left(\frac{E(z)}{E(z_\mathrm{piv})}\right)^2 \left(\frac{1+z}{1+z_\mathrm{piv}}\right)^{\gamma_\mathrm{X}}, \tag{4.2}$$

where A_X, B_X and γ_X have best values of 4.15×10^{44} erg s^{-1}, 1.91 and 0.252 respectively. The redshift pivot is 0.45 and the mass pivot is given as $6.35 \times 10^{14} M_\odot$ (Bulbul et al., 2019) .

The galaxies contributing to the richness measurement λ are not only limited to be within r_{500} but also within a certain luminosity range. Compared to our previous work we expand this range to be $m^* - 3$ to $m^* + 2$, where the characteristic magnitude m^* is the same as in our previous work and is based on a star formation model with an exponentially decaying starburst at a redshift z = 3 with a Chabrier initial mass function and a decay time of 0.4 Gyr (Bruzual & Charlot, 2003). We therefore now consider galaxies 0.75 mag fainter than before, which leads to a $\sim$ 1.5 times increase in λ compared to runs with the previous RS galaxy luminosity cut. We trace the local imaging depth at the location of a cluster candidate and account for missing sources in cases where the local imaging depth does not reach $m^* + 2$ (for further details see Sec. 3.5 in Klein et al. (2019)).

The distribution of richness given redshift is then searched for peaks and then fitted by so-called peak profiles, which are derived from stacks of clusters with spectroscopic redshifts (see Klein et al., 2019, 2022, for examples). With this approach we directly perform a calibration against spectroscopic redshifts and include other effects such as the contribution of blue cluster members or the evolution of the used aperture (r_{500}) as function of redshift into the profile shape.

4.3.1.1 Identifying and removing the contamination

Separating non-cluster sources from real X-ray clusters requires that we estimate the probability that a source with measured redshift and richness is a contaminating source. A contaminating source would be an AGN, star or 2RXS noise fluctuation that happens to lie along the line of sight toward a physically unassociated optical system. With the MCMF algorithm, we use the differences between the richness distributions as a function of redshift $f(\lambda, z_i)$ toward contaminants and toward galaxy clusters, to then assign an estimate that each matched pair is a random superposition of X-ray and optical source rather than a true cluster. The cluster candidates with the highest probability of being contamination are excluded from the catalog.

To estimate the richness distribution of the contaminants $f(\lambda, z_i)$ one wants to apply MCMF to a catalog that represents as much as possible the characteristics of these contaminants. In the case of 2RXS the vast majority of the sample ($\sim$ 85%) is either AGN, star or noise fluctuation, and we therefore make use of the 2RXS catalog itself. In previous work on the MARD-Y3 sample (Klein et al., 2019), randomly redistributed 2RXS positions were used to characterize the richness

distribution of the contaminants. In fact, it is important to trace possible changes in the properties of the contaminants as a function of survey characteristics such as exposure time or location on the sky or with respect to the galactic plane. Therefore, in the current analysis we systematically shift the 2RXS source positions along ecliptic latitude, because this follows the scan direction of the RASS survey. We create four shifted versions of the 2RXS catalog shifted by plus and minus one and two degrees. This largely preserves characteristics such as the exposure time distribution, source density and flux distributions. We apply MCMF to those shifted catalogs after removing any shifted locations that by chance correspond to the locations of real 2RXS sources in positional and redshift space, because the goal is to measure the richness distributions of non-2RXS selected sources and in particular to avoid biasing of these richness distributions by galaxy clusters. We refer to the resulting richness distribution as $f_{\rm rand}(\lambda, z_i)$.

Similarly, we measure the richness distribution of the 2RXS candidates $f_{\rm obs}(\lambda, z_i)$. With the set of richness distributions from the 2RXS catalog and the shifted catalogs we calculate for each candidate i the contamination estimator $f_{\rm cont,i}$, which is defined as

$$f_{{\rm cont},i} = \frac{\int_{\lambda_i}^{\infty} f_{\rm rand}(\lambda, z_i) d\lambda}{\int_{\lambda_i}^{\infty} f_{\rm obs}(\lambda, z_i) d\lambda}, \tag{4.3}$$

where z_i and λ_i are the redshift and richness of the cluster candidate, $f_{\rm obs}(\lambda, z_i)$ is the richness distribution of 2RXS candidates and $f_{\rm rand,z}(\lambda, z_i)$ is the richness distribution extracted from the shifted catalogs. The integrands are evaluated at the candidate redshift z_i and the integral is carried out from the candidate richness λ_i.

The first step we take in defining a cluster sample is to adopt an $f_{\rm cont}$ threshold $f_{\rm cont}^{\rm cut}$. Such a sample has an associated $\lambda_{\rm min}(z)$ that marks the minimal richness a cluster at redshift z must have to be included in the sample. For cluster samples that extend to very low mass ($\sim 10^{14} M_{\odot}$), this $f_{\rm cont}$ selection introduces incompleteness, which can be accounted for using the λ-mass relation describing the sample.

By design, a cluster sample selected with $f_{{\rm cont},i} < f_{\rm cont}^{\rm cut}$ has its contamination reduced by the factor $f_{\rm cont}^{\rm cut}$ as long as the shifted catalogs (i.e., random positions) produce richness distributions $f_{\rm rand}(\lambda, z)$ that are representative of the contaminating sources. In this work we may expect some small difference between the richness distributions $f_{\rm rand}(\lambda, z)$ we use and that of the contaminants. Only $\sim$ 35% of the contaminants are expected to be noise fluctuations and another $\sim$ 15% are stellar X-ray sources. Neither of these source types are correlated in any way with the passive galaxy population in the Universe. However, the remaining ($\sim$ 65%) of the contaminants are X-ray AGN, which are hosted by galaxies and therefore trace the large scale structure as do the cluster passive galaxies. Therefore, in Section 4.4.1 we carry out a validation by measuring the contamination of the catalog as function of $f_{\rm cont}^{\rm cut}$.

4.3.1.2 Constraining the impact of different band combinations

As described in Section 4.2.2, the Legacy Survey can be broadly divided into three parts, Legacy Survey North, Legacy Survey South grz and Legacy Survey South gi. Using overlapping regions between the north and south parts, we calibrate the galaxy photometry of the north part to the

system in the south. We further create individual peak profiles for each of the three MCMF runs based on spectroscopic clusters in those subsamples.

The photometric redshift and richness measurements for overlapping sources of the grz runs of the North and South surveys are in decent agreement with scatter in redshift of $\sigma_{z/(1+z)} = 0.003$ and $\sigma_{\ln(\lambda_{\rm North}/\lambda_{\rm South})} = 0.12$. We further compared $f_{\rm cont}$ measurements between both survey patches. We also found agreement too, which means that also the distribution of the richness estimates around randoms is comparable. We therefore merge both of the grz based MCMF runs to one big catalog and calculate $f_{\rm cont}$ from the merged 2RXS and shifted catalogs.

We additionally did an MCMF run using the full set of $griz$ bands. The performance improvements in photometric redshifts were minor, not justifying yet another subdivision of the survey.

The sole reason for running MCMF in gi-mode is to increase the footprint solid angle, in particular to fill a $\sim$1000 deg^2 hole in the eastern galactic hemisphere with low rz-band coverage. There is a large overlap between the MCMF run in grz-mode and that using the gi-mode.

In Figure 4.2 we show the redshift and richness comparison between the two runs. We do not find any significant redshift bias between both systems and an outlier fraction of 0.2% (0.8%), for candidates with $\lambda/\sigma_\lambda > 3$ and defining offsets of $\Delta z/(1+z) > 0.1(0.05)$ as outliers. We find a richness trend of $\lambda_{grzw1w2} = 1.1 * \lambda_{giw1w2}$ that we correct before we include sources from gi-mode into the main sample.

Given that the sky area and the number of clusters identified in the gi-mode run are small, and given that the gi-mode performance is similar to that of the grz-mode, we adopt the grz-mode mapping of richness and redshift to $f_{\rm cont}$. This means that for a given $f_{\rm cont}$ based selection, the same redshift dependent richness cut applies to all survey regions, independent of whether it lies in the Legacy Survey North, the South or the region with the gi-mode measurements. Note that we provide footprint related flags in the catalog to allow users to compare different parts of the survey.

4.3.2 Rejecting multiple detections of the same cluster

One known problem with the 2RXS catalog and other X-ray catalogs based on the same detection pipeline is that bright, extended X-ray sources are multiply detected and listed in the catalog. 2RXS provides an optical screening flag (S_FLAG; Boller et al., 2016) to partially address this problem, but we have found that this flagging is neither complete nor is it always correct. We therefore make use of the optical cluster centres to identify multiple detections, which we then remove from the catalog. Important in this process is to tune the removal so that resolved galaxy cluster mergers and pairs remain in the catalog.

As described in previous MCMF analyses (Klein et al., 2018, 2019, 2022), each cluster has two possible centres: 1) the BCG position and 2) the position of the peak of the galaxy density map. The galaxy density map is created using the same color weights as for the richness measurements, and therefore represents the density of red sequence like galaxies at the cluster redshift. The density peaks are searched over the full size of the map ($\sim$ 5 Mpc) and the nearest peak is recorded as the galaxy density based centre. The BCG is identified as the brightest galaxy within 1.5 Mpc, 3σ from the mean red sequence color and brighter than $m^* - 1$. If no source is

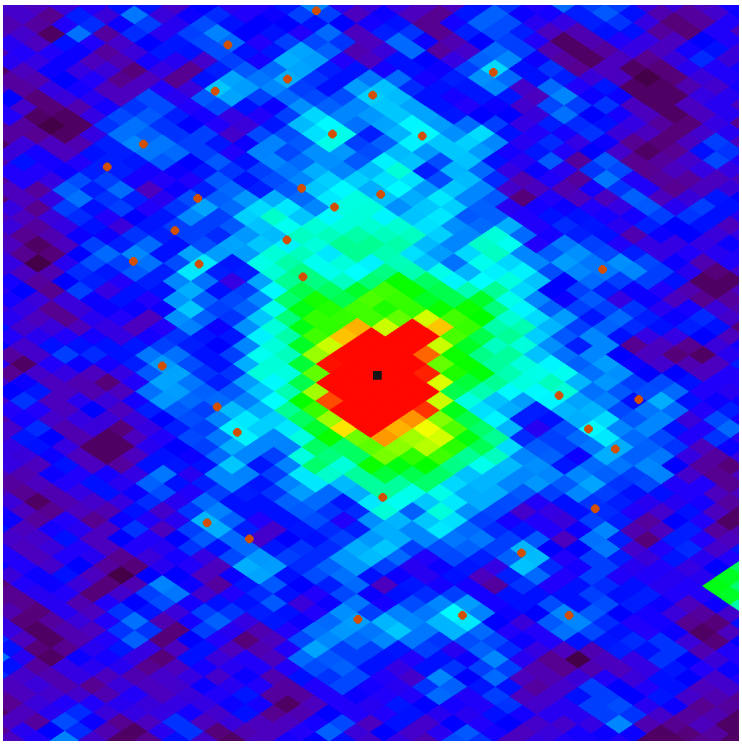

Figure 4.3: Smoothed RASS X-ray count rate map of a $40' \times 40'$ region centered on the Centaurus cluster (ACO 3526). There are 36 2RXS sources in this field. Only one source (black square) survives the $f_{\rm cont}$ selection and the rejection of multiple detections of the same source. Orange circles mark those sources identified as multiple detections.

found we expand the search to 4σ and $m^* - 0.5$. If still no source is found, we consider the BCG search to be unsuccessful. We consider the galaxy density centre to be the more robust estimate, while the BCG position is the more accurate, if correctly identified. We consider the BCG to be correctly identified if it agrees within 250 kpc with either the X-ray position or the galaxy density centre. In 81% of the cases we consider the BCG to be correctly identified, for the remaining cases we use the galaxy density centre.

To flag multiple detections of the same cluster we first flag all 2RXS sources where the optical centre has a nearer match to another 2RXS source whose redshift differs by $\Delta z \leq 0.05$. We then run through the list of flagged sources and check if this criterion results in flagging all sources of a given system. We then include in the catalog the system with highest count rate and exclude the others. An example for multiple detections and their rejection is shown in Figure 4.3. Finally, we merge the measured count rates of all multiple detections closer than 1 Mpc that are not flagged as likely point-like sources from the dedicated point-like follow-up discussed in the following section.

4.3.3 Identifying residual point source contamination

The high-purity cluster samples created from an MCMF run still have residual point source contamination– these are essentially all random superpositions of X-ray point sources with red sequence optical systems. The unresolved X-ray sources have been studied extensively to identify the 2RXS AGN and stars and to assign the best associated optical or infrared counterparts. In the following subsections we summarize this work, because we use the results to devise a method for removing many of the contaminating AGN and stars that initially make it into our cluster sample.

4.3.3.1 2RXS detected Stars

Young and fast rotating stars show a high ratio of X-ray to bolometric luminosity of up to $L_X/L_{bol} = 10^{-3}$, which seems to be a saturation limit (Vilhu, 1984; Wright et al., 2011). For slow rotating old stars the ratio of luminosities can be as low as $L_X/L_{bol} = 10^{-8}$ (Güdel, 2004; Testa et al., 2015). Given the 2RXS flux limit, these luminosity ratios imply that 2RXS stellar sources should be brighter than 15 mag and within a distances of ~750 kpc (Freund et al., 2022); therefore, they should be included in the GAIA catalogs (Prusti et al., 2016; Brown et al., 2021). In fact, the limitation in identifying optical counterparts for 2RXS stars is that some of those stars will be too bright for GAIA, making it sensible to augment the GAIA catalog with the Tycho-2 catalog (Høg et al., 2000).

In their analysis of 2RXS stars, Freund et al. (2022) calculated matching probabilities using positional offsets. The quantity p_{ij} gives the probability that a given stellar source is the right counterpart to the 2RXS source and $p_{stellar}$ is the probability that any of the given stellar counterparts is a match to the 2RXS source and therefore the probability that the 2RXS source is a star. Due to the poor angular resolution of RASS, there are often many possible optical counterparts, and thus two further observables– the X–ray to Gaia flux ratio and the stellar distance– were used to help improve the counterpart selection. Because source densities characteristics of true and contaminating sources change with sky position, the whole analysis and calibration is performed independently on multiple patches on the sky, depending on galactic coordinates. With improved selection algorithm, it is possible to create a $\sim$ 93% pure and complete stellar sample using $p_{stellar} > 0.51$ and $p_{ij} > 0.5$ (Freund et al., 2022). The fraction of stars in 2RXS depends on galactic latitude (see Figure 5; Freund et al., 2022), but over the region of interest of our study (excluding galactic plane), the stars constitute only 10 – 15% of the 2RXS sources.

4.3.3.2 2RXS counterparts

A dedicated search for AllWISE (Cutri et al., 2013) counterparts to 2RXS sources (Salvato et al., 2018) focused on the identification of extragalactic point sources, limiting the analysis to $|b| > 15$ and excluding regions of radius 6° and 3° around the Large and Small Magellanic Clouds (similar to the region in Figure 4.1).

To identify counterparts, a Bayesian statistics based algorithm called NWAY was used, that adds priors from counterpart magnitudes and colors to the typical counterpart search based on position and source density (Salvato et al., 2018). To inform the color-magnitude prior, 2,349

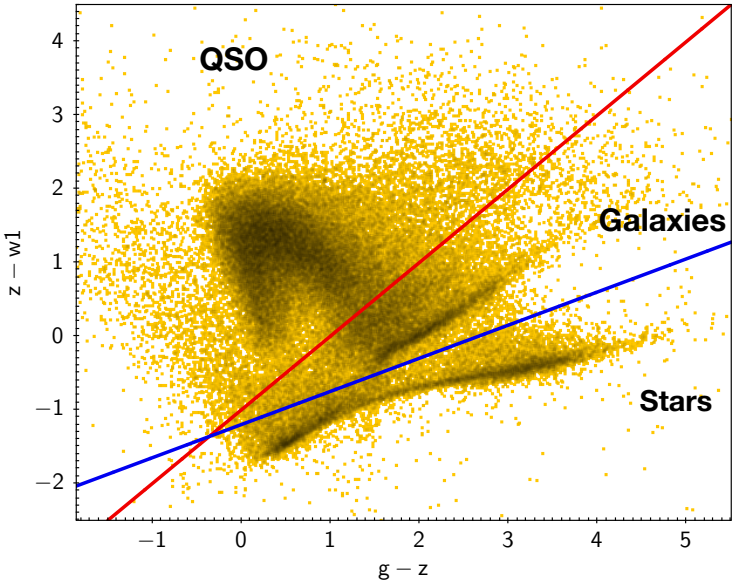

Figure 4.4: Color distribution of NWAY selected counterparts of 2RXS sources. Several high-density regions are visible, which can be separated by two lines. Sources above the red line are predominantly QSO. The high-density region below the blue line belongs to stellar sources. Sources in between the two lines correspond to galaxies, where red sequence galaxies build a dense stripe starting at $g - z \approx 1.5$

secure counterparts from 3XMM-DR5 (Rosen, S. R. et al., 2016) were used. The application of NWAY to 2RXS resulted in at least one ALLWISE candidate for 99.9% of the 2RXS catalog in the previously defined footprint and within the a maximum offset of 2 arcminutes.

Key NWAY measurements used to identify good counterparts are *p_any* and *p_i*. The estimator *p_any* provides the probability that any of the considered ALLWISE counterparts are the correct 2RXS counterpart. The second estimator *p_i* provides the probability of a given ALLWISE candidate to be the correct counterpart. High values of *p_i* therefore indicate clear one-to-one matches between 2RXS and ALLWISE. The analysis found $\sim$ 59% of the 2RXS sources with *p_any* > 0.5, while only 5% of randomized 2RXS positions show *p_any* > 0.5 and *p_i* > 0.8. Given that 2RXS is expected to have 30% spurious sources, the expected contamination by spurious sources after applying this cut to 2RXS is $\leq$ 2%. Based on the same assumption, this cut provides counterparts to 84% of real sources in 2RXS. Considering multiple detections caused by extended sources ($\sim$ 4.5%) the fraction of identified sources increases to 90%. Given the optimization to point-like sources and the presence of clusters with significant positional offsets, this suggests that the completeness of identifying AGN in 2RXS is likely significantly higher than 90% (Salvato et al., 2018).

4.3.3.3 Photometric properties of 2RXS counterparts

Thanks to the much better positional accuracy from the WISE counterpart, compared to 2RXS, we can simply match the best ALLWISE counterparts to the LS DR10 data set using a 1.5 arcsec maximum offset. With that, we have g, r, z and forced WISE $w1, w2$ photometry for the point-like

counterparts in our footprint. As the NWAY catalog contains counterparts to various source types, including the BCG of galaxy clusters, the color information available allows us to split between source types. In Figure 4.4 we show the color distribution of NWAY counterparts in $z - w1$ vs $g - z$ color. Visible are three over dense regions, that can be roughly separated by the two lines drawn in the plot. The over density below the blue line ($z - w1 = (g - z) \times 0.4 - 1.2$) is the stellar locus, which is dominated by stars. The over density above the red line ($z - w1 = (g - z) - 1$) is dominated by AGN, mostly by QSOs. The over density in between the lines is dominated by passive, red-sequence like galaxies. In our subsequent analysis of contamination by non-cluster sources in Section 4.4.1 we make use of the LS DR10 photometry to define a clean AGN subsample.

4.4 RASS-MCMF cluster catalog

There are three key criteria that affect the usefulness of a cluster catalog: sample purity, sample size and the difficulty of modelling the selection function. Typically there is a trade off between sample size and purity, because methods to remove contamination from the sample often also remove some real clusters. In addition, more complicated cleaning methods will likely minimize this loss but may make the sample selection more challenging to model. That can then impact one's ability to do cosmological studies with the sample.

To make an educated choice of the final sample definition, we first characterise the sample purity and estimate the impact of the additional point source rejection step. We then define the multi-component matched filter RASS cluster catalog confirmed with DESI Legacy Surveys (RASS-MCMF) sample in Section 4.4.3, presenting the characteristics of the 90% purity RASS-MCMF catalog together with two subsets of the catalog that have 95% and 99% purity. In Section 4.4.4 we describe the cross-matching required to assign spectroscopic redshifts to over half the RASS-MCMF clusters. Finally, in Section 4.4.5 we discuss some properties of the sample, including mass estimates, redshifts and richnesses.

4.4.1 Measuring catalog contamination

To measure the level of contamination by non-cluster sources (AGN, stars, noise fluctuations) we follow previous MCMF analyses (Klein et al., 2019, 2022) and model the distribution of the contamination and the clusters in an observable-observable space $\log(\lambda/M_X)$. As noted previously, M_X is an X-ray based mass estimate that uses the object X-ray flux and redshift (see equation 4.2).

Figure 4.5 shows the distribution of a clean sample, defined using an $f_{\rm cont}$ threshold as desribed in the following subsection and that of the shifted 2RXS catalogs (i.e., random sky locations within the survey). Real clusters follow the power-law observable-observable relation, which for $\lambda - M_X$ is approximately a relation with slope of one, and exhibit considerable scatter. In contrast, the distribution of the random sky locations in this space lies significantly lower than the clusters. Essentially, the clusters are the 2RXS sources with the highest richnesses. The

density of the contamination in this space peaks well below the clusters, but at a given mass the tails of the cluster and contamination distributions overlap somewhat in richness.

In the top panel of Figure 4.6 we show the distribution in $\log(10^{14}\lambda/M_X)$ for 2RXS sources in Legacy Survey DR10 (black data points), excluding only the multiple detections of the same clusters. Given the relative behavior of clusters and contamination in Figure 4.5, it is clear that clusters prefer higher $\log(10^{14}\lambda/M_X)$.

For a measurement of the contamination fraction of the full candidate list, we model the $\log(10^{14}\lambda/M_X)$ distribution with a contamination model only (green line; described below), limiting the fit region to low-enough $\log(10^{14}\lambda/M_X)$ values that contamination by real clusters is minimal. We do not simultaneously fit for a cluster model, because clusters compose only a fraction of the 2RXS catalog, and the statistics are not adequate to produce a good model.

For the contamination model we consider three different populations: noise fluctuations, AGN and stellar sources. As already mentioned, the AGN are hosted in galaxies that trace the large scale structure, whereas the noise fluctuations and bright stars are uncorrelated with it.

The 2RXS catalog paper provides estimates for the fraction of noise fluctuations as a function of existence likelihood. To create a model for noise fluctuations, we select a subset of the sources along random lines of sight that follows this expected distribution in existence likelihood (see Boller et al., 2016). The X-ray AGN model is based on sources directly selected from 2RXS sources using NWAY selection thresholds as described in Section 4.3.3.2 and a color cut shown in Figure 4.4 that excludes passive galaxies and stars. While the NWAY thresholds ensure a sample of $\sim 98\%$ real sources, the color cut excludes non-AGN like sources from the sample. For the stellar subsample we use the selection described in Section 4.3.3.1, yielding 93% purity and completeness.

We then combine all three models to build a composite contamination model, assuming the constituent fractions are 30%, 43% and 12%, which leaves space for $\sim$ 15% clusters. We do not attempt to fit for the relative contributions of AGN, stars and spurious sources, because all three models of contaminants are similar enough that they cannot be independently constrained with the given data. When restricting the fit to $\log(10^{14}\lambda/M_X) < 0$ (see Figure 4.6 top) we find a contamination of $87 \pm 2\%$, which provides an estimate of the contaminant population and expected number of clusters in the sample. Previous work on RASS data showed that between 16% (Hasinger, G. et al., 2021) and 20.5% (Böhringer et al., 2013) of the RASS sources should be galaxy clusters. Accounting for the fact that $\sim$ 30% of 2RXS siources are noise fluctuations, we obtain the fraction of clusters among the real X-ray sources to be $18.5 \pm 3.0\%$. This is in good agreement with previous estimates and suggests that 2RXS should contain $> 10,000$ galaxy clusters in the extragalactic sky.

To enable a test of the true contamination fraction in $f_{\rm cont}$ selected samples such as RASS-MCMF, we must estimate the level of contamination of a cluster dominated subsample. To do this we first exclude all likely stellar and AGN sources (using the same selection methods described above), and reduce the spurious sources by increasing the existence likelihood cut to 8.08, corresponding to a reduction from 30% to 10% (according to Boller et al., 2016). The stellar and AGN rejection excludes 93-98% of real point sources leaving a sample with only $\sim$15% residual contamination. Creating a cluster sample from this cleaned catalogue using a threshold $f_{\rm cont}$ <0.3 would then further reduce the contamination by a factor of three, creating

a ~95% pure cluster sample. At even smaller $f_{\rm cont}$ thresholds, the contamination in the cluster sample quickly becomes irrelevant.

We use this clean cluster sample to create a cluster model at the high (cluster dominated) end of the $\log(10^{14}\lambda/M_X)$ distribution, which will enable us to estimate the number of real clusters in any subsample selected with a particular threshold $f_{\rm cont}$.

We measure the contamination fraction for multiple $f_{\rm cont}$ thresholds and show them in Figure 4.7. If we would assume the real contamination is well described by the richness distribution extracted using the shifted 2RXS catalogs, we would expect the final contamination of the sample to be the $f_{\rm cont}$ threshold times the initial contamination of the candidate list ($\sim 87 \pm 2\%$). The measurements shown in Figure 4.7 suggest that the actual contamination within $f_{\rm cont}$ selected samples is at the expected level for $f_{\rm cont}$ thresholds between 0.05 and 0.3. In other words, there is no evidence that richness distributions $f_{\rm rand}(\lambda, z)$ derived from the shifted 2RXS catalogs (corresponding therefore to random lines of sight) differ from the true richness distributions of the non-cluster contaminants in the 2RXS catalog.

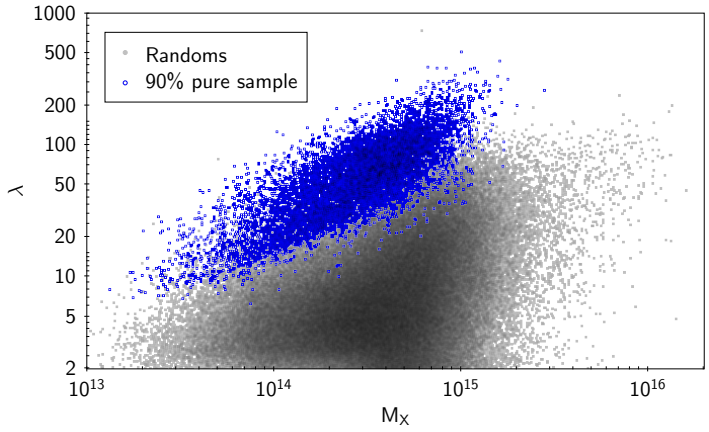

Figure 4.5: Distribution of RASS-MCMF clusters (blue) and non-clusters selected along random lines of sight (gray) in λ versus M_X, where M_X is an X-ray mass estimate that assumes the source is a cluster at the redshift of the best optical counterpart. Real clusters scatter about the λ-M_X relation while non-clusters extracted along random lines of sight are predominantly distributed below that relation.

4.4.2 Additional point source removal

From the cluster fits to the observed $\log_{10}(10^{14}\lambda/M_X)$ distribution, we know the amount of contamination for any given threshold in $f_{\rm cont}$, and we know that contamination predominantly lies at low $\log_{10}(10^{14}\lambda/M_X)$ values (see bottom panel of Figure 4.6). The distribution is dominated by the contamination for $\log_{10}(10^{14}\lambda/M_X) < 1.0$, indicating that excluding identifiable stars and AGN with QSO colors that exhibit low $\log_{10}(10^{14}\lambda/M_X)$ would be effective at reducing sample contamination while having only a minor impact on the real cluster content of the sample.

We test this explicitly by examining the multiband optical images of low $\log_{10}(10^{14}\lambda/M_X)$ sources, finding the first clear cluster cases at $\log_{10}(10^{14}\lambda/M_X) \sim 1.1$. We therefore set

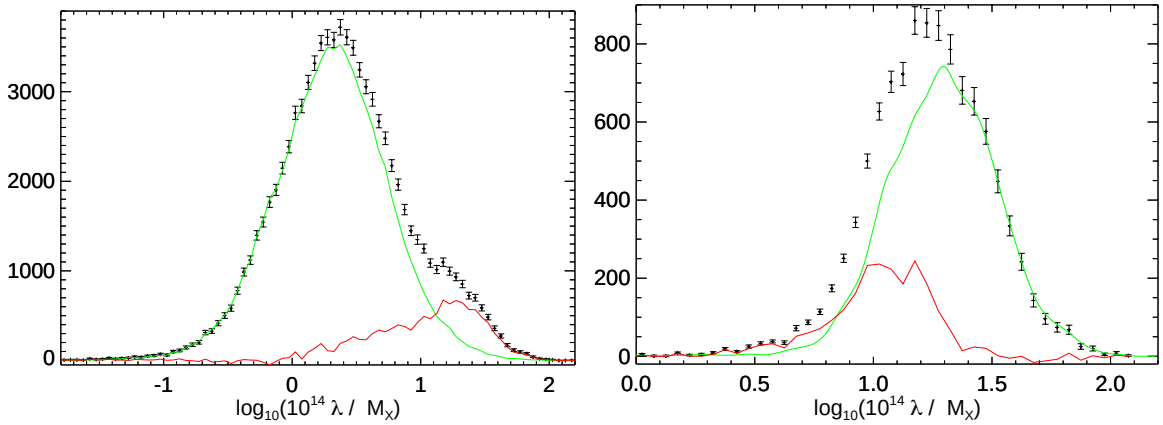

Figure 4.6: Histogram (top) of the richness over X-ray mass ratio in $\log_{10}(10^{14}\lambda/M_X)$ for all 2RXS sources in the optical footprint, excluding multiple detections of the same cluster. Green line show the model of the contaminant population; the red line shows the residual between the contaminant model and total distribution, which is the estimate for the cluster population. Similar plot (bottom) but for the subsample with $f_{\rm cont} < 0.2$. Here a cluster model (green) is fit to the data and the residual (red) is showing the estimated distribution of non-clusters.

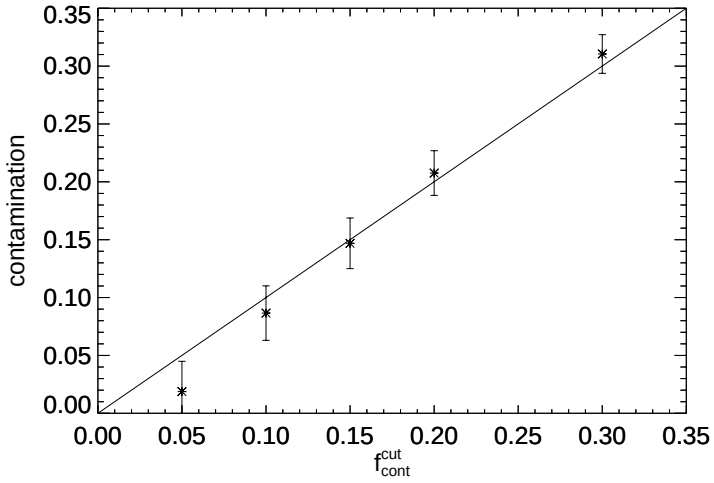

Figure 4.7: Estimated contamination from the cluster model fit to the λ/M_{500} distribution (see Figure 4.6) versus the $f_{\rm cont}$ selection threshold $f^{\rm cut}_{\rm cont}$. The line marks the expected contamination of each subsample, given the initial contamination measured to be 87 ± 2%. The measurements are in good agreement with expectations.

$\log_{10}(10^{14}\lambda/M_X) \approx 1.1$ as the upper limit for point source rejection and exclude all those NWAY identified sources with QSO colors and all Gaia selected stars from the sample in this region.

We emphasize that some sources identified as AGN or stars could be associated with clusters. With the reported purities of the stellar sample (93%) and the NWAY counterparts (98%) (see

discussion of stellar and AGN selection in Section 4.3.3), we estimate that ~95% of the excluded sources are true contamination, with the remainder being clusters. Using this information, we can estimate the fraction of lost clusters coming from direct point source exclusion. For the RASS-MCMF 90%, 95% and 99% purity cluster samples presented below, this introduces an effective cluster incompleteness of 0.7%, 0.3% and 0.2%, respectively.

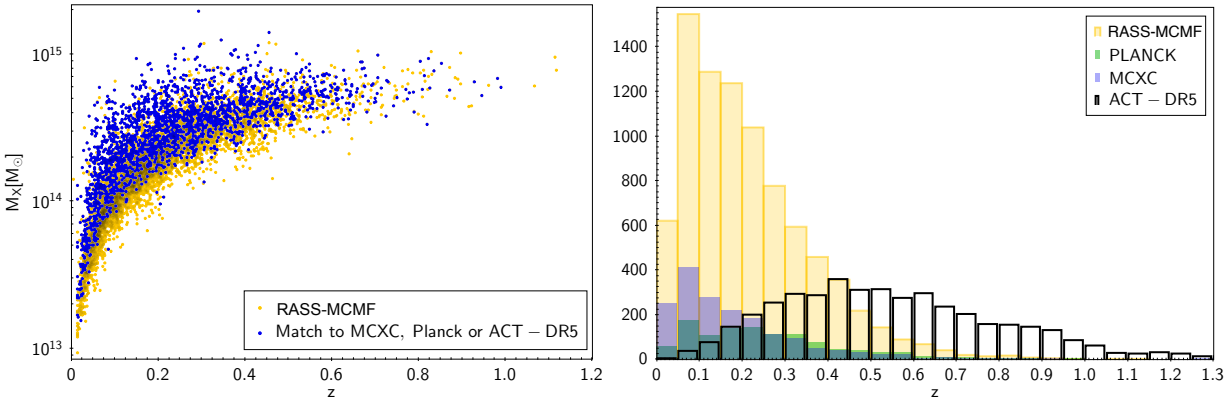

Figure 4.8: Left: X-ray mass proxy M_X versus redshift for the 90% pure RASS-MCMF sample. Highlighted in blue are matches to the MCXC, Planck PSZ2 and the ACT-DR5 cluster catalogs. Right: Redshift distribution of the 90% pure RASS-MCMF sample, as well as for MCXC, Planck PSZ2 and the ACT-DR5. The RASS-MCMF catalog contains more clusters per redshift interval (dN/dz) than ACT-DR5 out to $z \sim 0.4$ and more clusters overall than all three external cluster surveys put together.

4.4.3 Catalog definition

We present here a clean cluster catalog of 90% purity that is built from 2RXS by applying the MCMF algorithm to the Legacy Survey DR10 dataset. The key contamination removal steps include 1) removal of sources with low significance optical counterparts (Section 4.3.1), 2) removal of multiple detections of the same cluster (Section 4.3.2) and 3) an additional point source rejection (Section 4.4.2).

Table 4.1 contains the properties defining the 90% pure sample as well as the 95% and 99% subsamples. From left to right in the table we present the sample purity and the number of clusters in the sample. In addition, we present the $f_{\rm cont}$ selection threshold applied, the number of 2RXS candidates selected, the number of multiple detections rejected and the number of point sources rejected.

We consider point source rejection that impacts the sample selection at the subpercent level as small enough to be ignored in most studies. However we acknowledge the fact that some of these sources might indeed be of astrophysical interest like clusters with strong central AGN emission or star formation like the Phoenix cluster (McDonald et al., 2015). In fact, the Phoenix cluster is likely the most famous source excluded by our point source rejection step, although it almost avoided the point source rejection threshold with $\log(10^{14}\lambda/M_X) = 1.096$. Other known clusters with similar features such as A1835 and Zwicky 3146 (Allen et al., 1992), A2667 (Rizza et al., 1998) or CHIPS1356-3421 (Somboonpanyakul et al., 2021) are retained in our sample. We

Table 4.1: Properties of the three RASS-MCMF galaxy cluster samples. From the left are the sample purity, the final number of galaxy clusters, the $f_{\rm cont}$ selection threshold, the total number of 2RXS sources selected, the number excluded due to being multiple detections of the same source, and the number of sources excluded due to point source rejection.

Sample Purity	Number of Clusters	$f_{\rm cont}$ Selection Threshold	Number of Candidates	Rejected Multiples	Rejected Point Sources
90%	8449	0.17	11585	2092	1044
95%	6912	0.11	9214	1856	446
99%	5506	0.06	7352	1652	194

therefore believe that the RASS-MCMF sample presented here represents an excellent resource for most galaxy cluster studies.

We emphasize that the 99% pure sample has negligible contamination, and the point source rejection step plays a smaller role here than in the 90% and 95% pure samples.

4.4.4 Spectroscopic redshifts

We estimate spectroscopic redshifts for the best optical counterpart identified using MCMF by employing public spectroscopic galaxy redshifts. The galaxy redshifts are drawn from a merged catalog of the SDSS DR17 (Blanton et al., 2017), 2dFGRS (Colless et al., 2001), 6dFGS (Jones et al., 2009), 2MRS (Huchra et al., 2012) and the spectroscopic subset of GLADE+ (Dálya et al., 2022). As a first step, we match this catalog with the positions of successfully estimated BCGs using a maximum positional offset of 2 arcsec. As a second step, we search for all spectroscopic galaxies within 2 Mpc and $|z_{\rm cluster} - z_{\rm spec}| < 0.025(1 + z_{\rm cluster})$. From the selected galaxies, we derive the median redshift and finally derive the cluster redshift using all galaxies within $|\delta z| < 0.015$ from the median redshift. In the case where BCG redshifts exist, we select galaxies within $|\delta|\, z < 0.015$ from the BCG redshift. In our final cluster catalog, we only list spectroscopic redshifts based on at least two members or with a BCG redshift.

In total we provide spectroscopic redshifts for $\sim$ 53% of the RASS-MCMF cluster sample, which reduces to 40% when requiring five or more spectroscopic members. Requiring five or more spectroscopic members, we then find that the scatter between MCMF photo-z and spectroscopic redshift ($\delta z/(1 + z)$) is well described by a Gaussian distribution with standard deviation of $\sigma = 0.0048 \pm 0.0001$. Due to the various depths and redshift ranges covered by the different spectroscopic surveys we employ, the fraction of clusters with spectroscopic redshifts changes significantly over the footprint. In the area covered by SDSS, we are able to provide spectroscopic redshifts for 93.5% of the RASS-MCMF clusters.

4.4.5 Properties of the cluster catalog

With 8,449(6,912 and 5,506) clusters in the RASS-MCMF 90% (95% and 99%) purity sample, this catalog contains the largest ICM-selected cluster sample to date. By covering ~25,000 deg^2

of extragalactic sky, the survey area covers >90% of the sky that is not significantly impacted by high stellar density or high Galactic N_H column density.

In the left panel of Figure 4.8 we show the distribution of the 90% pure RASS-MCMF sample in estimated mass M_X (see equation 4.2) and redshift. As an approximately all-sky survey, the RASS-MCMF sample has overlap with most previous cluster surveys. Restricting to the largest ICM-selected cluster catalog ACT-DR5 (Hilton et al., 2021) and the two largest ICM-selected all-sky catalogs Planck PSZ2 (Planck Collaboration et al., 2016a) and MCXC (Piffaretti et al., 2011), we find ∼2,000 clusters in common with RASS-MCMF; these are marked in blue in Figure 4.8. Each of the three surveys individually contains ∼ 900 clusters in common with RASS-MCMF. As visible by the mass range covered by the blue points with respect to the full sample, the RASS-MCMF sample reaches lower masses than the other surveys out to z ∼0.4. Above this redshift ACT-DR5 clusters probe to lower mass than RASS-MCMF, causing the overlap between the two samples to cover the full dynamic range probed by RASS-MCMF.

In the right panel of Figure 4.8 we show the redshift distribution of the RASS-MCMF sample with respect to the aforementioned samples. The RASS-MCMF redshift distribution peaks at $z \sim 0.1$ and shows a strong decrease in the number of clusters with redshift as expected for an X-ray selected cluster sample. It also shows that RASS-MCMF significantly outnumbers each of the three samples. Thanks to the weak mass dependence of the SZE signature on redshift, ACT-DR5 exceeds RASS-MCMF in the number of clusters per redshift interval (dN/dz) above $z = 0.45$ and can therefore be seen as complementary to the X-ray-based RASS-MCMF sample.

In Figure 4.9 we show the sky distribution of RASS-MCMF clusters in the narrow redshift range $0.05 < z < 0.1$. This redshift slice contains ∼1,500 clusters, almost as many as the entire MCXC or Planck PSZ2 samples. At this redshift range the RASS-MCMF catalog allows one to nicely sample the large scale structure or so-called cosmic web, which in this figure is traced using the 2MASS Photometric Redshift Catalog galaxy density map (Bilicki et al., 2014). For highlighting the advances over previous RASS-based cluster surveys, we also show clusters from the REFLEX (Böhringer et al., 2004) and NORAS (Böhringer et al., 2000) samples as magenta squares.

From our previous work on MARD-Y3 (Klein et al., 2019) and subsequent work on the validation of the selection function of that sample (Grandis et al., 2020), we do expect reasonable scaling of the two mass observables (M_X, λ) of this sample with the underlying true halo mass. This is also illustrated in Figure 4.10, which shows the scaling with masses taken from ACT-DR5, MCXC and Planck PSZ2 for the clusters in common. In the case of Planck clusters, we updated the masses using the correction found in Salvati et al. (2022). This plot highlights the large dynamic range in masses covered by RASS-MCMF, reaching well into the group mass regime at the low redshift end. We remind the reader that the low level of scatter for MCXC masses against M_X shown in Figure 4.10 is very likely due to the fact that both masses are derived from the same RASS data. When comparing the ACT-DR5 datapoints in both panels of Figure 4.10, one can see some indication of data points at $M_{500,\mathrm{public}} \approx 3 \times 10^{14} M_\odot$ scattering to either higher M_X or lower $M_{500,\mathrm{public}}$. Given that we do not see this in scaling with richness, this suggests that M_X might be scattered high at higher redshifts. This might be evidence of Eddington and Malmquist bias coming into play at the low count rate regime. Alternatively the redfshift evolution assumed in the derivation of M_X from count rate might be biased.

This highlights the importance of a dedicated mass calibration of the RASS-MCMF sample including the modeling of the selection function. This has recently been done using HSC-SSP weak lensing in an MCMF-based X-Ray survey (Chiu et al., 2023). RASS-MCMF essentially covers all surveys with dedicated weak gravitational lensing programs, making such a weak lensing mass calibration a natural next step for this sample.

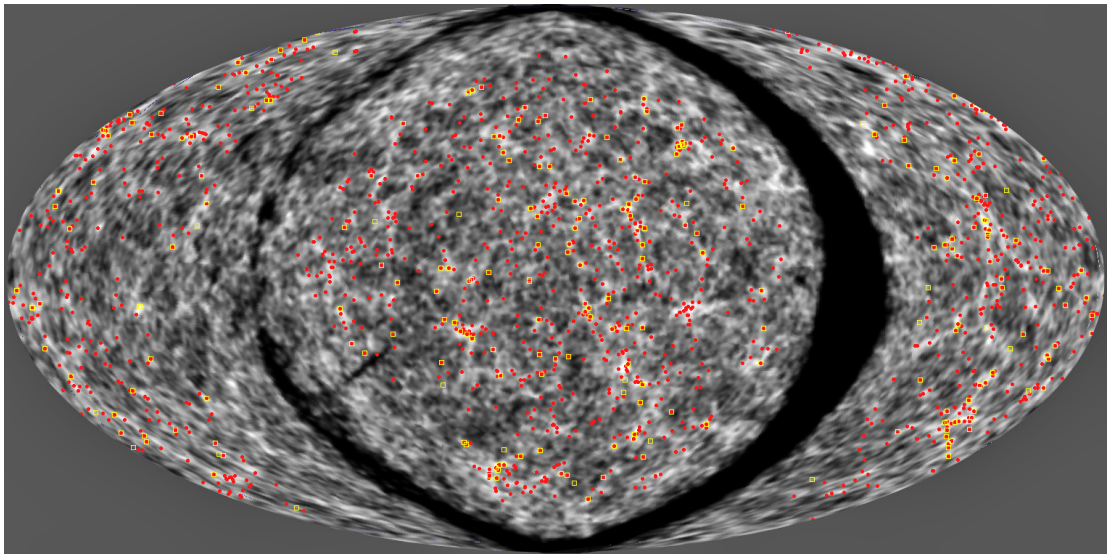

Figure 4.9: Aitoff projection of the galaxy density at $0.07 < z < 0.1$ from the 2MASS Photometric Redshift Catalog (Bilicki et al., 2014) centred on the North Galactic Pole. Red points show a $0.05 < z < 0.1$ redshift slice of the 90% pure RASS-MCMF sample, containing $\sim 1,500$ clusters. For comparison, we show clusters over the same redshift range from the combined ROSAT-based NORAS and REFLEX catalogs (241 clusters) as yellow squares.

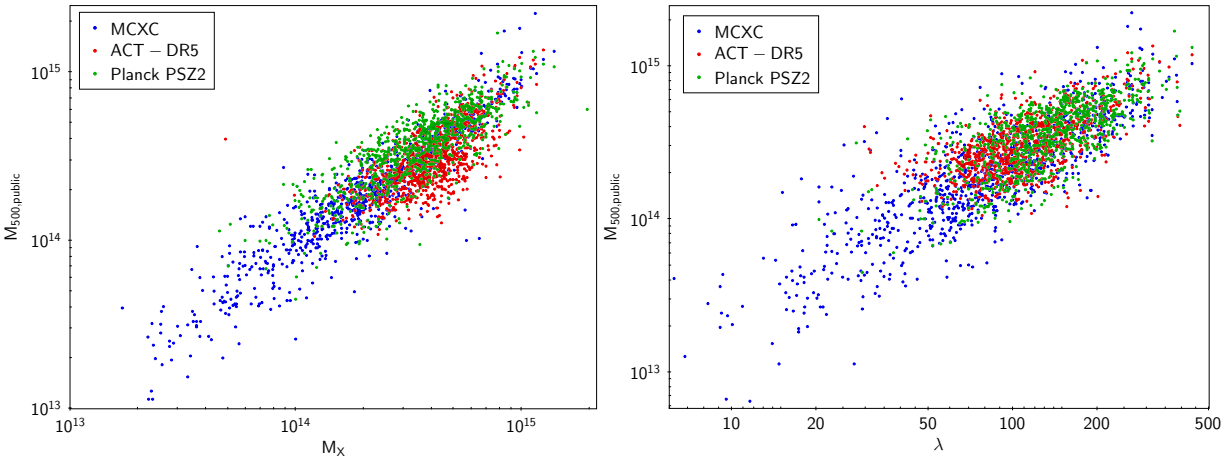

Figure 4.10: Left: Comparison of the mass estimator M_X employed here with mass estimates from ACT-DR5, MCXC and Planck for clusters in the 90% pure RASS-MCMF sample. Right: Same but in comparison to richness.

4.5 Comparison to other cluster catalogs

In this section we investigate the RASS-MCMF catalog properties through comparison to external X-ray or SZE selected samples. These include: 1) the eROSITA Final Equatorial-Depth Survey X-ray catalog (eFEDS), 2) the CODEX X-ray selected sample, 3) the Planck-PSZ2 SZE selected catalog, 4) the MCXC X-ray selected catalog, 5) the ACT-DR5 SZE selected catalog, and 6) the MARD-Y3 X-ray selected catalog.

4.5.1 eFEDS

The extended ROentgen Survey with an Imaging Telescope Array (eROSITA; Predehl et al., 2021) observed during its performance verification phase a ∼140 deg^2 region with an average net exposure time of ∼1.3 ks. This survey was called the eROSITA Final Equatorial-Depth Survey (eFEDS; Brunner et al., 2022). eROSITA and its dedicated all-sky survey (eRASS) are the successors of ROSAT and RASS, showing much higher sensitivity and improved imaging resolution. eFEDS is therefore an ideal testing ground to investigate the expected purity and completeness of the RASS-MCMF sample. The eFEDS source catalog is divided into an unresolved source sample (Brunner et al., 2022) and an extended source sample (Liu et al., 2022), with corresponding optical confirmation presented in follow-on papers (Salvato et al., 2021; Klein et al., 2022; Bulbul et al., 2022). The majority of the RASS-MCMF clusters can be expected to be included in the sample of extended sources, however matches to clusters in the point source sample cannot be excluded because several hundreds of clusters in the eFEDS catalog do not fulfill the requirements to be included in the extend source sample (Bulbul et al., 2022; Chiu et al., 2022).

Of the 90% purity RASS-MCMF sample, only 39 lie within the eFEDS footprint. Using a maximum separation of three arcminutes between 2RXS and eFEDS position we find 32 matches to the sample of eFEDS extended sources, 5 matches to eFEDS unresolved sources, and two sources do not have a match to any eEFDS source. Of the 5 matches to the unresolved sources, two are clearly galaxy clusters and were identified as such in the eFEDS papers. One detection is associated with the unrelaxed cluster eFEDS J085751.6+03103, where 2RXS finds two sources that can be associated with two clumps of galaxies, while the eFEDS detection pipeline classifies the system as one. Another RASS-MCMF source is dominated by a point source, that is likely a cluster member of eFEDS J084544.3-002914, which is located four arcminutes away from the 2RXS position. The remaining source associated with an unresolved eFEDS source and the two sources without a match to any eFEDS sources are likely random superpositions, which serve as contamination within the RASS-MCMF sample. These 3 contaminating sources out of 39 total are in good agreement with the expected 10% contamination of the RASS-MCMF sample.

4.5.2 CODEX

The COnstraining Dark Energy with X-ray cluster survey (CODEX: Finoguenov et al., 2020) is based on the same ROSAT raw data as the 2RXS catalog in our work but uses a different source detection algorithm, namely a wavelet decomposition method (Vikhlinin et al., 1998). A total of 24,788 X-ray sources were found over the ∼10,500 deg^2 of the SDSS BOSS footprint. To identify

optical counterparts for clusters in the X-ray source list, the CODEX team uses the redMaPPer optical cluster finder (Rykoff et al., 2014) run at the X-ray positions using SDSS photometric data. After obtaining redshifts and richnesses of possible optical counterparts, they perform a redshift dependent richness cut that is similar to that used in the MCMF study of 2RXS sources over the DES region (Klein et al., 2019). With this cut they produce what they estimate to be a 95% pure cluster sample. Those X-ray sources making this cut are flagged as clean in the CODEX catalog. The publicly available CODEX catalog contains 10,382 sources of which 2,815 are flagged as clean.

Over the BOSS footprint we find ∼42,000 2RXS sources, 70% more than the CODEX candidate list, which is a reflection of the different X-ray selection techniques used to construct the two candidate lists. Out of the list of CODEX confirmed sources, only 2,033 (72%) have a match to 2RXS within a 3 arcminutes radius. From the matched sources we find ∼ 85% of the sources to be part of the 90% pure RASS-MCMF sample.

We test the purity of the clean CODEX subsample by repeating the test we performed for the 90% purity RASS-MCMF sample on eFEDS (Section 4.5.1). We find 33 CODEX sources in the eFEDS footprint, of which 27 match to eFEDS extended sources, and one of the remaining 6 CODEX sources match to a cluster in the eFEDS unresolved source sample. The remaining CODEX sources typically match to a bright point source in the vicinity of a cluster, but showing offsets greater than 3.5 arcmin. The contamination noted for the clean CODEX sample would translate into an expectation of ∼1.6 false sources, which is small compared to the 5 we find with matches to point sources in eFEDS, suggesting that the purity of the clean CODEX sample is overestimated. The full CODEX sample over the eFEDS footprint contains 123 sources. Only 42 match to extended eFEDS sources and 6 to clusters in the point source catalog. This suggests that ∼60% of the CODEX sources provided in the full CODEX sample are not X-ray selected clusters.

The comparison of total numbers and the redshift distribution over the CODEX footprint is also interesting. The 90% purity RASS-MCMF sample contains ∼4,000 clusters over the CODEX footprint, while the 95% pure RASS-MCMF sample contains ∼3,300. Thus, the 95% purity RASS-MCMF sample provides 17% more sources than CODEX, while the 90% purity sample shows 50% more sources than CODEX. Part of this difference between CODEX and RASS-MCMF arises because the Legacy Survey data are deeper than SDSS, and therefore it is possible to confirm clusters at higher redshift in RASS-MCMF.

By cross matching CODEX clean sources with RASS-MCMF sources lying within 3 arcminutes, we find generally good agreement between the samples in the redshift range $0.1 < z_{\rm CODEX} < 0.5$. We see the known bias at redshifts below $z_{\rm CODEX} = 0.1$ (Clerc et al., 2016) and some trend for redshift underestimation at high redshifts where SDSS imaging depth is reaching its limits for cluster member galaxy detection. We see only ∼1.9% outliers in photometric redshifts for the crossmatched 90% purity sample, and all but 2 outliers (0.1%) show a redshift match between the CODEX redshift and the 2nd or third ranked RASS-MCMF counterpart redshift. This highlights the fact that outliers between both surveys are typically not caused by redshift measurement errors, per se, but by the choice of which optical structure to associate with an X-ray source.

For sources with consistent redshifts we find a significant redshift trend in the ratio of richnesses, suggesting a factor two increase in $\lambda_{\rm CODEX}$ from redshift of z= 0.2 to z=0.55 relative

to the RASS-MCMF richnesses. A similar redshift trend was reported in Ider Chitham et al. (2020), where they remeasured richnesses for CODEX clusters in the SPIDERS subsample (Clerc et al., 2016) using Legacy Survey DR8 data. Using a redshift dependent richness cut from Klein et al. (2019) without accounting for this redshift trend would lead to an increasing contamination of the CODEX clean sample with redshift.

4.5.3 Planck-PSZ2

The Planck PSZ2 (Planck Collaboration et al., 2016a) catalog is the second catalog of SZE selected sources derived from the Planck dataset. It contains 1,653 sources with an expected purity of at least 83%; of these, 1,203 are considered to be optically confirmed. Only 1,261 PSZ2 sources fall within our RASS-MCMF footprint. The median positional uncertainties of the PSZ2 sources is comparably large (≈2.4 arcmin); we therefore adopt a cross-matching radius of 5 arcminutes. We find 881 matches to 2RXS sources, of which 842 have a match to the 90% purity sample and an additional 12 systems make the $f_{\rm cont}$ selection threshold but are rejected as likely AGN or stellar sources. From the remaining 27 PSZ2 sources matched to 2RXS, only 10 systems have a measured redshift and can be considered to be optically confirmed. We visually inspect those systems and find two obvious missed clusters (PSZ2 G047.71-59.47, PSZ2 G046.26-70.47), both located in the patchy *gi*-band part of the footprint. The remaining 8 systems are either at large separation from the 2RXS source (> 300 arcsec) or are potentially low richness systems or perhaps chance superpositions. From repeating the matching with the random catalogs, we estimate a ∼ 15 chance of matching physically unassociated PSZ2 and 2RXS systems. This provides an explanation for the remaining 8 systems not being confirmed.

From the 854 sources that either have a match with the 90% purity sample or make the $f_{\rm cont}$ cut but are then excluded as likely AGN, 761 have a PSZ2 redshift. Only 8 of these sources show redshift offsets of $\Delta z/(1+z) > 0.1$, which are very large by galaxy cluster standards. In all these cases, the RASS-MCMF redshift seem to be the correct redshift to the 2RXS candidate. In three cases the matched PSZ2 cluster corresponds to another cluster and refers to the 2RXS candidate in question. Finally we find five cases where the redshifts listed in PSZ2 likely need to be reconsidered. One noteworthy case of a likely wrong redshift is PSZ2 G181.71-68.65 (ACO 305) with $z_{\rm PSZ2} = 0.1529$, which is likely at $z_{\rm MCMF} = 0.293$. We list the five cases with discrepant redshifts in Appendix 4.4.

We summarise the comparison as follows: we find 842 matches to the 90% purity sample, with only ∼ 5 being redshift outliers where Planck likely lists incorrect redshifts. We find two cases where the local quality of the optical data do not allow us to confirm the clusters.

4.5.4 MCXC

The Meta-Catalog of X-Ray Detected Clusters of Galaxies (MCXC Piffaretti et al., 2011) is a collection of various ROSAT-based cluster surveys, including RASS-based (BCS, CIZA, MACS, NEP, NORAS, REFLEX and SGP) as well as samples obtained from serendipitous detections within pointed ROSAT PSPC observations. Within a distance of 2.5 arcmin we find 985 matches

to the 90% pure RASS-MCMF sample and an additional 45 sources making the $f_{\rm cont}$ cut but marked as possible point sources or multiple detections.

The majority of these systems ($\sim$ 90%) are matches to RASS based surveys. Further, we find 79 matches to 2RXS that do not make the RASS-MCMF selection thresholds. Of those, 27 (33%) are matched to the NEP survey (Henry et al., 2006). Visual inspection of these suggest that the majority ($\sim$ 67%) of these matches are not clusters, while two have an $f_{\rm cont}$ close to the selection threshold. A list of NEP sources likely misclassified as clusters is provided in Appendix 4.5. This further stands in strong contrast the matches to the SGP survey (Cruddace et al., 2002), for which we confirm 48 out of 49 clusters or the BCS sample (Ebeling et al., 1998) for which confirm all matches. From the remaining matches, 41 are RASS based with 37 of those coming from the REFLEX or NORAS surveys, which share flux limit and source identification methods. Visual inspection of these sources using optical and auxiliary X-ray data indicate that $\sim 1/3$ of those matches are either point sources or strongly dominated by point sources. Those sources typically lie at the high luminosity and redshift regime of the non-confirmed systems, where the optical cleaning in our catalog should have a smaller impact on the completeness. Where we do expect our selection methods to lead to the loss of real systems is predominantly at low redshifts and masses. Here the LS photometry and red sequence techniques hit their limit with respect to galaxy size and contrast to non-collapsed structures. This is reflected in the fact that 30% of the missed systems are groups at $z < 0.02$. The median of all unconfirmed systems is $z = 0.055$. In total we find four clusters, where visual inspection, cluster redshift and X-ray luminosity would lead us to expect them to be detected in our sample. Two of them fall into the patchy region where only gi-band data are available. A third case is RXC J0105.0+0201 (z=0.197), which lies projected behind the nearby (z=0.006) galaxy IC1613. It is likely in this case that the $>$ 15 arcmin size of the galaxy resulted in over-subtraction of the background in the Legacy Survey DR10 images and may also have impacted photometric calibration at the location.

Comparing redshifts between MCXC and our 90% pure sample, we find that redshifts agree for 98% of the cases within $\Delta z < 0.05$. Investigating the remaining 21 sources with larger offsets, we find 16 cases where RASS-MCMF redshifts are correct as confirmed by publicly available spectroscopic redshifts. The remaining five show at least two clusters along the line of sight, all indicated with a second strong counterpart with $f_{\rm cont} < 0.1$. In all except one case the primary counterpart listed in RASS-MCMF is the better or an equally good counterpart. In only one case, RXC J1036.6-2731, at z=0.013, the second ranked counterpart from RASS-MCMF seems to be the preferred counterpart.

We summarise the comparison to MCXC as follows, we find ~1000 clusters in common with the 90% purity RASS-MCMF sample. We further find 79 positional matches which do not make our selection cuts. From those, the majority of sources are either point sources, point source dominated clusters or low redshift and low mass systems. The small number of missed massive clusters fall either in the gi-band footprint or are affected by foreground galaxies. For redshifts we find generally good agreement between the catalogs, the majority of the 2% redshift outliers are caused by MCXC listing a foreground source rather than the true cluster at higher redshift.

4.5.5 ACT-DR5

The ACT-DR5 cluster catalog (Hilton et al., 2021) is currently the largest SZE-selected cluster catalog, containing 4,195 clusters over a sky area of 13,211 deg^2. More than 98% of the sources are within the Legacy Survey DR10 footprint. Using a matching radius of 2.5 arcmin, we find 1,074 matches to 2RXS sources, of which 915 appear in the 90% pure RASS-MCMF sample, 23 were excluded as likely point sources, and 136 simply did not make the $f_{\rm cont}$ selection threshold meant to exclude likely random superpositions. Many of the SZE selected clusters are at $z > 0.5$, where ACT probes significantly lower masses than our catalog.

Among the cross-matched cluters, we find good redshift agreement for 98.5% of the cases. From the remaining sources 80% have a match to the second ranked source in our sample, suggesting multiple clusters along the line of sight as the main source for redshift discrepancies. In contrast to our previous exercise on MCXC, ACT-DR5 redshifts tend to lie above our redshifts. This is likely a consequence of the different redshift dependencies of the cluster selection methods. Because of the approximately redshift independent mass threshold in its survey, the ACT-DR5 team prefers the richest system along the line of sight at the position of their SZE candidate. Our RASS-MCMF selection depends on richness and redshift, identifying the highest significance optical counterpart along the line of sight toward the X-ray selected candidate. In principle, it would even be possible for the SZE and X-ray selected systems to be different along a line of sight, in which case the redshifts could be correctly assigned and still not agree.

Of greater scientific interest than the small fraction of redshift outliers is the question what 2RXS selected clusters matched to ACT-DR5 do we miss in our 90% pure RASS-MCMF sample. Looking at the ACT masses and redshifts of those systems unconfirmed systems, we find three ACT clusters that are massive enough that they should be well above our selection thresholds. Of those three, two– ACT-CL J0105.0+0201 and ACT-CL J2248.5-1606– were already found to be missing in comparison to the MCXC catalog. In case of ACT-CL J2248.5-1606 being one of two two clusters lying in the patchy region where only gi-band data are available. The last cluster– ACT-CL J1355.1+0430 ($z = 0.185$)– is a more complicate case. Our analysis finds an additional cluster at $z = 0.81$ with richness $\lambda = 68$ at the ACT position, suggesting this is the correct counterpart compared to the $\lambda = 17$ system at lower redshift. In addition, we find a good QSO point like counterpart for this 2RXS source. Placing the ACT cluster at $z = 0.81$ would likely put the cluster out of reach for 2RXS. The ACT catalog also contains a note about this cluster, indicating this it might be a projected system. The good AGN counterpart further supports that the 2RXS match is indeed a point source instead of a massive cluster. The remaining ACT matches show typically low masses, where scatter in the mass-richness relation could explain why these systems do not make our selection cut.

4.5.6 MARD-Y3

The MARD-Y3 catalog (Klein et al., 2019) is the result of the systematic MCMF follow-up of 2RXS over the DES footprint using the DES-Y3 data set (Abbott et al., 2022). It therefore shares similarities in method, the same X-ray catalog and in part imaging data. Key differences are additional (deeper) imaging data in case of LS DR10 and improved calibration at low redshifts.

Furthermore, the richness distribution for the contaminants are drawn over different areas (DES or full Legacy Survey DR10). The method to merge multiple detections also differs for the two catalogs.

To match the catalogs, we directly match the original 2RXS sources. This avoids ambiguity due to different choices in the rejection of multiple detections of the same cluster. For an $f_{\rm cont}$ selection threshold of 0.17, we find 2,626 (RASS-MCMF) and 2,599 (MARD-Y3) 2RXS sources over the same footprint, but only 81% of them make the selection cut in both surveys. That is, there are ∼500 sources in each catalog that don't appear in the other catalog. The reasons for this are 1) different richness thresholds at a given redshift for the same $f_{\rm cont}$ =0.17 selection and 2) scatter between the cluster richnesses extracted from DES and from LS DR10.

The richness thresholds are different because the average RASS exposure time over the DES footprint is higher than that over the full RASS-MCMF footprint. This leads to an increase in the probability that an optical cluster of a given richness is also a X-ray selected cluster to be higher over DES than on average over the full RASS-MCMF area. Consequently the minimum richness needed at fixed $f_{\rm cont}$ is lower for MARD-Y3 than for RASS-MCMF. In addition, the large solid angle and the improved coverage over SDSS in the RASS-MCMF sample allow for a better calibration of the cluster models as well as better statistics to model the $f_{\rm cont}$ selection toward low redshifts. As a result, RASS-MCMF systems not in MARD-Y3 but with $f_{\rm cont} < 0.17$ are predominantly at $z < 0.2$, while MARD-Y3 sources not in RASS-MCMF exhibit a broad redshift range.

A greater concern would be if there were clusters that have a high enough richness to not be affected by the $f_{\rm cont}$ threshold but nevertheless appear in only one catalog. Looking at the richness scatter and the richness thresholds given $f_{\rm cont}$ we find two MARD-Y3 clusters and 12 RASS-MCMF clusters that fit into this category. We find that both MARD-Y3 clusters lie close to the nearby galaxies NGC 300 (z=0.0005) and IC 1683 and therefore likely suffer from masking in the Legacy Survey. Similarly ten out of the twelve sources in RASS-MCMF show a lack of data in DES-DR3 at the cluster position. One of the remaining clusters is MACSJ0257.6-2209 (Ebeling et al., 2001), that was discussed as a special case in the MARD-Y3 catalog paper (Klein et al., 2019) as a rare case where the local DES photometry was impacted by an error in the photometric zeropoint estimation. For sources making the $f_{\rm cont}$ threshold in both samples we find 98.5% of the cases have consistent redshifts, and that the remaining 1.5% do have a match with a lower ranked optical counterpart in the other survey.

We summarise the comparison with MARD-Y3 as follows: for the same underlying 2RXS source and the same $f_{\rm cont}$ threshold we find 81% overlap between both samples. Sources appearing in just one of the samples can be explained by the different selections in richness as function of redshift and by scatter between the two richness measurements.

4.6 Cosmological forecast for RASS-MCMF×DES

Large, well understood cluster samples have been pursued over the last two decades largely because of the information they contain about the underlying physical processes responsible for the cosmic acceleration (Haiman et al., 2001). The RASS-MCMF sample together with the recent

developments toward an accurate understanding of the HMF over a broad range of cosmologies (e.g., Bocquet et al., 2020) create a situation where all requirements for precise and accurate cluster abundance cosmological studies are met (see discussion in Mohr, 2005).

To test the potential constraining power of the RASS-MCMF cluster sample when combined with DES weak lensing, we apply a cluster cosmology analysis code to representative mock datasets. The cosmology analysis code has been developed for the study of South Pole Telescope SZE selected clusters (Bocquet et al., 2015, 2019, in prep) in combination with gravitational weak lensing data. It is written as a CosmoSIS module (Zuntz et al., 2015).

To enable this RASS-MCMFxDES forecast, we have extended this code to work on X-ray selected cluster samples extracted from all-sky X-ray surveys like those from ROSAT and eROSITA. In the following subsections we review the analysis method and the mock observations and then present the parameter constraint forecasts.

4.6.1 Cluster cosmology analysis method

The likelihood is estimated as a multi-variate Poisson likelihood using the expected number of clusters given the observables. The observables to consider for each cluster are the X-ray count rate $\hat{\eta}$, the richness $\hat{\lambda}$ and the redshift z. Thus, our likelihood is closely related to that of Chiu et al. (2023):

$$\ln \mathcal{L}(\mathbf{p}) = \left[\sum_i \ln \left(C(\hat{\eta}, z) \frac{\mathrm{d}N(\hat{\eta}, \hat{\lambda}, z|\mathbf{p})}{\mathrm{d}\hat{\eta}\mathrm{d}\hat{\lambda}\mathrm{d}z} \right) \Bigg|_{i-\mathrm{thcluster}} - \int_{z_{\min}}^{z_{\max}} \mathrm{d}z \int_{\hat{\eta}_{\min}}^{\hat{\eta}_{\max}} \mathrm{d}\hat{\eta} \int_{\hat{\lambda}_{\min}(z)}^{\infty} \mathrm{d}\hat{\lambda}\, C_{\mathrm{HOF}}(\hat{\eta}, z) \frac{\mathrm{d}N(\hat{\eta}, \hat{\lambda}, z|\mathbf{p})}{\mathrm{d}\hat{\eta}\mathrm{d}\hat{\lambda}\mathrm{d}z} \right], \tag{4.4}$$

where $C(\hat{\eta}, z)$ and $C_{\mathrm{HOF}}(\hat{\eta}, z)$ represent the completeness function for each cluster i and for the ensemble, respectively (see Section 4.6.1.2), $\hat{\lambda}_{\min}(z)$ is the minimum observed richness of a cluster at redshift z (derived from the f_{cont} threshold) and $\frac{\mathrm{d}N(\hat{\eta},\hat{\lambda},z|\mathbf{p})}{\mathrm{d}\hat{\eta}\mathrm{d}\hat{\lambda}\mathrm{d}z}$ is the halo-observable function or HOF.

We note two differences with respect to the likelihood from Chiu et al. (2023): (1) our selection function includes a maximum count rate, $\hat{\eta}_{\max}$, and (2) we use only the likelihood of the number counts of our mock RASS-MCMF clusters, excluding the "mass calibration" likelihood (right-most term in their Eq. 6). To include mass information from the DES shear (Gatti et al., 2021) and photo-z (Myles et al., 2021) based weak lensing mass calibration, we adopt priors on the observable mass scaling relation parameters that come from posteriors derived from a separate, ongoing MARD-Y3×DES analysis (Singh et al. in prep.). The redshift range we assume for the analysis is $z_{\min} = 0.01$ and $z_{\max} = 1.1$, with the maximum count rate fixed at $\hat{\eta}_{\max} = 13$.

Similarly to Chiu et al. (2023), the HOF is calculated from the halo-mass function (HMF; Tinker et al., 2008), using the observable mass relations for the intrinsic richness λ and count rate η. Appropriate convolutions are carried out to model the intrinsic and sampling or measurement scatter of the two observables.

4.6.1.1 Observable mass relations

The underlying richness observable mass relation $\lambda - M - z$ has the form

$$\langle \ln(\lambda|M,z)\rangle = \ln A_\lambda + B_\lambda \ln\left(\frac{M}{M_{\rm piv}}\right) + C_\lambda \ln\left(\frac{1+z}{1+z_{\rm piv}}\right), \tag{4.5}$$

where $M_{\rm piv} = 1.6\times10^{14}h^{-1}{\rm M}_\odot$ and $z_{\rm piv} = 0.25$ are chosen to reflect the median mass and redshift of our mock cluster catalogue. We adopt a log-normal intrinsic scatter in λ at fixed mass and redshift that is the same for all redshifts and masses of

$$\sigma_{\ln\lambda} = ({\rm Var}[\ln\lambda|M,z])^{\frac{1}{2}}. \tag{4.6}$$

Furthermore, we model the sampling noise on the expectation value of the richness λ for a given mass and redshift as a Poisson distribution in the Gaussian limit. That is,

$$P(\hat{\lambda}|\lambda) = \frac{1}{\sqrt{2}\delta_\lambda}\exp\left(-\frac{(\hat{\lambda}-\lambda)^2}{2\delta_\lambda^2}\right), \tag{4.7}$$

where $P(\hat{\lambda}|\lambda)$ is the distribution of the observed richness $\hat{\lambda}$ given the measurement uncertainty $\delta_\lambda = \lambda^{1/2}$.

The underlying count rate observable mass relation $\eta - M_{500} - z$ has the form (Grandis et al., 2019; Chiu et al., 2022)

$$\begin{aligned}\left\langle \ln\left(\frac{\eta}{\rm counts/sec}\middle| M,z\right)\right\rangle = &\ln A_\eta - 2\ln\left(\frac{D_{\rm L}(z)}{D_{\rm L}(z_{\rm piv})}\right)\\ &+\left[B_\eta + \delta_\eta \ln\left(\frac{1+z}{1+z_{\rm piv}}\right)\right]\ln\left(\frac{M}{M_{\rm piv}}\right)\\ &+\gamma_\eta \ln\left(\frac{1+z}{1+z_{\rm piv}}\right) + 2\ln\left(\frac{E(z)}{E(z_{\rm piv})}\right),\end{aligned} \tag{4.8}$$

where $D_{\rm L}$ is the luminosity distance, $z_{\rm piv}$ and $M_{\rm piv}$ are the pivot redshift and mass (same as for the $\lambda - M - z$ scaling relation), and $E(z) = H(z)/H_0$ is the expansion history of the Universe. The form of this observable mass relation allows for the mass trend to evolve with redshift. It deviates from that presented in Chiu et al. (2023) through the missing eROSITA based bias factor b that we discuss further in Section 4.6.2.

As with the richness, we adopt a log-normal intrinsic scatter that is the same at all redshifts and masses

$$\sigma_{\ln\eta} = ({\rm Var}[\ln\eta|M,z])^{\frac{1}{2}}. \tag{4.9}$$

Similarly, we model the η measurement uncertainty as a Poisson probability with a mean corresponding to the expected number of photons $n_\gamma = \eta * t_{\rm exp}$, which corresponds to the expected count rate times the exposure time. Explicitly,

$$P(\hat{\eta}|\eta, t_{\rm exp}) = \frac{1}{t_{\rm exp}}\frac{n_\gamma^{\hat{n}_\gamma} e^{-n_\gamma}}{\hat{n}_\gamma!}, \tag{4.10}$$

where $n_\gamma = \eta t_{\rm exp}$ and $\hat{n}_\gamma = \hat{\eta} t_{\rm exp}$.

Because the third observable, the cluster redshift z has a high accuracy and precision ($\sigma(\delta z/(1+z) \sim 0.005$; see Figure 4.2 and associated discussion in Section 4.4.4) we do not model the redshift measurement uncertainty.

As can be seen in Fig. 4.1, the exposure time varies across the sky from 100 s to over 3,000 s, and thus the transformation from η to counts or photons also varies and impacts the Poisson measurement noise, which is needed to calculate the HOF. To account for this, we build a sequence of HOFs for different values of the exposure time, accounting for the survey solid angle at an exposure time by using the exposure time distribution of the RASS×LS-DR10 sky. Given an exposure time, particular values of the HOF $\frac{{\rm d}N(\hat{\eta},\hat{\lambda},z|\mathbf{p})}{{\rm d}\hat{\eta}{\rm d}\hat{\lambda}{\rm d}z}$ are extracted from the sequence using interpolation.

Explicitly, for each cluster the exposure time $t_{\rm exp}$ at the cluster sky location is employed in extracting the appropriate value of the HOF. The last term in the cluster counts likelihood (equation 4.4) is evaluated using the sum of the sequence of exposure-time dependent HOFs.

4.6.1.2 X-ray completeness function

The X-ray completeness function appears in our cluster abundance likelihood, because the X-ray selection on the 2RXS catalog is made using the existence likelihood EXI_ML, while the primary observable mass relation is modeled using the count rate, which is directly related to the cluster X-ray flux and, given the redshift and cosmological parameters, also the cluster X-ray luminosity. The existence likelihood of a cluster depends strongly on its flux or count rate, but it also has important dependencies on the ICM distribution (angular size and morphology of X-ray emission) as well as survey parameters such as the background and exposure time.

Conveniently, the 2RXS sample has measured count rates $\hat{\eta}$ and existence likelihoods EXI_ML for all objects, and the sample is drawn from the full range of exposure times and associated backgrounds. Previous X-ray cosmological analyses have used measured observables from the catalog to empirically defined the relationship between the count rate and the existence likelihood (Klein et al., 2019) or extent likelihood (Chiu et al., 2023) with good success, and therefore we proceed along this route for this forecast.

We model the $\hat{\eta}$-EXI_ML scaling relation needed for the completeness model therefore as follows

$$\begin{aligned}\left\langle \ln(\hat{\eta}|{\rm EXI\_ML}, t_{\rm exp}, {\rm bkg})\right\rangle =& A_{\rm C} + B_{\rm C}\ln {\rm EXI\_ML} \\ &+ C_{\rm C}\ln t_{\rm exp} + D_{\rm C}\ln {\rm Bkg},\end{aligned} \tag{4.11}$$

where $t_{\rm exp}$ is the exposure time and Bkg is the count rate of the background Bkg = Background$/t_{\rm exp}$. This relation is modeled using a log-normal intrinsic scatter in $\hat{\eta}$ of

$$s_{\ln\hat{\eta}} = ({\rm Var}[\ln\hat{\eta}|{\rm EXI\_ML}, t_{\rm exp}, {\rm Bkg}]^{\frac{1}{2}}) \tag{4.12}$$

We extract the best-fit values using the RASS-MCMF clusters with a value of EXI_ML close

to the threshold value of 6.5. We find

$$\begin{aligned} A_{\rm C} &= 1.356 \pm 0.043 \\ B_{\rm C} &= 0.755 \pm 0.004 \\ C_{\rm C} &= -0.624 \pm 0.004 \\ D_{\rm C} &= 0.387 \pm 0.008 \\ s_{\ln\hat{\eta}} &= 0.1819 \pm 0.0021. \end{aligned} \tag{4.13}$$

With this relationship between $\hat{\eta}$ and EXI_ML, a selection threshold in the observed existence likelihood ln EXI_ML introduces a selection in the observed count rate $\ln\hat{\eta}$ that is an error function. Therefore, we model the completeness function $C(\hat{\eta}, z)$ as

$$C(\hat{\eta}, z) = \frac{1}{2}\left(1 + \mathrm{erf}\left(\frac{\ln\hat{\eta} - (\ln\hat{\eta}_{50} + \delta\ln\hat{\eta}_{50})}{\sqrt{2}s_{\ln\hat{\eta}}}\right)\right), \tag{4.14}$$

where erf() is the error function with the scaling factor $s_{\ln\hat{\eta}}$, where $\hat{\eta}_{50}$ is the count rate which has 50% completeness (EXI_ML = 6.5), and $\delta\ln\hat{\eta}_{50}$ is a parameter that allows one to model deviations from this expected threshold count rate during the cosmological analysis. Similarly, during the cosmological analysis we fit for the parameter $s_{\ln\hat{\eta}}$ scaling parameter. Depending on the priors adopted on these two parameters during the cosmological analysis, it is possible to self-calibrate the completeness function (see discussion in Chiu et al., 2023).

As already noted, Equation (4.4) contains two different forms of the completeness functions: $C(\hat{\eta}, z)$ and $C_{\rm HOF}(\hat{\eta}, z)$. The former is the completeness function appropriate for a particular cluster with a given observed count rate, exposure time and background. The latter $C_{\rm HOF}(\eta, z)$ is the completeness of the full HOF used in the last term of the likelihood (see equation 4.4). This function is constructed first for each member of the sequence of HOFs created for different exposure time ranges (each corresponding to an equal width in $\Delta\log(t_{\rm exp})$). We adopt the mean exposure time $t_{\rm exp}$ for each member of this sequence, because the exposure time width of each bin is sufficiently small. To include the background Bkg dependence, we estimate a Bkg-weighted average completeness function $C_{{\rm HOF}_i}(\hat{\eta}, z)$ for exposure time bin i

$$C_{{\rm HOF}_i}(\hat{\eta}, z) = \frac{\sum_j p_j C_{{\rm HOF}_{i,j}}(\hat{\eta}, z)}{\sum_j p_j} \tag{4.15}$$

where $C_{{\rm HOF}_{i,j}}$ is the HOF of the i-th $t_{\rm exp}$ bin and the j-th Bkg bin. The factor p_j is the weight of the j-th Bkg bin.

4.6.2 Creating a RASS-MCMF mock catalog

We create mock RASS-MCMF cluster catalogues for use in forecasting the parameter constraints, adopting the Tinker et al. (2008) HMF for the given cosmology and imposing a mass range $10^{12.1}\ {\rm M}_\odot < M_{500} < 10^{15.5}\ {\rm M}_\odot$ and a redshift range $0.01 < z < 1.1$. In addition, we adopt the

Table 4.2: Summary of forecast parameters. First column corresponds to the name of the parameter, second column to the input value for the creation of the mock and third column corresponds to the priors adopted for the cluster cosmology analysis. The last two columns show the posteriors and uncertainties (1σ) for the two cosmologies: λCDM and wCDM.

Param.	Mock Input	Prior	Posterior	
			ΛCDM	wCDM
The $\lambda - M_{500} - z$ scaling relation (Eq. 4.5)				
A_λ	55.5	$\mathcal{N}(55.5, 2.235^2)$	55.9 ± 1.7	54.7 ± 1.6
B_λ	1.0	$\mathcal{N}(1, 0.101^2)$	$1.044^{+0.049}_{-0.057}$	$1.012^{+0.047}_{-0.054}$
C_λ	0.0	$\mathcal{N}(0, 0.395^2)$	-0.11 ± 0.26	-0.28 ± 0.27
$\sigma_{\ln\lambda}$	0.2	$\mathcal{U}(0.1, 0.4)$	0.181 ± 0.039	0.201 ± 0.039
The $\eta - M_{500} - z$ scaling relation (Eq. 4.8)				
A_η	0.19	$\mathcal{N}(0.19, 0.0076^2)$	$0.1949^{+0.0073}_{-0.0062}$	$0.1919^{+0.0077}_{-0.0066}$
B_η	1.9	$\mathcal{N}(1.9, 0.101^2)$	$1.859^{+0.066}_{-0.076}$	$1.874^{+0.062}_{-0.073}$
γ_η	-1.2	$\mathcal{N}(-1.2, 0.395^2)$	-0.98 ± 0.26	$-0.77^{+0.48}_{-0.41}$
$\sigma_{\ln\eta}$	0.332	$\mathcal{N}(0.332, 0.09^2)$	$0.295^{+0.067}_{-0.056}$	$0.298^{+0.062}_{-0.055}$
Completeness function $C(\eta, z)$ (Eq. 4.14)				
$\delta \ln\hat{\eta}_{50}$	0.0	$\mathcal{U}(-0.5, 0.5)$	-0.004 ± 0.017	-0.009 ± 0.017
$s_{\ln\hat{\eta}}$	0.1819	$\mathcal{U}(0.01, 0.46)$	0.193 ± 0.012	0.192 ± 0.012
Cosmology parameters				
Ω_m	0.28	$\mathcal{U}(0.15, 0.4)$	0.287 ± 0.028	$0.272^{+0.024}_{-0.027}$
h	0.7	$\mathcal{N}(0.7, 0.04^2)$	$0.693^{+0.030}_{-0.036}$	$0.707^{+0.028}_{-0.031}$
$\log 10^{10} A_s$	3.001	$\mathcal{U}(1, 4)$	3.00 ± 0.25	2.99 ± 0.26
σ_8	0.78	-	0.776 ± 0.031	0.784 ± 0.033
w	-1.0	$\mathcal{U}(-2, -0.5)$	-	-1.12 ± 0.15

observable–mass scaling relations presented in Section 4.6.1.1. The input values for the scaling relation parameters that we employ in creating the mock cluster catalogues are list in Table 4.2, where the first column contains the parameter name and the second column the mock input value adopted.

The λ-mass relation parameters are taken to be representative of an ongoing SPT× and MARD-Y3×DES cluster weak lensing analysis (Singh et al., in prep). Because the RASS-MCMF richnesses are on average 1.5× higher than the MARD-Y3 richnesses (due to using a larger portion of the cluster luminosity function), we adjust the normalization parameter A_λ appropriately.

For the η-mass relation parameters, we adopt best fit parameters from Chiu et al. (2023) with some changes to reflect the fact that we are working with RASS count rates rather. In particular, we absorb the parameters in the so-called bias function b into the parameters of equation (4.8). Moreover, we adopt a characteristic scale factor between eROSITA and ROSAT count rates for the clusters of 0.117 and use that to adjust the normalization parameter A_η.

The mock sample is an X-ray existence likelihood selected sample just like the RASS-MCMF

sample. Therefore, after transforming the mass and redshift into the observed count rate $\hat{\eta}$ (see discussion in Section 4.6.1.1), we use the $\hat{\eta}$-EXI_ML relation presented in Section 4.6.1.2 to estimate EXI_ML. Doing so requires that we have an exposure time and background value for each cluster. For these we sample the RASS-MCMF portion of the RASS sky by randomly selecting an equal area healpix pixel and using the RASS reported exposure time and background from that pixel. We then impose the EXI_ML threshold value of 6.5 by rejecting any cluster that falls below that limit.

The RASS-MCMF sample is also cleaned of likely random superpositions using an $f_{\rm cont}$ threshold that corresponds to a minimum value in $\hat{\lambda}(z)$. Using the derived values $\hat{\lambda}(z)$ for the RASS-MCMF sample (90% purity) or its subsamples (95% and 99% purity), we then reject any cluster that does not meet this observed richness threshold. We do not add contaminating sources to the mocks.

The process of generating a mock catalog can be summarized as follows. The first step is to construct the HOF from the HMF and observable mass scaling relations and to then integrate over the relevant ranges in observable space to estimate the expected total number of clusters. We then draw a Poisson deviate with this expectation value, and that sets the actual number of clusters in our mock RASS-MCMF sample. This step includes the effects of 1) the distribution of RASS exposure time and background over the RASS-MCMF sky, 2) the impact of the existence likelihood selection and 3) the impact of the $f_{\rm cont}$ optical cleaning.

Thereafter we cycle through the following set of steps creating the members of the mock RASS-MCMF sample: 1) We use the HMF within the specified mass and redshift ranges to randomly draw a cluster with halo mass M_{500} and redshift z. 2) Using the $\eta - M_{500} - z$ scaling relation (equation 4.8), we derive the count rate η and then randomly select a RASS sky cell that has an associated exposure time and background. With that information we predict the observed count rate $\hat{\eta}$. 3) Using the background, exposure time, and observed count rate we predict the existence likelihood (equation 4.11 and impose the RASS-MCMF existence likelihood threshold EXI_ML=6.5). 4) We use the $\lambda - M_{500} - z$ relation to assign a richness and the sampling noise to predict an observed richness $\hat{\lambda}$. Then we impose the appropriate richness cut using the function $\hat{\lambda}_{\rm min}(z)$ that corresponds to the $f_{\rm cont}$ selection for the RASS-MCMF sample we are modeling.

The cosmological parameters used to create the mocks are also listed in Table 4.2. With this approach the total number of clusters $N_{\rm tot}$ from our RASS-MCMF-like 99% purity sample ranges from 4,800 to 4,950, reflecting the Poisson variation on the expected total number of clusters for the survey. This is the sample we employ for the forecasts described below.

4.6.3 Forecast of Parameter Constraints

We explore two different cosmogonies in our analysis: a flat ΛCDM model and a flat wCDM, where the dark energy equation of state parameter w is a free parameter. The cosmological parameters we include are the mean dark matter density Ω_m, the present epoch value of the Hubble parameter H_0, modeled as the dimensionless Hubble parameter h where $H_0 = 100\,h$ km s^{-1} Mpc^{-1}, the linear power spectrum amplitude $\ln(10^{10}\mathrm{A_s})$ and also the dark energy equation of state parameter w. The present epoch amplitude σ_8 of matter density fluctuations on a scale of 8 h^{-1} Mpc is a derived parameter.

The priors and posteriors for both models are shown in Table 4.2. We adopt either flat priors within specified parameter limits $\mathcal{U}(\Theta_{\rm min}, \Theta_{\rm max})$ or Gaussian priors $\mathcal{N}(\langle\Theta\rangle, {\rm Var}(\Theta))$ defined by their mean and variance. For the cosmological parameters, the only Gaussian prior is for h with a mean of 0.7 and a variance 0.04^2, which comfortably spans the recently published values (e.g., Riess et al., 2019; Planck Collaboration et al., 2020b).

The priors adopted for the observable mass relation parameters reflect posteriors derived separately from an independent analysis of the MARD-Y3 and SPT cluster samples in combination with DES weak lensing (Singh et al., in prep.). The posteriors of the observable mass and cosmological parameters for both ΛCDM and wCDM are consistent with the input values of the mocks at a level of $\leq 1\sigma$. Fig. 4.11 shows the 1σ and 2σ contours of $\Omega_{\rm m}$ vs σ_8 and $\Omega_{\rm m}$ vs w for our RASS-MCMF mock cluster catalogue in red, with the top panel showing the results for a flat ΛCDM cosmogony and the bottom panel showing the results for a flat wCDM cosmogony.

The forecast constraining power for a RASS-MCMF-like survey is in red, while a selection of previously published results including the eFEDS cluster survey (yellow; Chiu et al., 2023), the SPT-SZ cluster survey (grey; Bocquet et al., 2019), the *Weighing the Giants* cluster survey (WtG with green lines; Mantz et al., 2015), the *Planck* primary CMB anisotropy using temperature and polarization (`TTTEE+lowE` in purple; Planck Collaboration et al., 2020b) and the DES 3×2-point analysis (blue; Abbott et al., 2022) are shown for comparison. Our results show tighter constraints than state-of-the-art cosmological analyses such as the 3×2-point weak lensing analysis of DES and other cluster analyses such as those from SPT-SZ and WtG.

It is worth noting that in the case of wCDM, the WtG analysis shows very similar constraints to those we forecast for the RASS-MCMF sample. The WtG contours include constraints from the assumption of constant ICM mass fraction with redshift that are impacted by tight priors adopted by the WtG team on the intrinsic evolution of the galaxy cluster ICM fraction.

In the case of *Planck* primary CMB constraints (bottom panel), we show the posteriors from an analysis that includes marginalization over the sum of the neutrino masses. Interestingly, for ΛCDM our results are weaker but quite competitive with *Planck*, whereas the forecast RASS-MCMF posteriors are tighter than *Planck* in wCDM.

4.7 Conclusions

In this analysis we present the RASS-MCMF cluster catalog, which has been created through a systematic search for galaxy clusters in the ROSAT 2RXS catalog (Boller et al., 2016) over the 25,000 deg^2 of the extragalactic sky covered by the Legacy Survey DR10 (Dey et al. in prep). With 8,449 clusters of galaxies, the RASS-MCMF sample is the largest ICM-selected cluster catalog to date.

Our analysis of the richness and X-ray count rates of all 2RXS counterparts indicates that the non-cluster population composes 87 ± 2% of the original 2RXS source list. Therefore, there are $\sim 11,000$ X-ray selected galaxy clusters in the extragalactic region of this analysis. The RASS-MCMF catalog contains $\sim$ 80% of the total detected cluster population in 2RXS. Additional clusters could be extracted from the 2RXS catalog using the same method, but at the cost of increasing the contamination level of the final cluster catalog.

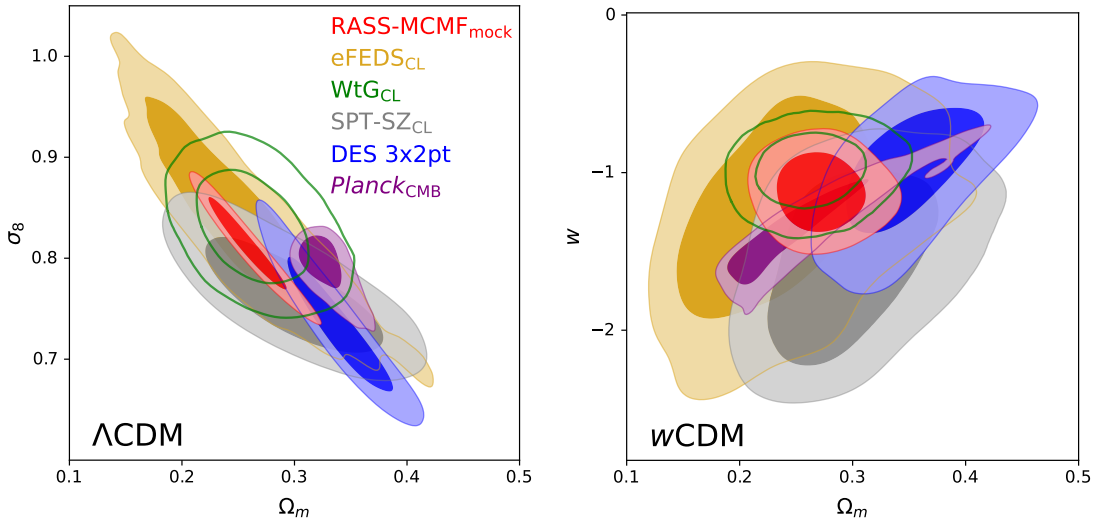

Figure 4.11: Forecast cosmological parameter constraints for both a ΛCDM (*top*) and a *w*CDM (*bottom*) cosmogony. In both panels the contours obtained using the RASS-MCMF mock cluster catalogue (red) are compared with those obtained from different datasets: eFEDS cluster cosmology (green; Chiu et al., 2023), SPT-SZ cluster cosmology (gray; Bocquet et al., 2019), DES 3×2-point constraints (blue; Abbott et al., 2022) and *Planck* primary CMB temperature and polarization anisotropy (purple; Planck Collaboration et al., 2020b). The contours indicate the 1σ and 2σ confidence intervals.

The RASS-MCMF sample of 8,449 clusters presented here has an estimated 10% contamination by residual non-cluster sources. We also present two subsets of the RASS-MCMF sample that contain 6,912 and 5,506 clusters with 5% and 1% contamination, respectively (see Table 4.1). The full extragalactic sky coverage of the RASS-MCMF cluster sample makes it particularly interesting for further analyses.

We include spectroscopic redshifts for $\sim$ 53% of the RASS-MCMF sample using public spectroscopic galaxy redshifts. For regions with deeper spectroscopic data, such as SDSS BOSS, we find spectroscopic redshifts for $\sim$ 93% of our sample. This RASS-MCMF subsample allows cluster studies that require spectroscopic redshifts over large contiguous areas. An analysis of the MCMF derived cluster photometric redshifts indicates a characteristic uncertainty of $\sigma_{\Delta z/1+z} = 0.0048 \pm 0.0001$.

The redshift distribution of the RASS-MCMF sample peaks at $z \sim 0.1$, and the new sample outnumbers any other ICM-selected cluster catalogs per redshift interval (dN/dz) out to $z \sim 0.4$, where the SZE selected ACT-DR5 clusters start to outnumber RASS-MCMF clusters. The RASS-MCMF sample probes the galaxy group mass regime ($M_X < 10^{14} M_\odot$) out to $z \sim 0.15$ and therefore densely samples the cosmic web at low redshifts.

When cross-matching RASS-MCMF with three large ICM selected (X-ray or SZE selected) cluster catalogs (Planck PSZ2, MCXC and ACT-DR5; Planck Collaboration et al., 2016a; Piffaretti et al., 2011; Hilton et al., 2021), we find between 800 and 1,000 matching clusters in each of the surveys and $\sim 2,000$ matched clusters in total. When cross-matching the 2,815 CODEX clusters flagged as clean with the full 2RXS catalog, we find just over 70% matches. Out of those matched sources we find $\sim$ 85% also in the final RASS-MCMF sample. We further match RASS-MCMF

with the 2RXS-based MARD-Y3 catalog (Klein et al., 2019) by directly matching to the same underlying 2RXS sources. We find generally good agreement between both catalogs. Sources appearing in only one of the catalogs can be explained by considering differences in catalog selection and scatter in observed richnesses. In the matched clusters from all these catalogs, the redshifts show generally good agreement with the small number of outliers being easily explained as a simple mistake in the older catalog or a case where there are multiple optical systems along the line of sight toward the source, and the highest significance peak chosen for the RASS-MCMF cluster is not the peak selected in the other catalogs (see discussion in Section 4.5).

We use X-ray selected clusters from the eROSITA Final Equatorial-Depth Survey (Brunner et al., 2022; Klein et al., 2022) to test and successfully confirm the purity of the RASS-MCMF sample to be 90%. The eROSITA X-ray survey mission (Predehl et al., 2021) is ongoing and has imaged the sky to greater depths than RASS. Early expectations were that as many as 10^5 X-ray selected clusters could be extracted from the eROSITA dataset (Merloni et al., 2012), and initial work in cluster catalog creation (Brunner et al., 2022; Klein et al., 2022; Liu et al., 2022) and also cluster cosmology using data from an early eFEDS test survey region (Chiu et al., 2023) have been very encouraging. We therefore look forward with excitement to the release of both the Russian and German parts of that rich X-ray dataset!

Our presentation of RASS-MCMF includes a cosmological forecast based on a RASS-MCMF-like mock catalog that includes various key aspects of the survey, such as X-ray existence likelihood selection, rejection of sources with low significance optical counterparts, exposure time and background variations across the sky and a realistic footprint. In addition to all these effects, we adopt realistic priors on the observable-mass scaling relation from an ongoing DES weak-lensing analysis of MARD-Y3 clusters. The cosmological parameter uncertainties from this forecast are 0.026, 0.033 and 0.15 (1σ) on the parameters Ω_m, σ_8 and w, respectively, making RASS-MCMF×DES a competitive dataset for cosmological analysis.

4.8 Appendix: Planck-PSZ2 clusters with potentially incorrect redshifts

In the table below we list RASS-MCMF matches to Planck PSZ2 clusters where the redshift listed in the PSZ2 catalog is likely incorrect.

4.9 Appendix: Sources from the NEP survey likely misclassified as clusters

In table 4.5 we list 18 sources from the NEP survey (Henry et al., 2006) classified as clusters that might not be real clusters after visual inspection. The first ten sources further have a counter part in the Million Quasars Catalog (Flesch, 2021).

Table 4.3: RASS-MCMF catalog column descriptions.

Column name	Description
NAME	Cluster name
RA_OPT	RA of optical centre for best counterpart in degrees
DEC_OPT	DEC of optical centre for best counterpart in degrees
CENT_TYPE	Type of optical centre: 1: GRZ galaxy density, 2:GRZ BCG, 3:GI galaxy density, 4:GI BCG
Z_1	Photo-z best counterpart
Z_2	Photo-z 2nd best counterpart
Z_3	Photo-z 3rd best counterpart
Z_SPEC_1	spectroscopic redshift best counterpart
LAMBDA_1	Richness best counterpart
LAMBDA_2	Richness 2nd best counterpart
LAMBDA_3	Richness 3rd best counterpart
F_CONT_1	$f_{\rm cont}$ best counterpart
F_CONT_2	$f_{\rm cont}$ 2nd best counterpart
F_CONT_3	$f_{\rm cont}$ 3rd best counterpart
M500_1	M_X for best counterpart in $M_\odot$
P_ANY	$p_{\rm any}$ from Salvato et al. (2018)
P_I	$P_{\rm I}$ from Salvato et al. (2018)
GRZ	LS DR10 GRZ footprint
GRZ_N	LS DR10 GRZ northern footprint
GI	LS DR10 GI footprint
PSTELLAR	$p_{\rm stellar}$ from Freund et al. (2022)
LIKELY_STELLAR	Likely stellar contaminant
LIKELY_QSO	Likely QSO contaminant
LG_LAM_MASS	$\log 10(10^{14}\lambda/M_X)$
MASKFRAC_120	DR10 Mask fraction within 120 arcsec radius
2RXS_NAME	2RXS source name
EXI_ML	2RXS existance likelihood
CTS	2RXS source counts
CERR	2RXS uncertainty on counts
RATE	count rate, including multiple detections in counts per second
2RXS_RATE	2RXS count rate in counts per second
2RXS_ERATE	2RXS uncertainty on count rate
EXPOSURE	2RXS exposure time in seconds
BGR	2RXS background in counts per pixel
2RXS_RA_DEG	2RXS RA
2RXS_DEC_DEG	2RXS DEC
LII	2RXS LII
BII	2RXS BII
EXT	2RXS source extend in image pixels
EXTERR	2RXS uncertainty on source extend
EXT_ML	2RXS likelihood of sources being extended
S_FLAG	2RXS screening flag

Table 4.4: PSZ2 clusters with possible incorrect redshifts. We list cluster name (PSZ2 name), position and redshifts from PSZ2 (RA_{PSZ2},DEC_{PSZ2},z_{PSZ2}) as well as 2RXS X-ray position (RA_{2RXS},DEC_{2RXS}) and MCMF based measurements of redshift (z), richness (λ) and $f_{\rm cont}$ for the best and second best counterpart. We finally provide a comment to each cluster.

PSZ2 NAME	RA_{PSZ2}	DEC_{PSZ2}	z_{PSZ2}	RA_{2RXS}	DEC_{2RXS}	z_1	z_2	λ_1	λ_2	$f_{\rm cont,1}$	$f_{\rm cont,2}$	Comment
PSZ2 G091.40-51.01	353.5028	7.0573	0.099	353.4740	7.0703	0.295	0.551	160.2	14.6	0.00	0.94	RASS-MCMF correct ($z_{\rm spec}$ = 0.2955)
PSZ2 G109.86+27.94	275.8330	78.3893	0.4	275.7792	78.3684	0.669	0.045	233.3	6.9	0.00	0.31	RASS-MCMF likely correct (λ = 233 system)
PSZ2 G181.71-68.65	31.6490	-14.8800	0.1529	31.6192	-14.8970	0.293	0.240	181.7	40.8	0.00	0.10	RASS-MCMF correct, PSZ redshfit from foregr. spiral
PSZ2 G281.09-42.51	56.2994	-66.5015	0.38	56.3505	-66.5000	0.557	0.130	153.1	6.4	0.00	0.59	RASS-MCMF likely correct (λ = 153 system)
PSZ2 G287.00-35.24	68.1017	-74.1685	0.43	68.0470	-74.1689	0.166	0.501	43.3	63.1	0.05	0.14	two clusters along the line of sight, PSZ redshift likely merges both clusters

Table 4.5: List of 2RXS matches to the ROSAT NEP survey that fail visual inspection. 2RXS sources that do have a match to the Million Quasars Catalog (Flesch, 2021) are listed first with names of the corresponding QSO, type and redshift listed, while sources without match to the Million Quasars Catalog are appended. We provide source name and position from the 2RXS catalog (2RXS Name, 2RXS_RA, 2RXS_DEC), redshift (z), richness (λ) and $f_{\rm cont}$ from MCMF measurements. Similar we provide source position and names of the matched NEP source

2RXS NAME	2RXS_RA	2RXS_DEC	z_1	λ_1	$f_{\rm cont}$	NEP_RA	NEP_DEC	NEP_NAME	z_NEP	QSO_Name	QSO_Type	QSO_z
2RXS J180606.6+681308	271.5276	68.2191	0.269	9.2	0.65	271.5275	68.2189	RX J1806.1+6813	0.303	WISEA J180609.00+681309.6	qX	0.3
2RXS J171640.0+641048	259.1669	64.1801	0.249	18.7	0.35	259.1654	64.1764	RX J1716.6+6410	0.251	WISEA J171636.31+641112.3	q	0.2
2RXS J180844.0+655705	272.1835	65.9516	0.495	5.5	1.00	272.1817	65.9514	RX J1808.7+6557	0.246	J180843.17+655705.4	X	0.3
2RXS J172839.2+704105	262.1634	70.6848	0.276	2.4	0.90	262.1646	70.6847	RX J1728.6+7041	0.551	RXS J17286+7041	QRX	0.551
2RXS J180732.2+642919	271.8846	64.4886	0.220	1.9	0.90	271.8846	64.4881	RX J1807.5+6429	0.239	3HSPJ180732.2+642926	BRX	0.239
2RXS J182237.8+664132	275.6575	66.6924	0.726	68.0	0.49	275.6558	66.6914	RX J1822.6+6641	0.089	J182237.52+664126.0	X	
2RXS J175857.7+652057	269.7407	65.3494	0.283	4.8	0.83	269.7400	65.3494	RX J1758.9+6520	0.365	WISEA J175856.74+652106.5	qX	0.5
2RXS J175406.5+645201	268.5272	64.8672	0.249	7.1	0.69	268.5221	64.8669	RX J1754.0+6452	0.246	CGRaBS J1754+6452	QRX	0.977
2RXS J174949.4+682319	267.4559	68.3886	0.522	9.0	1.00	267.4575	68.3875	RX J1749.8+6823	0.051	KUG 1750+683A	NRX	0.051
2RXS J183917.7+701820	279.8239	70.3058	0.103	6.6	0.55	279.8225	70.3056	RX J1839.2+7018	0.230	WISEA J183917.18+701823.7	qX	0.3
2RXS J172124.1+673313	260.3508	67.5537	1.491	3.5	0.99	260.3525	67.5539	RX J1721.4+6733	0.0861			
2RXS J172411.5+700027	261.0481	70.0078	0.325	5.2	0.86	261.0483	70.0075	RX J1724.1+7000	0.0386			
2RXS J174516.4+655617	266.3187	65.9382	0.612	22.9	0.89	266.3175	65.9381	RX J1745.2+6556	0.608			
2RXS J175130.6+701415	267.8779	70.2378	0.446	18.4	0.69	267.8779	70.2256	RX J1751.5+7013	0.4925			
2RXS J175211.8+652222	268.0494	65.3730	0.072	7.8	0.40	268.0500	65.3728	RX J1752.2+6522	0.3923			
2RXS J180416.1+672922	271.0672	67.4896	0.040	1.0	0.88	271.0650	67.4892	RX J1804.2+6729	0.0617			
2RXS J181119.1+644738	272.8297	64.7941	0.179	5.0	0.72	272.8296	64.7933	RX J1811.3+6447	0.451			
2RXS J181208.3+635336	273.0346	63.8935	0.046	6.3	0.36	273.0350	63.8931	RX J1812.1+6353	0.5408			

CHAPTER 5

CLUSTER COSMOLOGY USING RASS-MCMF: METHOD

5.1 Introduction

In Chapter 4 we used mocks to simulate a RASS-MCMF-like cluster catalog and test our initial likelihood model. One shortcoming of our previous work was the weak lensing mass calibration, which we did not apply directly. Instead, we used the parameter posteriors from the observable-mass relations derived from a similar cluster catalog, MARD-Y3 (Klein et al., 2019), as priors (Singh et al., in prep.). On the other hand, due to the way the abundance likelihood was derived, only the X-ray information was included in our analysis.

In our final analysis we want to explicitly include the weak lensing mass calibration likelihood. Additionally, we want to re-derive the abundance likelihood to make full use of our multi-observable halo observable function (HOF) and include adjustments to account for other significant effects, such as the absorption of X-ray photons by hydrogen. Thus, we are developing a RASS-MCMF specific method that builds upon the method previously developed for a SPTxDES study (Bocquet et al., 2023, 2024; Singh et al., 2024).

RASS-MCMF is an X-ray selected and optically confirmed cluster catalog, where the cosmology sample (with 99% purity) contains ~5000 galaxy clusters at $0.02 < z < 1.3$. For each cluster we have at least three observables: the observed X-ray countrate ($\hat{\eta}$), the observed optical richness ($\hat{\lambda}$), and the photometric redshift. A fourth observable, the weak lensing shear profile, is also available for a fraction of the clusters. This then allows us to use the RASS-MCMF cosmology sample together with the HMF, which describes the relation between the observables and the underlying halo masses as a function of redshift, to study cosmology by modeling our multi-observable HOF. Moreover, the weak lensing shear profiles provide essential mass information, enabling the calibration of the observable-mass relations.

In this chapter I will describe the derivation of the final likelihood that we will apply to the RASS-MCMF cosmology sample. In § 5.2 I introduce weak gravitational lensing, it's mathematical framework and how they are used to estimate cluster masses. In § 5.3 I describe the Poisson likelihood for cluster number counts, going into details describing both the abundance

and weak lensing mass calibration likelihoods. § 5.4 covers the creation of mocks with properties similar to those of the RASS-MCMF cosmology sample, which I use to validate the likelihood. In § 5.5 I present a validation of the likelihood using the mocks and compare our forecast with the latest cluster number counts cosmology results. Finally, in § 5.6 I summarize our results and discuss different improvements that can be applied to our analysis.

5.2 Mass calibration using weak gravitational lensing

In principle the parameters of the different observable-mass relations cannot be reliably predicted given our lack of detailed knowledge of, for example, the ICM and/or galaxy evolution. Thus, a good way of calibrating the observable-mass relation is to combine them with an effect that is given by the entire matter distribution of a halo: the weak gravitational lensing of background galaxies by the foreground galaxy cluster lens (e.g., Huterer & White, 2002; Rozo et al., 2009a; Bocquet et al., 2019; Grandis et al., 2024).

Following the theory of general relativity, space-time is curved by objects with mass. In particular for massive galaxy clusters, their gravitational potential curves the space-time through which photons are travelling, such that galaxies behind the clusters (in the line-of-sight) appear distorted for an observer (Bartelmann, 2010; Smit et al., 2017; Chan et al., 2017). These distortions produce what is known as gravitational lensing (Soucail et al., 1987; Mellier, 1999; Bartelmann & Schneider, 2001), which can be used to reconstruct the total mass of the cluster (Tyson et al., 1990; Clowe et al., 2004, 2006; Ma et al., 2010; Applegate et al., 2014; Gonzalez et al., 2018).

In general, weak gravitational lensing, which weakly distorts the shapes (shear; γ) of background galaxies and magnifies their images (convergence; κ), enables mass measurements that are independent of the dynamical state of galaxy clusters. It is worth noting, however, that strong gravitational lensing is also frequently employed in this context. Estimating the shear of the galaxies, which stretches the galaxies tangentially around the foreground matter (the cluster), is not straightforward. To quantify this "stretch", one measures the shape, e, of the galaxies, that also includes the intrinsic shape, which is a) not possible to accurately determine and b) not the same for all the source galaxies. To account for this particular difficulty, it is assumed that the shapes of the background galaxies are more or less random, and, thus, many background galaxies are combined to average down the shape noise.

A brief description of the mathematical framework for weak lensing goes as follows. In this framework, we define two planes: the source plane, which represents the true positions of background sources, and the lens plane, where a massive object (the lens) lies between the observer and the source. We start by imagining a source located at an angular position $\boldsymbol{\beta}$ in the source plane, representing its true location relative to the observer's line of sight. The gravitational field of the lens deflects the light from the source, causing it to appear at an angular position $\boldsymbol{\theta}$ in the observed sky. Fig. 5.1 shows a schematic representation of the angular positions of a source relative to an observer. The relationship between these angular positions, $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$, is described by the lens equation, which is given by

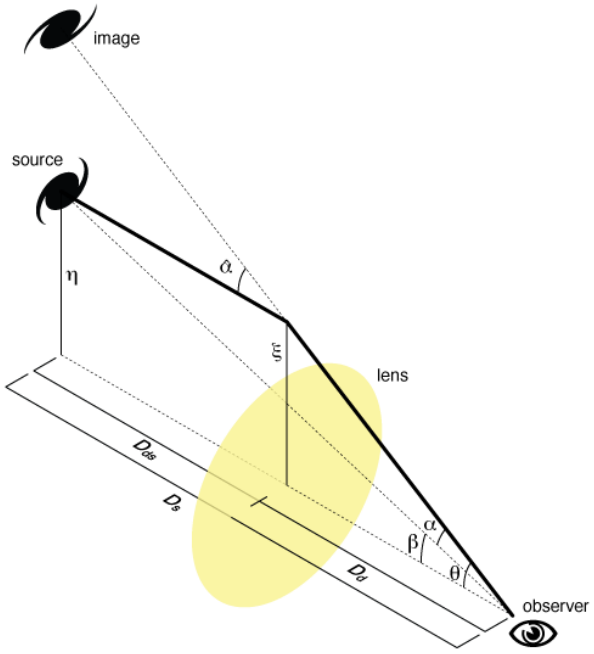

Figure 5.1: Schematic idea of light distortion of a source galaxy by a lens. β and θ are the true and observed angular positions of the source, respectively, with α being the deflection angle between the source and image angular positions. D_d, D_s and D_ds are the observer-lens, observer-source and lens-source distances, respectively. Figure created by Michael Sachs.

$$\boldsymbol{\beta} = \boldsymbol{\theta} - \boldsymbol{\alpha}(\boldsymbol{\theta}) \tag{5.1}$$

where $\boldsymbol{\alpha}(\boldsymbol{\theta})$ is the reduced bending angle, or deflection angle, between the source and image angular positions, and it depends on the lens mass distribution.

Given that we work in the weak lensing regime, the deflection angle is small, and thus we can consider the sky to be flat for small patches. Let us call the axis of the Cartesian coordinates of this small patch as θ_1 and θ_2, then the deflection angle (approximating to a thin lens) is defined as

$$\boldsymbol{\alpha}(\boldsymbol{\theta}) = \frac{1}{\pi} \int \mathrm{d}\boldsymbol{\theta}' \kappa(\boldsymbol{\theta}') \frac{\boldsymbol{\theta} - \boldsymbol{\theta}'}{\left|\boldsymbol{\theta} - \boldsymbol{\theta}'\right|^2}, \tag{5.2}$$

with κ, the convergence, being the dimensionless surface mass density of the lens, such that

$$\kappa(\boldsymbol{\theta}) = \frac{\Sigma(\boldsymbol{\theta})}{\Sigma_\mathrm{crit}}; \text{ with } \Sigma(\boldsymbol{\theta}) = \int \mathrm{d}z \rho(D_\mathrm{d}\boldsymbol{\theta}, z) \text{ and } \Sigma_\mathrm{crit} = \frac{c^2}{4\pi G} \frac{D_\mathrm{s}}{D_\mathrm{d}} \tag{5.3}$$

where Σ is the density profile projected along the line of sight and Σ_crit is the critical surface mass density (e.g., Schneider et al., 1992). D_s and D_d are the distances from the observer to the source and to the lens, respectively, and c is the speed of light.

For an extended source we can generalize the lens equation for a small deviation in the source plane $\delta\boldsymbol{\beta}$, which translates into a small deviation in the image plane $\delta\boldsymbol{\theta}$, so that $\boldsymbol{\beta} + \delta\boldsymbol{\beta} = \boldsymbol{\theta} + \delta\boldsymbol{\theta} - \boldsymbol{\alpha}(\boldsymbol{\theta} + \delta\boldsymbol{\theta})$. Then, the shape distortion of the distant galaxy by the lens can be written as

$$\delta\boldsymbol{\beta} = A\delta\boldsymbol{\theta}; A = I - \frac{\partial\boldsymbol{\alpha}}{\partial\boldsymbol{\theta}} = \begin{pmatrix} 1-\kappa-\gamma_1 & -\gamma_2 \\ -\gamma_2 & 1-\kappa-\gamma_1 \end{pmatrix} \tag{5.4}$$

where γ_1 and γ_2 are the two components of the gravitational shear defined as

$$\gamma_1(\boldsymbol{\theta}) = \frac{1}{2}\left(\frac{\partial\alpha_1}{\partial\theta_1} - \frac{\partial\alpha_2}{\partial\theta_2}\right); \gamma_2(\boldsymbol{\theta}) = \frac{\partial\alpha_1}{\partial\theta_2} = \frac{\partial\alpha_2}{\partial\theta_1}. \tag{5.5}$$

The shear can be measured by averaging the shapes of many galaxies. Let us call the intrinsic shape of galaxies as ϵ_i ($i = 1, 2$), then, assuming the orientations to be random on average, we get that $<\epsilon_i^j> \approx \frac{1}{N}\sum_j \epsilon_i^j = 0$, where ϵ_i^j is the intrinsic shape of the j-th galaxy. The observed shape of each galaxy is given by $\hat{\epsilon}_i = \epsilon_i + \gamma_i$, and thus the shear can be measured as $<\hat{\epsilon}_i^j> \approx \gamma_i$. However, the intrinsic shape of galaxies is 10 to 100 times bigger than the typical weak lensing shear, and thus a number of $10^3 - 10^4$ galaxies are needed for a significant detection.

Simulations of lensing distortions show that high density lens of massive objects distorts the shapes of background galaxies along the tangential direction: tangential shear (γ_t). For a given reference point, e.g. the center of a galaxy cluster, the tangential shear is defined as

$$\gamma_t = -\gamma_1 \cos 2\phi - \gamma_2 \sin 2\phi \tag{5.6}$$

where ϕ is the polar angle of the observed galaxy in the sky plane. Assuming an spherically symmetric lens (with the symmetry axis along the line of sight), the tangential shear profile caused by a projected mass distribution, $\Sigma(r)$, at a distance r from the cluster center is given by

$$\gamma_t(r) = \langle\kappa(r)\rangle - \kappa(r) = \frac{\langle\Sigma(<r)\rangle - \Sigma(r)}{\Sigma_{\rm crit}} = \frac{\Delta\Sigma(r)}{\Sigma_{\rm crit}}, \tag{5.7}$$

where it can be noted that the shear is determined by the integrated mass at distances $< r$ from the cluster center (e.g., Miralda-Escude, 1991). To measure the average tangential component of the shear of all galaxies within an annulus, one averages the tangential shear of each j-th galaxy using weights given by, for example, their luminosities.

Given that source galaxies experience both shear and magnification, the weak-lensing observable is the reduced tangential shear g_t, defined as

$$g_t = \frac{\gamma_t}{1-\kappa}. \tag{5.8}$$

To estimate the mass density of a given cluster, it is common practice to fit a model to the data. The most widely used model is the so called Navarro, Frenk and White profile (NFW; Navarro et al., 1996, 1997), given by

$$\rho(r) = \delta_{\rm s}\rho_{\rm crit}\left[\frac{r}{cr_{200c}}\left(1 + \left(\frac{r}{cr_{200c}}\right)^2\right)\right]^{-1} \tag{5.9}$$

where c the halo concentration parameter defined as $c = r_{200c}/r_s$, with r_s being a characteristic radius of the cluster, and δ_s is a characteristic overdensity that depends on c

$$\delta_s = \frac{200}{3} \frac{c^3}{\ln(1+c) - c/(1+c)} \tag{5.10}$$

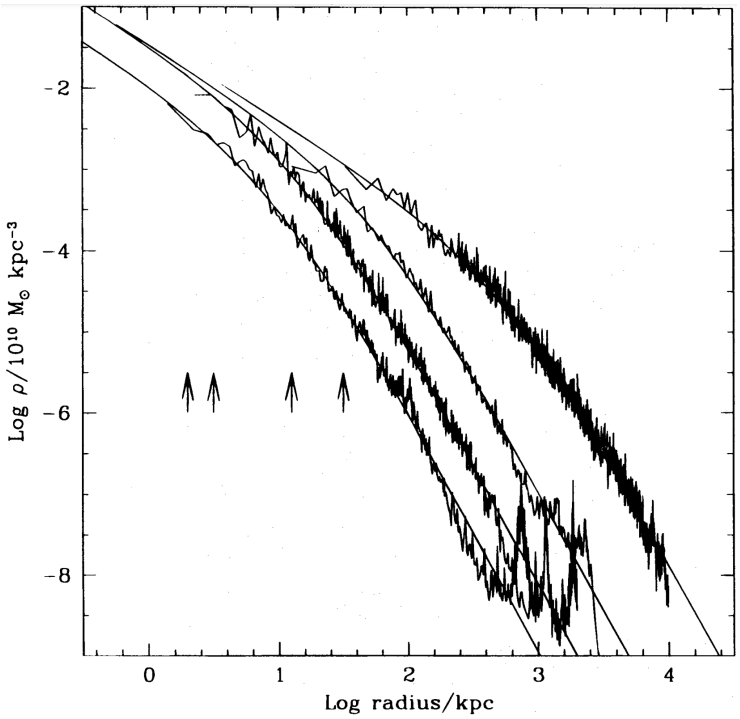

Figure 5.2: Density profiles of four simulated galaxy clusters spanning masses from 3×10^{11} to 3×10^{15}. The NFW fit from eq. 5.9 is also shown. Taken from Navarro et al. (1996).

Fig. 5.2 shows the fit of eq. 5.9 to the density profiles of four simulated galaxy clusters with different masses. Bartelmann (1996) gives a relation between an NFW lens and the surface mass density profile; Σ_{NFW}. Replacing in eq. 5.7, a tangential shear profile can then be described by a series of analytical equations (e.g., Wright & Brainerd, 2000).

5.3 Likelihood function

When studying the number of galaxy clusters in a given survey, the expected counts follow Poisson statistics due to the discrete nature of the clusters. The probability mass function (PMF) of the Poisson distribution is given by

$$P(k|\lambda) = \frac{\mu^k e^{-\mu}}{k!} \tag{5.11}$$

which describes the probability of observing k events given the rate parameter μ. In statistical context, the likelihood function, $\mathcal{L}$, represents the probability of the observed data given a set of parameters, such that

$$\mathcal{L}(\mu|k) = \prod_i P(k_i|\mu_i) = \prod_i \frac{\mu_i^{k_i} e^{-\mu_i}}{k_i!} \tag{5.12}$$

with the log-likelihood function being

$$\ln \mathcal{L}(\mu|k) = \sum_i (k_i \ln \mu_i - \mu_i - \ln k_i!) \tag{5.13}$$

here k_i is the observed count of events (which, for our intents and purposes is the observed number of galaxy clusters), in the i-th bin, and μ_i is the expected number of clusters for the same bins, predicted by a given model.

In cosmology, we often deal with continuous parameters (like X-ray countrate and redshift) rather than discrete bins. Instead of k_i, we consider the observed distribution of galaxy clusters over these continuous variables by assuming a space of fine bins such that each bin contains at most one cluster.

Given that each observed cluster contributes individually, $k_i = 1$. We can then rewrite Eq. 5.13 as

$$\ln \mathcal{L} = \sum_j \ln \mu_j - \sum_i \mu_i \tag{5.14}$$

where the sum over j accounts for all the bins for which we have at least one cluster, whereas the sum over i runs over all bins. If we assume a infinitesimally small binning dx, Eq. 5.14 can be written as

$$\ln \mathcal{L}(\mu|k) = \sum_j \ln \left(\frac{\mathrm{d}\mu}{\mathrm{d}x} \mathrm{d}x \right) \Big|_{x_j} - \int \frac{\mathrm{d}\mu}{\mathrm{d}x} \mathrm{d}x = \sum_j \ln \frac{\mathrm{d}\mu}{\mathrm{d}x} \Big|_{x_j} - \int \frac{\mathrm{d}\mu}{\mathrm{d}x} \mathrm{d}x + \text{const.} \tag{5.15}$$

Assuming a set of observed parameters for our cluster sample, O, we can represent the expected number of clusters per unit O and z as

$$\frac{dN(O, z|\mathbf{p})}{\mathrm{d}O\mathrm{d}z} \Theta_s(O, \mathbf{p}) \tag{5.16}$$

where $\mathbf{p}$ is a vector containing all observable-mass relations, cosmological and nuisance parameters and $\Theta_s(O, \mathbf{p})$ is the selection function applied on the observable O for a given set of parameters $\mathbf{p}$.

Finally, we can write 5.15 in terms of our observables and selection functions such that

$$\ln \mathcal{L}(\mathbf{p}) = \sum_i \ln \left(\frac{dN(O_i, z_i)}{\mathrm{d}O\mathrm{d}z} \Theta_s(O_i, \mathbf{p}) \right) \Big|_{O_i, z_i} - \int \frac{dN(O, z)}{\mathrm{d}O\mathrm{d}z} \Theta_s(O, \mathbf{p}) \mathrm{d}O\mathrm{d}z \tag{5.17}$$

where the integral term represents the total expected number of clusters over the entire parameter space given our model, whereas the sum is over the clusters i with parameters O_i, z_i, $\mathbf{p}_i$.

Finally, given a model of our data governed by observable-mass relations between observables and mass, with Eq. 5.17 we can get the likelihood of the parameters that describe our observable-mass relations, as well as the cosmological parameters.

5.3.1 Poisson likelihood for cluster number counts

We can reformulate Eq. 5.17, to explicitly incorporate our observables: countrate ($\hat{\eta}$), richness ($\hat{\lambda}$), redshift (z) and the reduced tangential shear (g_t)

$$\begin{aligned}\ln \mathcal{L}(\mathbf{p}) = \Bigg[\sum_i \ln \left(\Theta_s(\hat{\eta}, z, \hat{\lambda}) \frac{\mathrm{d}^4 N(\mathbf{p})}{\mathrm{d}\hat{\eta}\mathrm{d}\hat{\lambda}\mathrm{d}z\mathrm{d}\mathbf{g}_t} \right) \Bigg|_{i-\text{th cluster}} \\ - \int \cdots \int \mathrm{d}z\, \mathrm{d}\hat{\lambda}\, \mathrm{d}\hat{\eta}\, \mathrm{d}\mathbf{g}_t\, \Theta_s(\hat{\eta}, z, \hat{\lambda}) \frac{\mathrm{d}^4 N(\mathbf{p})}{\mathrm{d}\hat{\eta}\mathrm{d}\hat{\lambda}\mathrm{d}z\mathrm{d}\mathbf{g}_t} \Bigg],\end{aligned} \tag{5.18}$$

with $\Theta_s(\hat{\eta}, z, \hat{\lambda})$ representing the selection function of the different observables. We refer to $\frac{\mathrm{d}^4 N(\mathbf{p})}{\mathrm{d}\hat{\eta}\mathrm{d}\hat{\lambda}\mathrm{d}z\mathrm{d}\mathbf{g}_t}$ as the halo observable function (HOF), given by

$$\begin{aligned}\frac{\mathrm{d}^4 N(\mathbf{p})}{\mathrm{d}\hat{\eta}\mathrm{d}\hat{\lambda}\mathrm{d}z\mathrm{d}\mathbf{g}_t} &= \frac{\mathrm{d}^3 N(\mathbf{p})}{\mathrm{d}\hat{\eta}\mathrm{d}\hat{\lambda}\mathrm{d}z} \frac{\mathrm{d}N(\mathbf{p})}{\mathrm{d}\mathbf{g}_t} \\ &= \int \cdots \int \mathrm{d}\Omega_s \mathrm{d}M \mathrm{d}\eta \mathrm{d}\lambda \mathrm{d}M_{\mathrm{WL}} P(\hat{\eta}|\eta) P(\hat{\lambda}|\lambda) P(\eta, \lambda, M_{\mathrm{WL}}|M, z, \mathbf{p}) \\ &\quad \times P(\mathbf{g}_t|M_{\mathrm{WL}}, z) \frac{\mathrm{d}^3 N(\mathbf{p})}{\mathrm{d}M\mathrm{d}z\mathrm{d}V} \frac{\mathrm{d}V(z, \mathbf{p})}{\mathrm{d}\Omega_s}\end{aligned} \tag{5.19}$$

where Ω_s is the survey solid angle, which accounts for the sky area of interest, $\frac{\mathrm{d}^3 N}{\mathrm{d}M\mathrm{d}z\mathrm{d}V}$ is the halo mass function for a given mass and redshift bin per volume unit. We integrate over the different intrinsic quantities that we relate to mass (η and λ) and over the weak-lensing mass (M_{WL}). The HMF is weighted both by the probabilities of the intrinsic and the observed richness and countrate (more details in § 5.3.2).

Since there is no lensing selection in our cluster catalog, we can rewrite Eq. 5.18 as

$$\begin{aligned}\ln \mathcal{L}(\mathbf{p}) = \Bigg[\sum_i \ln \left(\Theta_s(\hat{\eta}, z, \hat{\lambda}) \frac{\mathrm{d}^3 N(\mathbf{p})}{\mathrm{d}\hat{\eta}\mathrm{d}\hat{\lambda}\mathrm{d}z} \frac{\mathrm{d}N(\mathbf{p})}{\mathrm{d}\mathbf{g}_t} \right) \Bigg|_{i-\text{th cluster}} \\ - \iiint \mathrm{d}z\, \mathrm{d}\hat{\lambda}\, \mathrm{d}\hat{\eta}\, \Theta_s(\hat{\eta}, z, \hat{\lambda}) \frac{\mathrm{d}^3 N(\mathbf{p})}{\mathrm{d}\hat{\eta}\mathrm{d}\hat{\lambda}\mathrm{d}z} \Bigg],\end{aligned} \tag{5.20}$$

where

$$\begin{aligned}\frac{\mathrm{d}^3N(\mathbf{p})}{\mathrm{d}\hat{\eta}\mathrm{d}\hat{\lambda}\mathrm{d}z} &= \int \mathrm{d}\mathbf{g}_t \frac{\mathrm{d}^4N(\mathbf{p})}{\mathrm{d}\hat{\eta}\mathrm{d}\hat{\lambda}\mathrm{d}z\mathrm{d}\mathbf{g}_t} \\ &= \int \cdots \int \mathrm{d}\Omega_s \mathrm{d}M \mathrm{d}\eta \mathrm{d}\lambda \; P(\hat{\eta}|\eta) P(\hat{\lambda}|\lambda) P(\eta, \lambda|M, z, \mathbf{p}) \frac{\mathrm{d}^3N(\mathbf{p})}{\mathrm{d}M\mathrm{d}z\mathrm{d}V} \frac{\mathrm{d}V(z, \mathbf{p})}{\mathrm{d}\Omega_s}\end{aligned} \tag{5.21}$$

We separate our likelihood to make it easier to describe its different components. Specifically, we refer to the portion of the likelihood that handles the expected number of clusters, given the observables $\hat{\eta}$ and $\hat{\lambda}$, as the *abundance likelihood*. Similarly, the portion of the likelihood that incorporates the weak-lensing observable is referred to as the *mass calibration likelihood*. In other words,

$$\begin{aligned}\ln \mathcal{L}(\mathbf{p}) = &\sum_i \ln \left(\Theta_s(\hat{\eta}, z, \hat{\lambda}) \frac{\mathrm{d}^3N(\mathbf{p})}{\mathrm{d}\hat{\eta}\mathrm{d}\hat{\lambda}\mathrm{d}z} \right) \Bigg|_{i-\text{th cluster}} - \iiint \mathrm{d}z \, \mathrm{d}\hat{\lambda} \, \mathrm{d}\hat{\eta} \; \Theta_s(\hat{\eta}, z, \hat{\lambda}) \frac{\mathrm{d}^3N(\mathbf{p})}{\mathrm{d}\hat{\eta}\mathrm{d}\hat{\lambda}\mathrm{d}z} \\ &+ \sum_i \ln \left(\frac{\mathrm{d}N(\mathbf{p})}{\mathrm{d}\mathbf{g}_t} \right) \Bigg|_{i-\text{th cluster}} \\ = &\ln \mathcal{L}_{\text{abundance}} + \ln \mathcal{L}_{\text{lensing}}.\end{aligned} \tag{5.22}$$

It is important to emphasize that Eq. 5.22 describes the usage of individual cluster mass profiles for the mass calibration likelihood. Such method has been used in previous studies using galaxy clusters as cosmological probes (e.g., Bocquet et al., 2023; Mazoun et al., 2024; Vogt et al., 2024). However, this is not the exact approach we take. Instead, we use stacks of cluster matter profiles for the weak-lensing mass calibration. Even though the mass calibration likelihood will then be, in principle, different, Eq. 5.22 is still valid. In the following sections I will go into deeper details and discuss both the abundance and the mass calibration likelihoods.

5.3.2 Abundance likelihood

The abundance likelihood is given by

$$\begin{aligned}\ln \mathcal{L}_{\text{abundance}} &= \sum_i \ln \left(\Theta_s(\hat{\eta}, z, \hat{\lambda}) \frac{\mathrm{d}^3N(\mathbf{p})}{\mathrm{d}\hat{\eta}\mathrm{d}\hat{\lambda}\mathrm{d}z} \right) \Bigg|_{i-\text{th cluster}} - \iiint \mathrm{d}z \, \mathrm{d}\hat{\lambda} \, \mathrm{d}\hat{\eta} \; \Theta_s(\hat{\eta}, z, \hat{\lambda}) \frac{\mathrm{d}^3N(\mathbf{p})}{\mathrm{d}\hat{\eta}\mathrm{d}\hat{\lambda}\mathrm{d}z} \\ &= \sum_i \ln \left(\Theta_s(\hat{\eta}, z, \hat{\lambda}) \frac{\mathrm{d}^3N(\mathbf{p})}{\mathrm{d}\hat{\eta}\mathrm{d}\hat{\lambda}\mathrm{d}z} \right) \Bigg|_{i-\text{th cluster}} - N_{\text{tot}}.\end{aligned} \tag{5.23}$$

The treatment of Eq. 5.23 is pretty similar to what was described in Chapter 4 (§4.6.1). However, the way the likelihood is derived is slightly different, and we expect a higher constraining power from our new method. In this subsection I will explain our new method while also discussing the differences with respect to the likelihood used in Chapter 4 (see §5.3.2.1).

We start from the HMF per unit volume, $\frac{\mathrm{d}^2N}{\mathrm{d}M\mathrm{d}z}$, in mass and redshift bins, estimated following a Tinker et al. (2008) HMF. Then, we iterate each redshift bin and transform the HMF into a HOF for our two observables, η and λ, by weighting the HMF by the probabilities of the observables given the mass. We can effectively express this integral, at fixed redshift, as

$$\frac{\mathrm{d}^2N}{\mathrm{d}\eta\mathrm{d}\lambda} = \int \mathrm{d}M P(\eta, \lambda|M, z, \mathbf{p}) \frac{\mathrm{d}N}{\mathrm{d}M}, \tag{5.24}$$

where $(P(\eta, \lambda|M, z, \mathbf{p})$ represents the probability density function of the intrinsic richness (λ) and countrate (η) given the mass and redshift. This distribution is modeled as a 2-dimensional lognormal with means defined by the observable-mass relations

$$\langle \ln \lambda | M_{200\mathrm{c}}, z \rangle = \ln \lambda_0 + \lambda_M \ln\left(\frac{M_{200\mathrm{c}}}{M_{\mathrm{piv}}}\right) + \lambda_z \ln\left(\frac{1+z}{1+z_{\mathrm{piv}}}\right), \tag{5.25}$$

and

$$\begin{aligned}\left\langle \ln\left(\frac{\eta}{\mathrm{counts/sec}}\middle| M_{200c}, z\right)\right\rangle = \ln\eta_0 &- 2\ln\left(\frac{D_{\mathrm{L}}(z)}{D_{\mathrm{L}}(z_{\mathrm{piv}})}\right) \\ &+ \left[\eta_M + \eta_{M-z}\ln\left(\frac{1+z}{1+z_{\mathrm{piv}}}\right)\right]\ln\left(\frac{M_{200c}}{M_{\mathrm{piv}}}\right) \\ &+ \eta_z \ln\left(\frac{1+z}{1+z_{\mathrm{piv}}}\right) + 2\ln\left(\frac{E(z)}{E(z_{\mathrm{piv}})}\right),\end{aligned} \tag{5.26}$$

where $M_{\mathrm{piv}} = 2.5\times10^{14}h^{-1}\mathrm{M}_\odot$ and $z_{\mathrm{piv}} = 0.2$ are chosen to reflect the median mass and redshift of the RASS-MCMF cosmology sample, D_{L} is the luminosity distance and $E(z) = H(z)/H_0$ is the expansion history of the Universe. For both relations, the parameter pairs $\ln\lambda_0$ and $\ln\eta_0$ correspond to the normalization; λ_M and η_M describe the mass trend; and λ_z and η_z capture the redshift trend of their respective relations. Following Chiu et al. (2023), we also allow for cross-scaling between mass and redshift, given by the η_{M-z} parameter.

For λ we adopt a log-normal intrinsic scatter, at fixed mass and redshift, that is the same for all redshifts and masses, of

$$\sigma_{\ln\lambda} = (\mathrm{Var}[\ln\lambda|M, z])^{\frac{1}{2}}, \tag{5.27}$$

whereas for η we model the log-normal intrinsic scatter with a redshift dependency (see §5.4) as

$$\sigma_{\ln\eta} = \sigma_{\ln\eta_0}\left(\frac{1+z}{1+z_{\mathrm{piv}}}\right)^{\sigma_{\ln\eta_z}}. \tag{5.28}$$

Additionally, we account for the possibility of correlated scatter between η and λ, $\rho_{\eta,\lambda}$, in Eq. 5.24. Thus, $P(\eta, \lambda|M, z, \mathbf{p})$ is defined as the probability density function of a 2-dimensional multivariate Gaussian given by

$$f(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \times \exp\left(-\frac{1}{2[1-\rho^2]}\left[\left(\frac{x-\mu_X}{\sigma_X}\right)^2 - 2\rho\left(\frac{x-\mu_X}{\sigma_X}\right)\left(\frac{x-\mu_X}{\sigma_X}\right) + \left(\frac{x-\mu_X}{\sigma_X}\right)^2\right]\right) \tag{5.29}$$

where x and y correspond to arrays in $\ln\eta$ and $\ln\lambda$, respectively, with means μ_X and μ_Y given by their observable-mass relations (Eq. 5.26 and 5.25). In this case, the covariance matrix is given by

$$\mathbf{\Sigma} = \begin{pmatrix} \sigma^2_{\ln\eta} & \rho_{\eta,\lambda}\,\sigma_{\ln\lambda}\,\sigma_{\ln\eta} \\ \rho_{\eta,\lambda}\,\sigma_{\ln\lambda}\,\sigma_{\ln\eta} & \sigma^2_{\ln\lambda} \end{pmatrix} \tag{5.30}$$

Next, we need to take into account that we don't work with intrinsic quantities, but rather with the observed richness ($\hat{\lambda}$) and observed countrates (η), which include the measurement noise. At fixed redshift, we transform the HOF from intrinsic to observed quantities as

$$\frac{\mathrm{d}^2N}{\mathrm{d}\hat{\eta}\mathrm{d}\hat{\lambda}} = \iint \mathrm{d}\eta\mathrm{d}\lambda P(\hat{\eta}|\eta)P(\hat{\lambda}|\lambda)\frac{\mathrm{d}^2N}{\mathrm{d}\eta\mathrm{d}\lambda}, \tag{5.31}$$

where $P(\hat{\lambda}|\lambda)$ is the probability of the observed richness given the intrinsic richness, which we model as a Poisson likelihood in the Gaussian limit, with mean $\hat{\lambda}$ and scatter $\sqrt{\hat{\lambda}}$. On the other hand, we model the probability of the observed countrates given the intrinsic countrates, $P(\hat{\eta}|\eta)$, using the Poisson likelihood for the expected number of photons, $n_\gamma = \eta t_{\mathrm{exp}}$.

Eq. 5.31 gives us the final HOF we use, including the redshift dimension. However, the HOF is estimated for a 1 deg^2 area. To account for the survey solid angle we use the exposure time distribution of the sky and construct a PDF using exposure time bins. Integrating the PDF over all exposure time bins gives us the total survey area of ∼22,300 deg^2. The exposure time PDF is shown in Fig. 5.3, with mean at $t_{\mathrm{exp}} \sim 400$ s. We interpolate the sky area for each exposure time bin and at the cluster's exposure time to get the total number of clusters and the expectation value of each cluster, respectively.

5.3.2.1 Previous abundance likelihood estimation

In Chapter 4 we use the abundance likelihood, along with observable-mass relation priors, to get a forecast of the constraining power on the cosmological parameters. However, we estimated the likelihood following the method applied to the SPT cluster sample (Bocquet et al., 2023, 2024). Although the likelihood is in principle similar, the application is different. Here I'll give a brief summary of the main difference of the method used in Chapter 4 and the one presented here.

The likelihood derived in Chapter 4 also starts from the HMF. This HMF is then convolved in mass space using the intrinsic scatter in richness and countrate. Then, the convolved HMF is further convolved to account for $P(\hat{\lambda}|\lambda)$ and integrated from $\hat{\lambda} \geq \hat{\lambda}\mathrm{min}(z)$. This final convolved HMF is transformed into a 2D HOF in η and z space. The measurement uncertainty in η is then

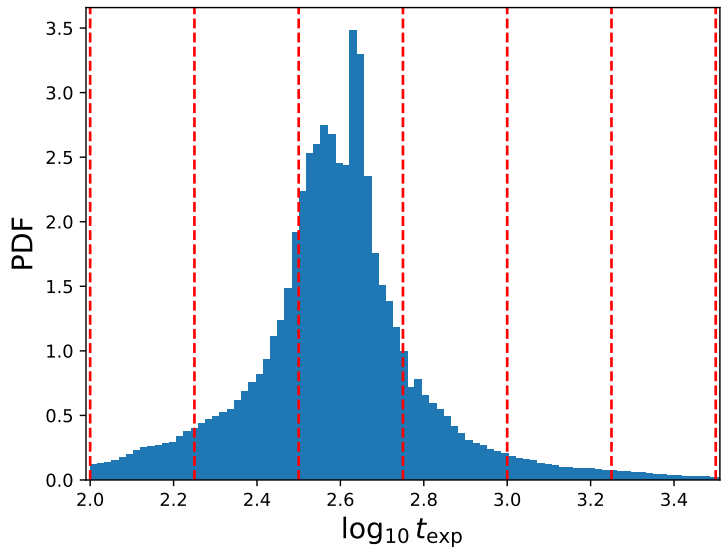

Figure 5.3: Probability density function of the exposure time distribution of 2RXS. Red dashed lines mark the exposure times for which we estimate the HOFs that are used to built an interpolator for the other $t_{\rm exp}$ bins.

applied using the same method described in this chapter. Following the notation used in this chapter, this procedure can be written as

$$\begin{aligned}\frac{\mathrm{d}^2N}{\mathrm{d}\hat{\eta}\mathrm{d}z} &= \int_{\hat{\lambda}_{\rm min}}^{\infty} \mathrm{d}\hat{\lambda}\frac{\mathrm{d}^3N}{\mathrm{d}\hat{\eta}\mathrm{d}\hat{\lambda}\mathrm{d}z} \\ &= \int \mathrm{d}\eta P(\hat{\eta}|\eta)\left(\int_{\hat{\lambda}_{\rm min}}^{\infty}\mathrm{d}\hat{\lambda}\mathrm{d}\lambda P(\hat{\lambda}|\lambda)P(\eta,\lambda|M,z,\mathbf{p})\frac{\mathrm{d}^2N}{\mathrm{d}M\mathrm{d}z}\right)\frac{\mathrm{d}M}{\mathrm{d}\eta}\end{aligned} \tag{5.32}$$

This means that the individual richness information of each cluster is not used in the abundance likelihood, and the only richness information comes from the integration at $\hat{\lambda} > \hat{\lambda}_{\rm min}(z)$. In the SPT case, the richness information is used in the weak-lensing mass calibration likelihood which includes the probability of observing a richness value given their observable (SZ significance: ζ) and redshift (Bocquet et al., 2023; Singh et al., 2024). Since in our new method we are including the richness as well as the countrates of the clusters, we expect our abundance-only analysis to give better constraints than the previous method.

5.3.2.2 N_H column density absorption

One important factor to consider when working in the X-ray regime is the absorption of photons by atomic elements in the interstellar medium, primarily hydrogen (e.g., Henke et al., 1982). This absorption effect, quantified by the hydrogen column density (N_H), impacts observed X-ray fluxes by reducing the number of low-energy photons that reach the detector. The Portable Interactive Multi-Mission Simulator (PIMMS) is a tool developed to simulate X-ray observations

across multiple telescopes accounting for the effects of N_H absorption. PIMMS uses different models to estimate how interstellar gas absorbs photons Brown & Gould (1970); Morrison & McCammon (1983), providing conversions between intrinsic and observed countrates and fluxes. By specifying a source model (e.g., thermal bremsstrahlung or power-law), energy range, N_H value, and instrument response, it is possible to predict how absorption affects X-ray observations.

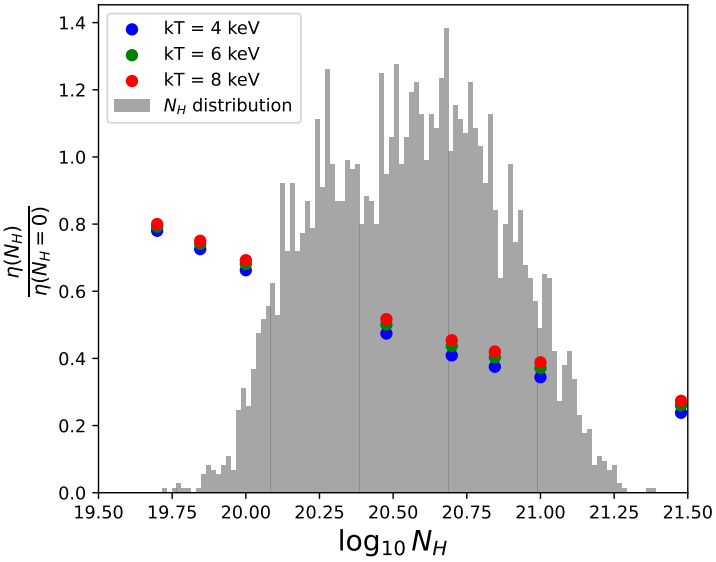

Figure 5.4: Ratio of unabsorbed to absorbed X-ray countrates ($\eta(N_H)/\eta(N_H = 0)$), estimated using PIMMS, as a function of hydrogen column density (N_H), shown for plasma temperatures of 4 keV (blue), 6 keV (green), and 8 keV (red). The background gray histogram shows the normalized distribution of N_H values in our survey.

We use PIMMS to model the X-ray countrates with and without N_H absorption as a function of N_H values, assuming a 0.1-2.4 keV energy range (which is the one used by 2RXS). For this, we use the thermal bremsstrahlung model at different temperature values of $kT = 4, 6$ and 8 keV. Fig. 5.4 shows the ratio of absorbed to unabsorbed countrate values as a function of N_H for the different temperatures. The gray histogram shows the normalized distribution of N_H values within our survey. The countrate ratios for temperature values of 4 keV and 8 keV have an absolute difference of $\sim 1.6\%$ at $N_H = 3 \times 10^{19}$ cm^{-2} and $\sim 13\%$ at $N_H = 3 \times 10^{21}$ cm^{-2}. We use the ratios at a temperature $kT = 4$ keV to build an interpolator of the ratio given a N_H value.

The observed countrates of the RASS-MCMF clusters are indeed affected by the N_H absorption, however this effect would be applied to the intrinsic countrates, which are then subjected to the measurement noise (modeled using Poisson) as they are observed. This can be described as

$$\eta_{N_H} = \frac{\eta}{\mathrm{Corr}_{N_H}} \Rightarrow \hat{\eta}_{N_H} = \frac{\mathrm{Poisson}(\eta_{N_H} \times t_{\mathrm{exp}})}{t_{\mathrm{exp}}}, \tag{5.33}$$

where $\mathrm{Corr}_{N_H} = \frac{\eta(N_H)}{\eta(N_H=0)}$ is the N_H correction factor given a temperature and the N_H value in the sky at the position of the clusters.

Estimating N_H unabsorbed countrates is, therefor, not an easy task since we cannot estimate the value of the intrinsic N_H-absorbed countrates before they are observed (i.e., without Poisson noise). For our model, when estimating the expectation value of each cluster (the sum in Eq. 5.23), we correct the observed countrate $\hat{\eta}$ by interpolating the ratio of absorbed to unabsorbed countrates using the N_H value at the position of the cluster. This procedure can be described as

$$\hat{\eta} = \hat{\eta}_{N_H} \times \mathrm{Corr}_{N_H} \tag{5.34}$$

where $\hat{\eta}$ is the observed countrate corrected by the N_H absorption. We note that this will generate a bias, however we expect it to be negligible in comparison to the other sources of uncertainty.

5.3.2.3 Selection functions

In the case of the selection function in Eq. 5.23, Θ_s, this is separated for each observable, $\hat{\eta}$, $\hat{\lambda}$ and z, such that $\Theta_s(\hat{\eta}, z, \hat{\lambda}) = C(\hat{\eta}, H)\Theta(\hat{\lambda} > \hat{\lambda}\mathrm{min}(z))$, where $C(\hat{\eta}, H)$ is the X-ray selection function with H being cluster parameters that depend on the survey (such as exposure time and background). The X-ray selection on the 2RXS catalog is done using the existence likelihood parameter, EXI_ML > 6.5, whereas our X-ray observable is the countrate $\hat{\eta}$. We use a scaling relation to relate $\hat{\eta}$ to EXI_ML as

$$\begin{aligned}\left\langle \ln(\hat{\eta}|\mathrm{EXI\_ML}, t_{\mathrm{exp}}, \mathrm{bkg}) \right\rangle =& A_{\mathrm{C}} + B_{\mathrm{C}} \ln \mathrm{EXI\_ML} \\ &+ C_{\mathrm{C}} \ln t_{\mathrm{exp}} + D_{\mathrm{C}} \ln \mathrm{Bkg},\end{aligned} \tag{5.35}$$

where t_{exp} and Bkg are the exposure time and background countrate (Bkg = Background/t_{exp}), respectively. $s_{\ln\hat{\eta}}$ is the log-normal intrinsic scatter. See § 4.6.1.2 for further details on the parameters that describe the EXI_ML-$\hat{\eta}$ relation.

We model the completeness function, $C(\hat{\eta}, H)$, as an error function (erf) given by

$$C(\hat{\eta}, z) = \frac{1}{2}\left[1 + \mathrm{erf}\left(\frac{\ln\hat{\eta} - (\ln\hat{\eta}_{50} + \delta\ln\hat{\eta}_{50})}{\sqrt{2}s_{\ln\hat{\eta}}}\right)\right], \tag{5.36}$$

where $\ln\hat{\eta}_{50}$ is the countrate at EXI_ML = 6.5 and $\delta\ln\hat{\eta}_{50}$ is a free parameter that captures any deviations from the expected threshold.

In the case of $\hat{\lambda}$, the selection function, $\Theta(\hat{\lambda} > \hat{\lambda}\mathrm{min}(z))$, is a Heaviside step function defined as

$$\begin{cases} 1 & \text{if } \hat{\lambda} \geq \hat{\lambda}\mathrm{min}(z) \\ 0 & \text{otherwise,} \end{cases}$$

which uses the same minimum richness as a function of redshift, $\lambda_{\mathrm{min}}(z)$, that is used to create the RASS-MCMF cosmology sample ($f_{\mathrm{cont}} < 0.06$; see Chapter 4).

5.3.2.4 Integrated total number of clusters: $N_{\rm tot}$

The integral term on Eq. 5.23 gives us the expected total number of clusters. However, many parts of our likelihood depend on sky values such as the exposure time (to model the measurement uncertainty of the countrates), the background (for the completeness function) or the N_H column density (to quantify the countrate absorption). Thus, to compute the total number of clusters we divide the sky using the exposure time distribution shown in Fig. 5.3 and integrate the total number of clusters within each exposure time bin.

First, to account for the measurement uncertainty of the countrates, we compute 7 HOFs at different exposure time values that map the entire distribution of exposure times in our survey. These exposure times are shown as red lines in Fig. 5.3. We chose 7 as a compromise between interpolation accuracy and speed. Then, we iterate the exposure time bins and interpolate a HOF, using the PDF to account for the sky area that the bin covers. For each bin, we account for variations in the completeness function and the countrate absorption at different N_H values.

The variation of the completeness function within a exposure time bin is no different than what we did in § 4.6.1.2. Given that the completeness function depends on both the exposure time and the background, we study the variation of $\ln \hat{\eta}_{50}$, within an exposure time bin, by fixing each of them. Within an exposure time bin, the exposure time variation has little effect on $\ln \hat{\eta}_{50}$, whereas the background variation has a noticeable effect. Thus, for each bin we estimate a Bkg-weighted average completeness function $C_{\rm HOF}(\hat{\eta}, H)$ (see Eq. 4.15).

Similarly, we weight the total number of clusters within each exposure time bin using the N_H values. For each bin, we use the distribution of N_H values to interpolate the ratio of the absorbed to unabsorbed countrate. We use the ratios to boost the values of $\ln \hat{\eta}_{50}$, thus effectively reducing the number of clusters. Finally, we estimate a N_H-weighted average $N_{\rm tot}$ for each exposure time bin. The final $N_{\rm tot}$ is just the sum over all bins.

5.3.3 Lensing likelihood

The lensing likelihood, $\mathcal{L}_{\rm lensing}$, is estimated following Singh et al. (2024) by averaging the cluster matter profiles, re-scaled by the cluster radius (R/R_{200}), in bins of $\hat{\lambda}$, $\hat{\eta}$ and z. In this subsection I will give a summary of the method presented in Singh et al. (2024) and the final likelihood estimation for our cluster sample.

5.3.3.1 Re-scaled matter profiles

We start by reminding ourselves of the tangential reduced shear profile, $\gamma_t(R)$, induced by a projected mass distribution $\Sigma(R)$, from Eq. 5.7. In this case, the critical surface mass density depends on the source-lens system, $\Sigma_{\rm crit}(z_s, z_l)$, with z_s and z_l the source and lens redshift, respectively.

The numerator on Eq. 5.7, $\Delta\Sigma(R) = \langle\Sigma(< R)\rangle - \Sigma(R)$, is composed of the mean surface density enclosed within a radius $< R$ and the surface mass density at the projected radius, which

are given by

$$\langle \Sigma(< R) \rangle = \frac{2}{R^2} \int_0^R \mathrm{d}R' R' \Sigma(R') \tag{5.37}$$

and

$$\Sigma(R) = \int_{-\infty}^{\infty} \mathrm{d}\chi \rho \left(\sqrt{R^2 + \chi^2} \right), \tag{5.38}$$

respectively, with χ being the line-of-sight comoving distance. The density, $\rho(r)$, is taken to follow a NFW profile (Eq. 5.9). To average the cluster mass profiles, it is necessary to re-scale the projected matter profile by the clusters sizes, so that

$$\widetilde{\Delta\Sigma}\left(\frac{R}{R_{200_c}, z}\right) = \frac{\Delta\Sigma\left(\frac{R}{R_{200c}, z}\right)}{R_{200_c}\rho_{\mathrm{crit}}(z)} \tag{5.39}$$

where R_{200_c} is the 3D radius that encloses the mass M_{200_c}, estimated as

$$R_{200_c} = \left[\frac{3M_{200_c}}{200\rho_c(z)4\pi} \right]^{1/3}. \tag{5.40}$$

The value of R_{200_c} is estimated individually for each cluster following the weak-lensing mass posterior weighted by the observable-mass relations and by the measurement uncertainties in the observables (see Eq. 5.46 in §5.3.3.3).

For the stacking method, we average the matter profiles of individual clusters within j bins of (R/R_{200_c}), so that

$$\left\langle \widetilde{\Delta\Sigma}\left(\frac{R}{R_{200_\mathrm{c}}}\right)_j \right\rangle = \frac{\sum_{k,b,i} \Sigma_{\mathrm{crit},k,b} w_{k,b} \widetilde{\mathcal{W}}^{\mathrm{s}}_{k,b,i} e_{\mathrm{t},k,b,i} / \left(\rho_{\mathrm{crit},k} R_{200c,k}(1 - f_{\mathrm{cl},k,b})\right)}{\sum_{k,b,i} w_{k,b} \widetilde{\mathcal{W}}^{\mathrm{s}}_{k,b,i} (R_{\gamma_\mathrm{t},i} + R_{\mathrm{sel}})} \tag{5.41}$$

where the sum goes over each cluster k, each redshift tomographic bin b and for each source galaxy i. $w_{k,b} = \Sigma^{-1}_{\mathrm{crit}}, k, b'$ are tomographic bin weights, f_{cl} is the cluster member contamination (an estimate of the number of cluster members mistakenly taken for source galaxies, Gruen et al., 2014; Melchior et al., 2017; Paulus, 2021; Bocquet et al., 2023), $e_{\mathrm{t},k,b,i}$ is the ellipticity and $R_{\gamma_t,i}$ is the shear response of the i-th source galaxy. The selection response, R_{sel}, takes into account that lensing sources are selected based on their sheared observations (Bocquet et al., 2023). $\widetilde{\mathcal{W}}^{\mathrm{s}}_{k,b,i}$ is the re-scaled source weight, defined as

$$\widetilde{\mathcal{W}}^{\mathrm{s}}_{k,b,i} = w^{\mathrm{s}}_i \frac{\rho^2_{\mathrm{crit},k} R^2_{200c,k} (1 - f_{\mathrm{cl},k,b})^2}{\Sigma^2_{\mathrm{crit},k,b}} \tag{5.42}$$

where $w^s_i = \Sigma^{-1}_{\mathrm{crit},k,b'}$ is the individual source weight. The corresponding uncertainty for a radial bin $(R/R_{200_c})_j$ is given by

$$\widetilde{\sigma}^2_j = \frac{\sum_{k,b,i} \left(\Sigma_{\mathrm{crit},k,b} w_{k,b} \widetilde{\mathcal{W}}^{\mathrm{s}}_{k,b,i} \sigma_{\mathrm{eff},b} / \left(\rho_{\mathrm{crit},k} R_{200c,k}(1 - f_{\mathrm{cl},k,b})\right) \right)^2}{(\sum_{k,b,i} w_{k,b} \widetilde{\mathcal{W}}^{\mathrm{s}}_{k,b,i})^2} \tag{5.43}$$

5.3.3.2 Weak lensing mass - true halo mass scaling relation

A weak lensing mass ($M_{\rm WL}$) - mass scaling relation is defined to characterize the bias that may arise from the averaging of the cluster profiles. The scaling relation is defined as

$$\left\langle \ln\left(\frac{M_{\rm WL}}{M_{\rm piv}}\right)\right\rangle = \ln M_{\rm WL_0} + M_{\rm WL_M}\ln\left(\frac{M_{200_c}}{M_{\rm piv}}\right) + M_{\rm WL_z}\ln\left(\frac{1+z}{1+z_{\rm piv}}\right), \tag{5.44}$$

with $\ln M_{\rm WL_0}$ being the bias at $M_{200_c} = M_{\rm piv}$, $M_{\rm WL_M}$ is the mass trend and $M_{\rm WL_z}$ is the redshift trend. The uncertainty, $\sigma_{\ln \rm WL}$ is modeled as

$$\ln\sigma^2_{\ln \rm WL} = \ln\sigma^2_{\ln \rm WL_0} + \ln\sigma^2_{\ln \rm WL_M}\ln\left(\frac{M_{200_c}}{M_{\rm piv}}\right) + \ln\sigma^2_{\ln \rm WL_z}\ln\left(\frac{1+z}{1+z_{\rm piv}}\right). \tag{5.45}$$

The parameters of the $M_{\rm WL}$-mass relation are constrained using the *Magneticum* Pathfinder suite of cosmological hydrodynamical simulations[1] and the Illustring-TNG300-1 hydrodynamical simulation (Naiman et al., 2018; Marinacci et al., 2018; Pillepich et al., 2018; Nelson et al., 2018; Springel et al., 2018).

5.3.3.3 Single cluster re-scaled matter profile

There are several quantities which contribute to the posterior mass distribution $P(M|\hat{\eta},\hat{\lambda},z,\mathbf{p})$, such as the intrinsic scatters of the observable-mass relations, the measurement noise on the observables, the uncertainties on the observable-mass relation and cosmological parameters, etc. Even the $M_{\rm WL} - M_{200c}$ relation has an intrinsic bias and uncertainties on its parameters (Eq. 5.44 and 5.45).

Thus, the single cluster re-scaled matter profile, which accounts for all the aforementioned factors, can be written as

$$P(\widetilde{\Delta\Sigma}_k|\widetilde{\Delta\Sigma}_{\rm mod},z) = \int {\rm d}M_{\rm WL}P_k(M_{\rm WL}|\hat{\eta},\hat{\lambda},z,\mathbf{p})P(\widetilde{\Delta\Sigma}_k(M_{\rm WL})|\widetilde{\Delta\Sigma}_{\rm mod},z) \tag{5.46}$$

where $\widetilde{\Delta\Sigma}_{\rm mod}$ is the lensing model profile within a radial bin. $P_k(M_{\rm WL}|\hat{\eta}_k,\hat{\lambda},z,\mathbf{p})$ is the weak lensing mass posterior of the k-th cluster, estimated from the observables $\hat{\eta}$, $\hat{\lambda}$ and z using the Bayes theorem, accounting for both intrinsic and measurement uncertainties in the observables (neglecting the uncertainties on z), so that

$$P(M_{\rm WL}|\hat{\eta},\hat{\lambda},z,\mathbf{p}) = \frac{\iiint {\rm d}M{\rm d}\lambda{\rm d}\eta P(\hat{\lambda}|\lambda)P(\hat{\eta}|\eta)P(\eta,\lambda,M_{\rm WL}|M,z,\mathbf{p})P(M|z,\mathbf{p})}{\iiint {\rm d}M{\rm d}\lambda{\rm d}\eta P(\hat{\lambda}|\lambda)P(\hat{\eta}|\eta)P(\eta,\lambda|M,z,\mathbf{p})P(M|z,\mathbf{p})} \tag{5.47}$$

where $P(M|z,\mathbf{p})$ is the HMF and $P(M_{\rm WL}|M,z,\mathbf{p})$ is modelled as a lognormal distribution with mean and variance given by Eq. 5.44 and 5.45, respectively.

[1] http://www.magneticum.org/

$\widetilde{\Delta\Sigma}_k$ and $P(\widetilde{\Delta\Sigma}_k(M_{\rm WL})|\widetilde{\Delta\Sigma}_{\rm mod}, z)$ are the re-scaled matter profile and it's likelihood, respectively. The lensing likelihood for an averaged, re-scaled matter profile is calculated as a product of independent Gaussian probabilities, each representing the probability of observing the given matter profile within individual radial bins according to the model. Since the re-scaled matter profile $\widetilde{\Delta\Sigma}(R/R_{200_c})$ is influenced by cosmological parameters as well as the masses and radii of the clusters involved, the likelihood must be adjusted to incorporate these dependencies. The lensing likelihood for a re-scaled matter profile (defined in Eq. 5.41) is given by

$$P(\widetilde{\Delta\Sigma}(\boldsymbol{M}_{200_c}, \boldsymbol{p})|\widetilde{\Delta\Sigma}_{\rm mod}, z) = P_{\rm G} \prod_j \left(\left| \frac{\partial \widetilde{\Delta\Sigma}_j}{\partial \boldsymbol{e}_{t,j}} \left(\frac{\partial \widetilde{\Delta\Sigma}_j}{\partial \boldsymbol{e}_{t,j}} \right)^T \right|^{1/2} \right), \tag{5.48}$$

where the product goes over j radial bins of (R/R_{200_c}), $\boldsymbol{e}_{t,j} = [e_t^1, ..., e_t^n]$ contains the ellipticities of n sources within a radial bin and $P_{\rm G}$ is defined as

$$P_{\rm G} = \prod_j \left(\sqrt{2\pi}\sigma_{\widetilde{\Delta\Sigma},j} \right)^{-1} \exp\left[-\frac{1}{2} \left(\frac{\widetilde{\Delta\Sigma}_j - \widetilde{\Delta\Sigma}_{{\rm mod},j}}{\sigma_{\widetilde{\Delta\Sigma}_j}} \right)^2 \right], \tag{5.49}$$

where $\sigma_{\widetilde{\Delta\Sigma},j}$ is the re-scaled shape noise (Eq. 5.43). The transformation factor in Eq. 5.48 is calculated as

$$\left| \frac{\partial \widetilde{\Delta\Sigma}_j}{\partial \boldsymbol{e}_{t,j}} \left(\frac{\partial \widetilde{\Delta\Sigma}_j}{\partial \boldsymbol{e}_{t,j}} \right)^T \right|^{1/2} = \left[\left(\frac{\partial \widetilde{\Delta\Sigma}_j}{\partial \boldsymbol{e}^1_{t,j}} \right)^2, ..., \left(\frac{\partial \widetilde{\Delta\Sigma}_j}{\partial \boldsymbol{e}^n_{t,j}} \right)^2 \right]^{1/2}. \tag{5.50}$$

5.3.3.4 Stacked weak lensing likelihood

The stacked weak lensing likelihood, within a $\hat{\eta}$, $\hat{\lambda}$ and z bin, is given by the average lensing likelihood of the average re-scaled matter profile built from n clusters within said bin, such that

$$\begin{aligned} \mathcal{L}_{\rm bin} &= \langle P(\widetilde{\Delta\Sigma}(M_{{\rm WL}_1}, ..., M_{{\rm WL}_n}, \mathbf{p})|\widetilde{\Delta\Sigma}_{\rm mod}, z) \rangle \\ &= \int \cdots \int {\rm d}M_{{\rm WL}_1} ... {\rm d}M_{{\rm WL}_n} P(M_{{\rm WL}_1}|\hat{\eta}_1, \hat{\lambda}_1, z_1, \mathbf{p}) \times ... \\ &P(M_{{\rm WL}_n}|\hat{\eta}_n, \hat{\lambda}_n, z_n, \mathbf{p}) P(\widetilde{\Delta\Sigma}(M_{{\rm WL}_1}, ..., M_{{\rm WL}_n}, \mathbf{p})|\widetilde{\Delta\Sigma}_{\rm mod}, z) \end{aligned} \tag{5.51}$$

The final likelihood can be written as the product of the likelihood of the individual $\hat{\eta}$, $\hat{\lambda}$ and z bins as

$$\mathcal{L}_{\rm lensing} = \prod_{\rm bin=1}^{m} \mathcal{L}_{\rm bin} \tag{5.52}$$

5.3.4 Combined final likelihood

Combining all the previous parts we can rewrite Eq. 5.22 as

$$\begin{aligned}\ln \mathcal{L}(\mathbf{p}) = \sum_i \ln \left(C(\hat{\eta}, H) \frac{\mathrm{d}^3 N(\mathbf{p})}{\mathrm{d}\hat{\eta}\mathrm{d}\hat{\lambda}\mathrm{d}z} \right) \Bigg|_{i-\text{th cluster}} + \ln \mathcal{L}_{\text{lensing}} \\ - \iiint \mathrm{d}z\, \mathrm{d}\hat{\lambda}\, \mathrm{d}\hat{\eta}\, C_{\text{HOF}}(\hat{\eta}, H) \Theta(\hat{\lambda} > \hat{\lambda}_{\min}(z)) \frac{\mathrm{d}^3 N(\mathbf{p})}{\mathrm{d}\hat{\eta}\mathrm{d}\hat{\lambda}\mathrm{d}z},\end{aligned} \tag{5.53}$$

where $C(\hat{\eta}, H)$ and $C_{\text{HOF}}(\hat{\eta}, H)$ are the X–ray selection functions estimated as described in § 5.3.2.3 and § 5.3.2.4, respectively. Given that each cluster in our sample has an observed richness that is greater than the minimum richness at the clusters redshift ($\hat{\lambda}_i > \hat{\lambda}_{\min}(z_i)$), the first term on the right side of Eq. 5.53 doesn't need to explicitly account for the richness selection.

5.4 Mocks

5.4.1 Input values

In order to generate mocks that are similar to the RASS-MCMF cosmology sample, we did a weak-lensing mass calibration run on the real cluster catalog. For this we decided to follow the full analysis of Singh et al. (2024) and include $P(\hat{\lambda}|\hat{\eta}, z)$ in our likelihood, which is given by

$$P(\hat{\lambda}|\hat{\eta}, z) = \frac{\iiint \mathrm{d}M\mathrm{d}\lambda\mathrm{d}\eta P(\hat{\lambda}|\lambda) P(\hat{\eta}|\eta) P(\eta, \lambda|M, z, \mathbf{p}) P(M|z, \mathbf{p})}{\int_{\hat{\lambda}_{\min}}^{\infty} \mathrm{d}\hat{\lambda} \iiint \mathrm{d}M\mathrm{d}\lambda\mathrm{d}\eta P(\hat{\lambda}|\lambda) P(\hat{\eta}|\eta) P(\eta, \lambda|M, z, \mathbf{p}) P(M|z, \mathbf{p})} \tag{5.54}$$

Thus, the likelihood of the weak-lensing mass calibration run on RASS-MCMF was given by

$$\ln \mathcal{L}(\mathbf{p}) = \ln \mathcal{L}_{\text{lensing}} + \sum_i \ln P(\hat{\lambda}_i|\hat{\eta}_i, z_i, \mathbf{p}) \tag{5.55}$$

where the sum runs over all clusters, whereas the lensing information is only available for clusters within the DES region.

Since our interest was on the values of the parameters of the observable-mass relations, we used the best fit values of the cosmological parameters from the latest *Planck* run: $\Omega_m = 0.3166 \pm 0.0084$ and $\ln 10^{10} \mathrm{A_s} = 3.045 \pm 0.016$ (Planck Collaboration et al., 2020b). We adopted a Gaussian prior with a 1σ width equivalent to twice the uncertainties of the parameters. To test for biases in our parameters because of this choice, we created a mock and run it with and without priors on the cosmological parameters. The results showed a negligible discrepancy between the two runs, implying that the mass calibration has little to no dependency on the adopted cosmology.

We also applied different redshift cuts for our mass calibration run, finding a redshift trend on the intrinsic scatter of the $\eta - M_{200}$ relation, $\sigma_{\ln\eta}$, where the higher the low-redshift threshold, the lower the intrinsic scatter. To account for this, we added a redshift dependency on the log-normal intrinsic scatter of η (Eq. 5.28). The best-fit values, which we use as inputs for our mocks, are

reported in the second column of Table 5.1. We note, however, that the run we did in the real data was without accounting for the N_H correction, which yield a value of $\ln \eta_0 = -3.38$. Since the N_H correction effectively lowers the countrate values of the clusters, we increased the value to $\ln \eta_0 = -2.57$, which kept the number of clusters in the mock on similar levels with respect to the real data.

5.4.2 Mock creation

RASS-MCMF-like cluster mocks were created following the procedure described in § 4.6.2 and in Bocquet et al. (2023). We use a Tinker et al. (2008) mass function to estimate the number of clusters within mass and redshift bins, assuming a survey solid angle of Ω_s ~22,300 deg^2. From this, we get the expected theoretical total number number of clusters within a bin, which is given by

$$\langle N(M,z)\rangle = \frac{\mathrm{d}n(\mathbf{p},M,z)}{\mathrm{d}M}\frac{\mathrm{d}V(\mathbf{p},z)}{\mathrm{d}z}\mathrm{d}M\mathrm{d}z, \tag{5.56}$$

where the first term represents the HMF and the second term is the survey solid angle dependent differential volume. Then, for each mass and redshift bin we do a Poisson realization of the expected number of clusters to create clusters with the same true mass, M_{200_c}, and redshift.

For each of the clusters we use the observable-mass relations to estimate η and λ. We apply the intrinsic scatter to each observable by drawing η and λ from a normal distribution with scatter given by $\sigma_{\ln\eta}$ and $\sigma_{\ln\lambda}$, respectively. Additionally, for each cluster we randomly assign a sky position to get the exposure time, background and N_H values. Using N_H, we use the values estimated in §5.3.2.2 to interpolate the expected absorption on η. Then, using the exposure time, we estimate the number of X-ray photons from this absorbed countrate and draw from a Poisson distribution to get the observed counts, which are then divided by the exposure time to get $\hat{\eta}$. Using $\hat{\eta}$, the exposure time and background, we estimate a value for EXI_ML by applying the reverse scaling, and apply scatter by drawing EXI_ML from a normal distribution with $\sigma = s_{\ln\hat{\eta}}$. We apply a first cut by removing sources where EXI_ML < 6.5. Finally, we estimate $\hat{\lambda}$ by drawing λ from a normal distribution with $\sigma = \sqrt{\lambda}$ and apply a second cut by removing sources where $\hat{\lambda} < \hat{\lambda}_{\mathrm{min}}(z)$, given the cluster's redshift, using the $\hat{\lambda}_{\mathrm{min}}$ values of the RASS-MCMF cosmology sample.

Similarly to previous works (e.g. Bocquet et al., 2024; Ghirardini et al., 2024; Singh et al., 2024) we will use photometric data from the first three years of observations of the Dark Energy Survey (DES-Y3; Sevilla-Noarbe et al., 2021; Gatti et al., 2021) for the weak-lensing mass calibration. Thus, to mimic the expected number of clusters found within the DES area (~5000 deg^2) we randomly select ~20% of the clusters in our mock to create mock weak lensing following the method from Singh et al. (2024). First, we assume a source galaxy density of 6 arcmin^{-2} and estimate the total source galaxies for each cluster by calculating the area on the sky within $10h^{-1}$ Mpc from the cluster's center. This is done accounting for the cluster's redshift and assuming the redshift distribution of the DES-Y3 data for each redshift tomographic bin. The number of total source galaxies are the same for each tomographic bin. Weights are assigned to each source galaxy by randomly drawing source weights from real DES data. Then, for each i source

we randomly assign a distance to the cluster center, R_i, which we then use to calculate $\widetilde{\Delta\Sigma}$ by normalizing by R_{200_c} of the corresponding cluster. For each tomographic bin, we add cluster member contamination using recent measurements in DES data. Finally, we use $\widetilde{\Delta\Sigma}$ to estimate g_t and use a scatter of $\sigma_{\rm eff} = 0.3$, which is the effective shape noise for DES data (Gatti et al., 2021), to randomly draw a new value of g_t following a normal distribution.

5.5 Validation on Mocks

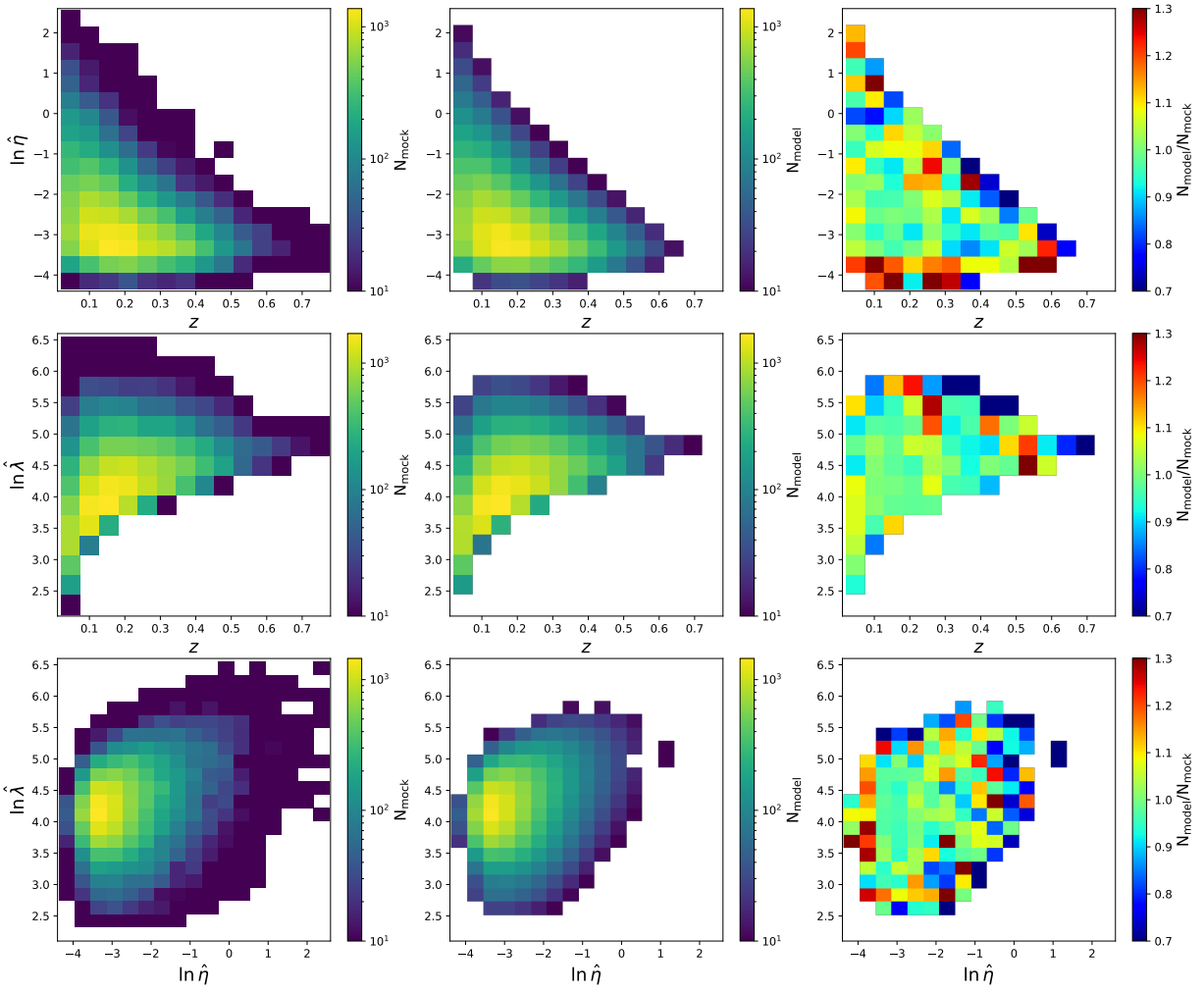

Figure 5.5: Expected versus observed number of clusters within an exposure time bin for a survey with 150× the sky area of RASS-MCMF. *Left:* Observed number of clusters from the mock. *Middle:* Expected number of clusters by integrating our model, for the same parameters used to create the mock. *Right:* Ratio of expected over observed number of clusters. The total number of clusters for any combination of two parameters is ∼28,000.

To validate the model we use to describe our data, we create a mock with 150× the survey area

and compare the predicted total number of clusters within an exposure time slice with that of the mock. The results are shown in Fig. 5.5, where the first and second columns show the number of clusters for the mocks and from integrating our model, respectively, within redshift, countrate, or richness bins. The last column shows the ratio between the number of expected over the number of observed clusters. The total number of clusters within this exposure time bin is ~28,000, with the difference between the expected and the observed number of clusters being less than 0.5σ for any combination of observables, where $\sigma = \sqrt{N_{\rm tot}}$.

Table 5.1: Input and posteriors of the observable-mass relations and cosmology parameters for the new likelihood (Eq. 5.53), both from an abundance only run and adding the weak lensing mass calibration likelihood.

Param.	Mock Input	Prior	Posterior	
			Abundance	Abundance + Mass cal.
$\lambda - M_{200} - z$ relation (Eq. 5.25)				
$\ln \lambda_0$	4.25	$\mathcal{U}(4.06, 4.41)$	$4.219^{+0.089}_{-0.099}$	4.275 ± 0.047
λ_M	1.03	$\mathcal{U}(0.7, 1.3)$	$1.009^{+0.038}_{-0.047}$	1.026 ± 0.032
λ_z	-2.6	$\mathcal{U}(-4, -1)$	$-2.67^{+0.18}_{-0.23}$	-2.71 ± 0.19
$\sigma_{\ln\lambda}$	0.103	$\mathcal{U}(0.05, 0.8)$	$0.131^{+0.046}_{-0.058}$	$0.087^{+0.075}_{-0.047}$
$\eta - M_{200} - z$ relation (Eq. 5.26)				
$\ln \eta_0$	-2.57	$\mathcal{U}(-4, -1)$	$-2.57^{+0.16}_{-0.18}$	-2.504 ± 0.091
η_M	1.85	$\mathcal{U}(1.2, 2.6)$	$1.804^{+0.076}_{-0.093}$	$1.817^{+0.050}_{-0.068}$
η_{M-z}	-0.7	$\mathcal{U}(-2.7, 1.3)$	-0.65 ± 0.17	-0.67 ± 0.17
η_z	0.08	$\mathcal{U}(-5, 3)$	0.28 ± 0.28	0.19 ± 0.28
$\sigma_{\ln\eta_0}$	0.651	$\mathcal{U}(0.05, 1.0)$	$0.650^{+0.048}_{-0.022}$	$0.676^{+0.035}_{-0.042}$
$\sigma_{\ln\eta_z}$	-2.01	$\mathcal{U}(-4, 0)$	$-2.31^{+0.39}_{-0.25}$	$-2.15^{+0.28}_{-0.21}$
Completeness function $C(\eta, z)$ (Eq. 5.36)				
$\delta \ln \hat{\eta}_{50}$	0.0	$\mathcal{U}(-1, 1)$	0.005 ± 0.013	0.003 ± 0.013
$s_{\ln\hat{\eta}}$	0.182	$\mathcal{U}(0.01, 0.46)$	0.187 ± 0.011	0.186 ± 0.011
Cosmology parameters				
$\Omega_{\rm m}$	0.28	$\mathcal{U}(0.1, 0.5)$	$0.267^{+0.019}_{-0.029}$	$0.265^{+0.018}_{-0.023}$
h	0.7	$\mathcal{N}(0.7, 0.04^2)$	0.702 ± 0.031	0.707 ± 0.031
$\ln 10^{10} A_{\rm s}$	3.001	$\mathcal{U}(2, 5)$	$3.14^{+0.31}_{-0.27}$	$3.08^{+0.24}_{-0.22}$
σ_8	0.78	-	0.800 ± 0.042	$0.787^{+0.021}_{-0.016}$

We run two likelihoods assuming a ΛCDM cosmology: abundance only and abundance with the weak-lensing mass calibration. All parameters are given a flat uniform prior, with a range wide enough to enclose at least 3σ, except for h which has a Gaussian prior with mean 0.7 and variance $(0.04)^2$. We note that, although our model allows for correlated scatter between the observables, for our current validation we have fixed $\rho_{\eta,\lambda} = 0$ for both the mocks and the model. Fig. 5.6 shows the posteriors of all the parameters we fit for the two likelihoods. Gray dashed lines mark the input values for a particular parameter. All the input values are within the 1σ region (darker colors) for both runs. The mock input values, prior ranges and posteriors

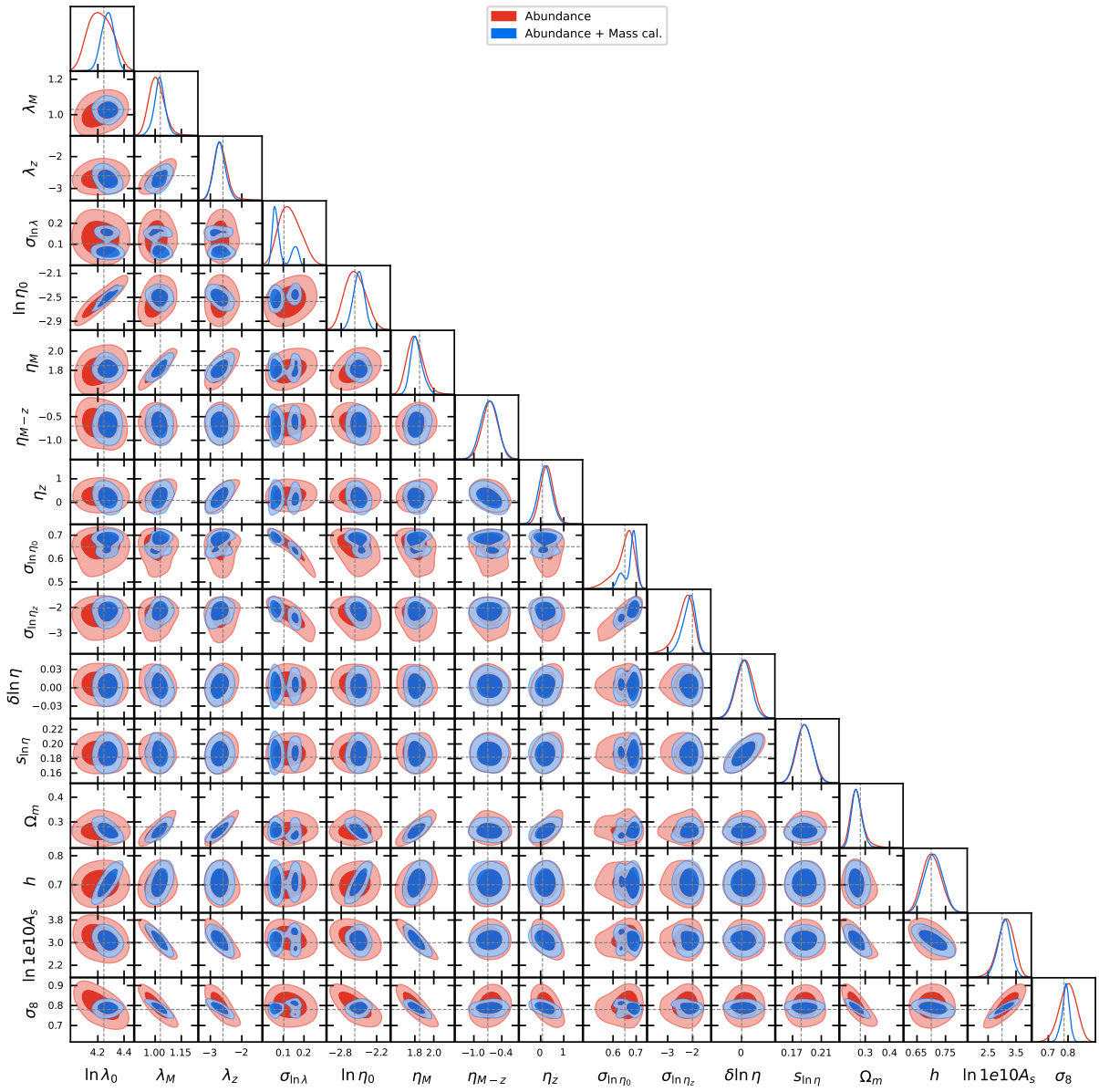

Figure 5.6: Posteriors of all the parameters in the abundance-only (red) and abundance with weak-lensing mass calibration (blue) runs on a RASS-MCMF-like mock (68% and 95% credible regions). The input values used to create the mock are shown as gray-dashed lines.

(with 1σ uncertainties) for both runs are shown in Table. 5.1. The impact of adding the mass calibration likelihood in the constraining power of the observable-mass relation parameters is noticeable in most of them, with some redshift evolution parameters not varying much (λ_z, η_{M-z} and η_z), except for the redshift evolution of the countrate intrinsic scatter, $\sigma_{\ln\eta_z}$. Given that the mass calibration likelihood does not require the completeness function, it is expected that the parameters $s_{\ln\eta}$ and $\delta\ln\eta$ are basically unchanged. On the other hand, the uncertainties in the

cosmological parameters are reduced, with the uncertainty in σ_8 shrinking by a factor of ~2-3. Interestingly, both $\sigma_{\ln\lambda}$ and $\sigma_{\ln\eta}$ exhibit a clear bimodal distribution when the mass calibration likelihood is included, with the input parameter positioned between the two peaks. This behavior may result from the specific seed used to generate the mock, suggesting that running additional mock realizations could help clarify the origin of this bimodality. Nevertheless, the current results indicate that the test can be considered successful, and we consider the pipeline to be validated.

A direct comparison of how we have improved our method and likelihood for a RASS-MCMF-like mock can be made with the results shown in Chapter 4. In the previous chapter we did not apply the mass calibration and instead used priors on the observable-mass relation parameters using the posteriors of the MARD-Y3 cluster sample, from a weak-lensing mass calibration only run (Singh et al., in prep.). The posteriors we got for a ΛCDM cosmology were $\Omega_m = 0.287 \pm 0.028$ and $\sigma_8 = 0.776 \pm 0.031$, meaning that we get a similar uncertainty on Ω_m but a factor of 1.5-2 improvement on the uncertainty of σ_8 when we compare to our latest results.

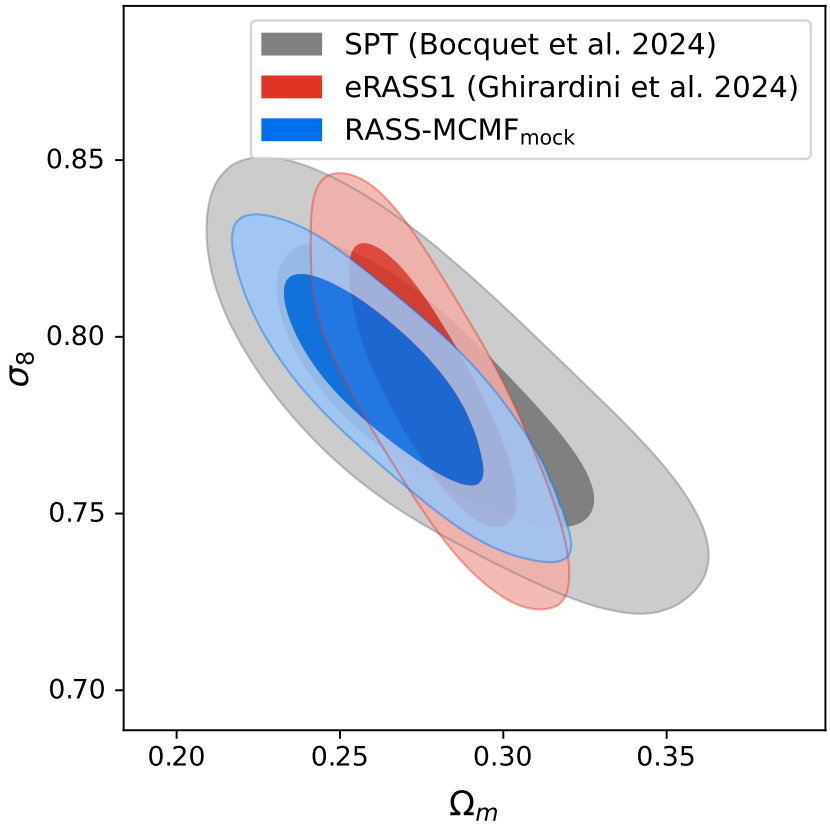

Figure 5.7: Constrains on Ω_m and σ_8 (68% and 95% credible regions) in a ΛCDM universe from a mock of RASS-MCMF-like clusters (blue contour). For comparison we show the latest constraints from SPT (gray contours; Bocquet et al., 2024) and eROSITA (red contours; Ghirardini et al., 2024) shifted to the input cosmology of the mock run.

We can also compare the constraining power of the cosmological parameters to that of the latest cluster cosmology results from eRASS1 (Ghirardini et al., 2024) and SPT (Bocquet et al.,

2024). Fig. 5.7 shows the contours in the Ω_m versus σ_8 space, where the contours from eRASS1 and SPT are shifted so that the centers match the center of the RASS-MCMF mock run. From our results, we expect the RASS-MCMF catalog, which is based on the ∼35 year old ROSAT satellite, to provide competitive results, being second in terms of constraining Ω_m and potentially the best out of those three surveys in constraning σ_8. However, it is important to note that both the eRASS1 and SPT analysis fit for other parameters that affect the results on Ω_m and σ_8, such as the sum of the neutrino masses, which is expected slightly increase the uncertainties (Bocquet et al., 2019).

5.6 Conclusions and outlook

In this chapter I have shown and explained the new method that will be applied to the RASS-MCMF cluster sample. The posteriors showing a remarkable agreement with the input values used is a sign that the likelihood and method, up to this point, are validated. The pipeline gives us promising results, with a constraining power of $\Omega_m = 0.265^{+0.018}_{-0.023}$ and $\sigma_8 = 0.787^{+0.021}_{-0.016}$, for a ΛCDM cosmology. For comparison, the latest SPT results report values of $\Omega_m = 0.286 \pm 0.032$ and $\sigma_8 = 0.817 \pm 0.026$, with eRASS1 estimating $\Omega_m = 0.29^{+0.01}_{-0.02}$ and $\sigma_8 = 0.88 \pm 0.02$ (Ghirardini et al., 2024), for the same ΛCDM cosmology. Our results suggest that the final run with RASS-MCMF is likely to yield competitive outcomes in comparison to the most recent analyses.

The pipeline is still lacking a few details that we are currently developing before testing with the real data, such as modeling the contamination of the cluster sample and accounting for correlated scatter between $\hat{\lambda}$ and $\hat{\eta}$. The latter is already being tested and has shown promising results so far. Fig. 5.8 shows an abundance only run for two mocks with different seeds and with a positive correlated scatter of $\rho_{\eta,\lambda} = +0.30$. For these runs we fixed the cosmology and two observable-mass relation parameters. We also didn't account for the N_H absorption, which is why the input value of $\ln \eta_0$ is different. All the input values are within the 1σ region of the posteriors, and so we consider this as a validation of our method. However, when using mocks with negative correlated scatter, $\rho_{\eta,\lambda} = -0.30$, one or more input values on each mock are over 2σ off from the best fit values, meaning that more tests are needed before we consider this a success. Similarly, we also want to allow for correlated scatter with $M_{\rm WL}$. Once some (or all) of the aforementioned improvements to the likelihood are included, we will start with blinded data runs on the cosmology cluster sample.

Alternatively, the method of using the sky distribution of exposure times to interpolate across different fields, while accounting for the solid angle covered by each exposure time bin, can be extended to other cluster samples. This approach is particularly useful for datasets with a non-uniform distribution of an observable, as it removes the need to assume fixed sky areas or constant values across multiple fields (Bocquet et al., 2023).

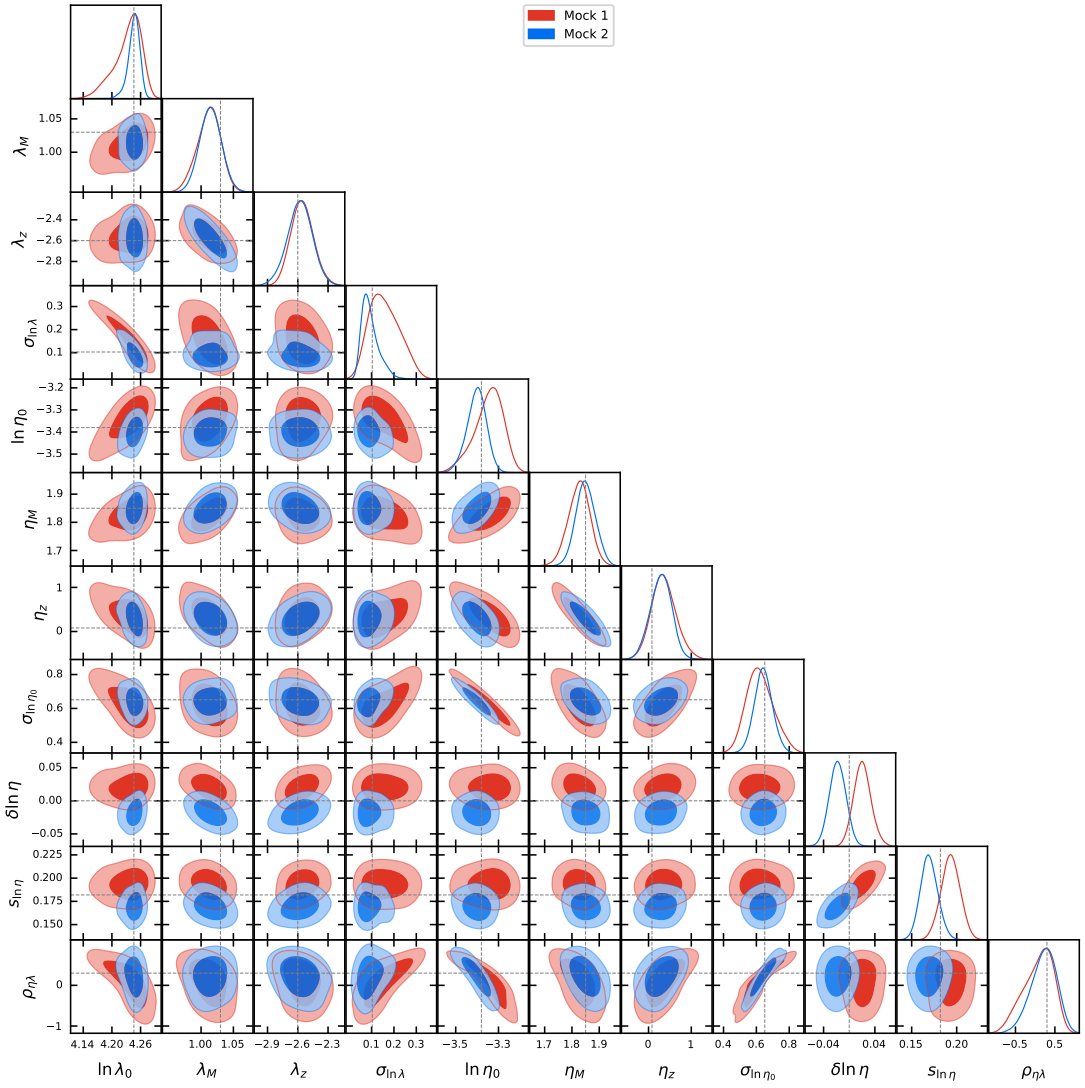

Figure 5.8: Posteriors of 9 observable-mass + 2 completeness parameters of an abundance-only run on two RASS-MCMF-like mocks with different seeds (68% and 95% credible regions). The input values used to create the mocks are shown as gray-dashed lines.

CHAPTER 6

CONCLUSIONS

Over recent decades, cosmological parameters have been constrained through a range of probes, from low-redshift ($z \leq 2$) probes such as galaxy weak lensing (e.g., Heymans et al., 2021; Abbott et al., 2022), galaxy clustering (e.g., Philcox & Ivanov, 2022; Yuan et al., 2022), and galaxy cluster number counts (e.g., Mantz et al., 2015; Planck Collaboration et al., 2016b; Bocquet et al., 2024), to predictions derived by extrapolating the CMB fluctuations ($z \sim 1100$) to the present day (e.g., Spergel et al., 2003; Planck Collaboration et al., 2020b). Although each of these probes has advanced our understanding of the cosmos, they do not always yield consistent results. In some cases, discrepancies between measurements, such as Ω_m, from low-redshift probes and the CMB exceed 4σ (Abdalla et al., 2022; Harnois-Déraps et al., 2024).

Historically, the tightest constraints have come from the CMB, with the latest results coming from the *Planck* mission. However, advancements in wide-field lensing and galaxy surveys, such as the Dark Energy Survey, along with larger catalogs of galaxy clusters with well-understood selection functions, have significantly improved the precision of local probes. As a result, cluster number counts are now reaching an accuracy in parameter estimation that begins to rival the precision of CMB anisotropy constraints.

Thus, with the advent of new surveys mapping large areas of the sky, the ability to generate and analyze galaxy cluster catalogs with thousands of clusters can prove highly valuable in constraining cosmological parameters with greater precision. In this thesis I started by studying structure growth by analyzing a cluster merger system. I followed up with the understanding of how galaxy cluster catalogs are created, moved to the creation of mock catalogs that follow the properties of a real catalog, and finally to the development of both a mathematical framework and pipeline that gives cosmological constraints for such a sample, with the pipeline being validated with the created mocks. I will now summarize the key points and future prospects of the different works.

The detailed multi-wavelength analysis of the $z = 0.58$ merging galaxy cluster SPT-CL J0307-6225, helps in our understanding of galaxy evolution and structure growth. Our results suggest that galaxy population is indeed affected by the dynamical state of their clusters, were we found rare red-galaxies with signs of star formation. The detailed analysis of the masses

we presented suggests that dynamical mass estimators are affected by the dynamical state of the cluster, particularly for lower mass clusters. For this particular system we were granted observations with the Hubble Space Telescope (PI: Zenteno, A.) in order to constrain the masses of the substructures using weak-lensing. Preliminary results from this new dataset suggest that the blue population of galaxies, which we determined to be recently accreted, are part of galaxy groups with 2 or 3 members. Using our MUSE data we are analyzing the possibility of interaction between the galaxies belonging to said groups. If we are able to confirm such claims, then we would be witnessing the pre-processing of said galaxies before they fall into the cluster's potential.

The PSZ-MCMF catalog is a promising result in the area of cluster confirmation, were we used a candidate list created from *Planck* data down to S/N=3. Given the low signal-to-noise, the initial catalog contamination was ∼50%. By confirming cluster candidates using the MCMF algorithm, we created the PSZ-MCMF catalog (which has a final contamination of 10%), expanding the number of previously confirmed *Planck* clusters (within the same area) by a factor of 4. A subsequent analysis of the latest SPT-SZ cluster catalog (Bleem et al., 2015) with the same cluster confirmation algorithm, have improved the number of SPT clusters by a factor of 1.5 down from S/N=4.5 to S/N=4 (Klein et al., 2024a). On the other hand, the method we developed to estimate the candidate list contamination (and thus the final contaminaton) of the sample has been subsequently applied to newer cluster catalogs (Klein et al., 2024a,b). When constructing the PSZ-MCMF catalog, we relied on DES photometric data to confirm the clusters, limiting us to a sky area of ∼5,000 deg^2. With the recent 10th data release of the Legacy Survey now available, we can expand the catalog to the full extragalactic sky (∼25,000 deg^2), as was done for the RASS-MCMF catalog, potentially resulting in a cluster catalog five times larger.

Following this, RASS-MCMF is an all-sky, X-ray-selected and optically-confirmed cluster catalog with over 8,000 confirmed clusters with the same contamination as that of PSZ-MCMF. We made use of our physical knowledge to generate a mock catalog with properties similar to RASS-MCMF, including the exposure time distribution, optical and X-ray selections, background variations, among others. We then adapted the likelihood method used in previous analyses of the SPT cluster sample (Bocquet et al., 2023; Mazoun et al., 2024; Vogt et al., 2024), to model an X-ray cluster catalog. The forecast showed promising results for the RASS-MCMF catalog, however the pipeline still needed development and we also did not run the mass calibration likelihood analysis ourselves.

In our latest work, we modified and expanded the likelihood model to better fit the RASS-MCMF data. We included the weak-lensing likelihood following a stacking technique (Singh et al., 2024) and also improved the number counts model. Running our model in mocks show a factor of 2 improvement on the uncertainty of σ_8 compared to our previous forecast. Our final results are also competitive with the latest cluster cosmology results. Once we are satisfied with our pipeline we will be able to get the real constraints for the RASS-MCMF catalog. These constraints could be further tightened by incorporating additional datasets, such as HST data for weak-lensing mass calibration (e.g., Bocquet et al., 2024), or by combining with other cluster probes, like galaxy cluster clustering (e.g., Fumagalli et al., 2024).

Although the results from the RASS-MCMF cluster catalog look promising, the combination with other cluster catalogs which rely on different observables may give the best constraints yet. A combination of the MARD-Y3 (Klein et al., 2019) and SPT-SZ + SPTpol-500d catalogs

(Bleem et al., 2015; Klein et al., 2024a; Bleem et al., 2024) is done in Singh et al. (in prep), where they explore the constraining power on scaling-relation parameters of the combined X-ray and SZ selected samples by using weak-lensing mass calibration. Their results show a factor of 2-3 improvement on the uncertainties of the scaling relation parameters. Thus, combining RASS-MCMF with other samples, such as the newly released ACT-DR5 MCMF cluster catalog (Klein et al., 2024b) or an all-sky version of PSZ-MCMF, are exciting new projects to look forward to.

BIBLIOGRAPHY

Abbott, T., Abdalla, F. B., Aleksić, J., et al. (2016), *The Dark Energy Survey: more than dark energy - an overview*, MNRAS, 460(2), 1270

Abbott, T. M. C., Abdalla, F. B., Allam, S., et al. (2018), *The Dark Energy Survey: Data Release 1*, ApJS, 239(2), 18

Abbott, T. M. C., Adamów, M., Aguena, M., et al. (2021), *The Dark Energy Survey Data Release 2*, ApJS, 255(2), 20

Abbott, T. M. C., Aguena, M., Alarcon, A., et al. (2020), *Dark Energy Survey Year 1 Results: Cosmological constraints from cluster abundances and weak lensing*, Phys. Rev. D, 102(2), 023509

Abbott, T. M. C., Aguena, M., Alarcon, A., et al. (2022), *Dark Energy Survey Year 3 results: Cosmological constraints from galaxy clustering and weak lensing*, Phys. Rev. D, 105(2), 023520

Abdalla, E., Abellán, G. F., Aboubrahim, A., et al. (2022), *Cosmology intertwined: A review of the particle physics, astrophysics, and cosmology associated with the cosmological tensions and anomalies*, Journal of High Energy Astrophysics, 34, 49

Aihara, H., Arimoto, N., Armstrong, R., et al. (2018), *The Hyper Suprime-Cam SSP Survey: Overview and survey design*, PASJ, 70, S4

Allen, S. W., Edge, A. C., Fabian, A. C., et al. (1992), *Optical spectroscopy of the ROSAT X-ray brightest clusters.*, MNRAS, 259, 67

Applegate, D. E., von der Linden, A., Kelly, P. L., et al. (2014), *Weighing the Giants - III. Methods and measurements of accurate galaxy cluster weak-lensing masses*, MNRAS, 439(1), 48

Arnaud, M., Pratt, G. W., Piffaretti, R., et al. (2010), *The universal galaxy cluster pressure profile from a representative sample of nearby systems (REXCESS) and the Y_{SZ} - M_{500} relation*, A&A, 517, A92

Bacon, R., Accardo, M., Adjali, L., et al. (2012), *News of the MUSE*, The Messenger, 147, 4

Bacon, R., Piqueras, L., Conseil, S., et al. (2016), *MPDAF: MUSE Python Data Analysis Framework*, Astrophysics Source Code Library, record ascl:1611.003

Bahcall, N. A. and Bode, P. (2003), *The Amplitude of Mass Fluctuations*, ApJ, 588(1), L1

Bahcall, N. A. and Fan, X. (1998), *The Most Massive Distant Clusters: Determining* Ω *and* σ_8, ApJ, 504(1), 1

Balogh, M. L., Morris, S. L., Yee, H. K. C., et al. (1999), *Differential Galaxy Evolution in Cluster and Field Galaxies at z~0.3*, ApJ, 527(1), 54

Barbosa, D., Bartlett, J. G., Blanchard, A., and Oukbir, J. (1996), *The Sunyaev-Zel'dovich effect and the value of {OMEGA}_0_.*, A&A, 314, 13

Bartelmann, M. (1996), *Arcs from a universal dark-matter halo profile.*, A&A, 313, 697

Bartelmann, M. (2010), *TOPICAL REVIEW Gravitational lensing*, Classical and Quantum Gravity, 27(23), 233001

Bartelmann, M. and Schneider, P. (2001), *Weak gravitational lensing*, Phys. Rep., 340, 291

Bayliss, M. B., Ruel, J., Stubbs, C. W., et al. (2016), *SPT-GMOS: A Gemini/GMOS-South Spectroscopic Survey of Galaxy Clusters in the SPT-SZ Survey*, ApJS, 227, 3

Bayliss, M. B., Zengo, K., Ruel, J., et al. (2017), *Velocity Segregation and Systematic Biases In Velocity Dispersion Estimates with the SPT-GMOS Spectroscopic Survey*, ApJ, 837(1), 88

Beers, T. C., Flynn, K., and Gebhardt, K. (1990), *Measures of location and scale for velocities in clusters of galaxies - A robust approach*, AJ, 100, 32

Bekki, K. (1999), *Group-Cluster Merging and the Formation of Starburst Galaxies*, ApJ, 510(1), L15

Bertin, E. (2011), *Automated Morphometry with SExtractor and PSFEx*, in *Astronomical Data Analysis Software and Systems XX*, edited by I. N. Evans, A. Accomazzi, D. J. Mink, A. H. Rots, volume 442 of *Astronomical Society of the Pacific Conference Series*, 435

Bertin, E. and Arnouts, S. (1996), *SExtractor: Software for source extraction.*, A&AS, 117, 393

Bilicki, M., Jarrett, T. H., Peacock, J. A., et al. (2014), *Two Micron All Sky Survey Photometric Redshift Catalog: A Comprehensive Three-dimensional Census of the Whole Sky*, ApJS, 210(1), 9

Bird, C. M. (1994), *Substructure in Clusters and Central Galaxy Peculiar Velocities*, AJ, 107, 1637

Blanchard, A. and Bartlett, J. G. (1998), *What does cluster redshift evolution reveal?*, A&A, 332, L49

Blanton, M. R., Bershady, M. A., Abolfathi, B., et al. (2017), *Sloan Digital Sky Survey IV: Mapping the Milky Way, Nearby Galaxies, and the Distant Universe*, AJ, 154(1), 28

Bleem, L. E., Bocquet, S., Stalder, B., et al. (2020), *The SPTpol Extended Cluster Survey*, ApJS, 247(1), 25

Bleem, L. E., Klein, M., Abbot, T. M. C., et al. (2024), *Galaxy Clusters Discovered via the Thermal Sunyaev-Zel'dovich Effect in the 500-square-degree SPTpol Survey*, The Open Journal of Astrophysics, 7, 13

Bleem, L. E., Stalder, B., de Haan, T., et al. (2015), *Galaxy Clusters Discovered via the Sunyaev-Zel'dovich Effect in the 2500-Square-Degree SPT-SZ Survey*, ApJS, 216(2), 27

Bocquet, S., Dietrich, J. P., Schrabback, T., et al. (2019), *Cluster Cosmology Constraints from*

the 2500 deg^2 SPT-SZ Survey: Inclusion of Weak Gravitational Lensing Data from Magellan and the Hubble Space Telescope, ApJ, 878(1), 55

Bocquet, S., Grandis, S., Bleem, L. E., et al. (2023), *SPT Clusters with DES and HST Weak Lensing. I. Cluster Lensing and Bayesian Population Modeling of Multi-Wavelength Cluster Datasets*, arXiv e-prints, arXiv:2310.12213

Bocquet, S., Grandis, S., Bleem, L. E., et al. (2024), *SPT Clusters with DES and HST Weak Lensing. II. Cosmological Constraints from the Abundance of Massive Halos*, arXiv e-prints, arXiv:2401.02075

Bocquet, S., Heitmann, K., Habib, S., et al. (2020), *The Mira-Titan Universe. III. Emulation of the Halo Mass Function*, ApJ, 901(1), 5

Bocquet, S., Saro, A., Mohr, J. J., et al. (2015), *Mass Calibration and Cosmological Analysis of the SPT-SZ Galaxy Cluster Sample Using Velocity Dispersion σ_v and X-Ray Y_X Measurements*, ApJ, 799(2), 214

Boese, F. G. (2000), *The ROSAT point spread functions and associates*, A&AS, 141, 507

Böhringer, H., Chon, G., Collins, C. A., et al. (2013), *The extended ROSAT-ESO flux limited X-ray galaxy cluster survey (REFLEX II) II. Construction and properties of the survey*, A&A, 555, A30

Böhringer, H., Schuecker, P., Guzzo, L., et al. (2004), *The ROSAT-ESO Flux Limited X-ray (REFLEX) Galaxy cluster survey. V. The cluster catalogue*, A&A, 425, 367

Böhringer, H., Voges, W., Huchra, J. P., et al. (2000), *The Northern ROSAT All-Sky (NORAS) Galaxy Cluster Survey. I. X-Ray Properties of Clusters Detected as Extended X-Ray Sources*, ApJS, 129(2), 435

Boller, T., Freyberg, M. J., Trümper, J., et al. (2016), *Second ROSAT all-sky survey (2RXS) source catalogue*, A&A, 588, A103

Bösch, B., Böhm, A., Wolf, C., et al. (2013), *Ram pressure and dusty red galaxies - key factors in the evolution of the multiple cluster system Abell 901/902*, A&A, 549, A142

Bower, R. G., Lucey, J. R., and Ellis, R. S. (1992), *Precision photometry of early-type galaxies in the Coma and Virgo clusters : a test of the universality of the colour-magnitude relation - II. Analysis.*, MNRAS, 254, 601

Brown, A. G., Vallenari, A., Prusti, T., et al. (2021), *Gaia early data release 3-summary of the contents and survey properties*, Astronomy & Astrophysics, 649, A1

Brown, R. L. and Gould, R. J. (1970), *Interstellar Absorption of Cosmic X Rays*, Phys. Rev. D, 1(8), 2252

Brüggen, M., van Weeren, R. J., and Röttgering, H. J. A. (2012), *Simulating the toothbrush: evidence for a triple merger of galaxy clusters*, MNRAS, 425(1), L76

Brunner, H., Liu, T., Lamer, G., et al. (2022), *The eROSITA Final Equatorial Depth Survey (eFEDS). X-ray catalogue*, A&A, 661, A1

Bruzual, G. and Charlot, S. (2003), *Stellar population synthesis at the resolution of 2003*, MNRAS, 344(4), 1000

Bulbul, E., Chiu, I.-N., Mohr, J. J., et al. (2019), *X-Ray Properties of SPT-selected Galaxy*

Clusters at 0.2 < z < 1.5 Observed with XMM-Newton, ApJ, 871, 50

Bulbul, E., Liu, A., Pasini, T., et al. (2022), *The eROSITA Final Equatorial-Depth Survey (eFEDS). Galaxy clusters and groups in disguise*, A&A, 661, A10

Bykov, A. M., Churazov, E. M., Ferrari, C., et al. (2015), *Structures and Components in Galaxy Clusters: Observations and Models*, Space Sci. Rev., 188, 141

Caldwell, N. and Rose, J. A. (1997), *The Butcher-Oemler Effect at Low Redshift: Spectroscopy of Five Nearby Clusters of Galaxies*, AJ, 113, 492

Cao, L., Wei, J. Y., and Hu, J. Y. (1999), *High X-ray-to-optical flux ratio RASS-BSC sources. I. The optical identification*, A&AS, 135, 243

Carlstrom, J. E., Ade, P. A. R., Aird, K. A., et al. (2011), *The 10 Meter South Pole Telescope*, PASP, 123(903), 568

Carlstrom, J. E., Holder, G. P., and Reese, E. D. (2002), *Cosmology with the Sunyaev-Zel'dovich Effect*, ARA&A, 40, 643

Carrasco, E. R., Gomez, P. L., Verdugo, T., et al. (2010), *Strong Gravitational Lensing by the Super-massive cD Galaxy in Abell 3827*, ApJ, 715(2), L160

Chan, B. M. Y., Broadhurst, T., Lim, J., et al. (2017), *Geometric Corroboration of the Earliest Lensed Galaxy at $z \simeq 10.8$ from Robust Free-form Modelling*, ApJ, 835, 44

Chandrasekhar, S. (1931), *The Maximum Mass of Ideal White Dwarfs*, ApJ, 74, 81

Chiu, I., Dietrich, J. P., Mohr, J., et al. (2016), *Detection of enhancement in number densities of background galaxies due to magnification by massive galaxy clusters*, MNRAS, 457, 3050

Chiu, I. N., Ghirardini, V., Liu, A., et al. (2022), *The eROSITA Final Equatorial-Depth Survey (eFEDS). X-ray observable-to-mass-and-redshift relations of galaxy clusters and groups with weak-lensing mass calibration from the Hyper Suprime-Cam Subaru Strategic Program survey*, A&A, 661, A11

Chiu, I. N., Klein, M., Mohr, J., and Bocquet, S. (2023), *Cosmological constraints from galaxy clusters and groups in the eROSITA Final Equatorial Depth Survey*, MNRAS

Clerc, N., Merloni, A., Zhang, Y.-Y., et al. (2016), *SPIDERS: the spectroscopic follow-up of X-ray-selected clusters of galaxies in SDSS-IV*, Monthly Notices of the Royal Astronomical Society, 463(4), 4490

Clowe, D., Bradač, M., Gonzalez, A. H., et al. (2006), *A Direct Empirical Proof of the Existence of Dark Matter*, ApJ, 648, L109

Clowe, D., Gonzalez, A., and Markevitch, M. (2004), *Weak-Lensing Mass Reconstruction of the Interacting Cluster 1E 0657-558: Direct Evidence for the Existence of Dark Matter*, ApJ, 604, 596

Collaboration:, D. E. S., Abbott, T., Abdalla, F. B., et al. (2016), *The Dark Energy Survey: more than dark energy – an overview*, MNRAS, 460(2), 1270

Colless, M., Dalton, G., Maddox, S., et al. (2001), *The 2dF Galaxy Redshift Survey: spectra and redshifts*, MNRAS, 328(4), 1039

Cortese, L., Gavazzi, G., Boselli, A., et al. (2004), *Multiple merging in the Abell cluster 1367*, A&A, 425, 429

Costanzi, M., Rozo, E., Rykoff, E. S., et al. (2019), *Modelling projection effects in optically selected cluster catalogues*, MNRAS, 482(1), 490

Crocce, M., Castander, F. J., Gaztañaga, E., et al. (2015), *The MICE Grand Challenge lightcone simulation - II. Halo and galaxy catalogues*, MNRAS, 453(2), 1513

Crossett, J. P., Pimbblet, K. A., Stott, J. P., and Jones, D. H. (2014), *Environments and morphologies of red sequence galaxies with residual star formation in massive clusters*, MNRAS, 437(3), 2521

Cruddace, R., Voges, W., Böhringer, H., et al. (2002), *The ROSAT All-Sky Survey: a Catalog of Clusters of Galaxies in a Region of 1 steradian around the South Galactic Pole*, ApJS, 140(2), 239

Cutri, R. M., Wright, E. L., Conrow, T., et al. (2013), *Explanatory Supplement to the AllWISE Data Release Products*, Explanatory Supplement to the AllWISE Data Release Products, by R. M. Cutri et al.

Dálya, G., Díaz, R., Bouchet, F. R., et al. (2022), *GLADE+ : an extended galaxy catalogue for multimessenger searches with advanced gravitational-wave detectors*, MNRAS, 514(1), 1403

Dawson, W. A. (2013), *The Dynamics of Merging Clusters: A Monte Carlo Solution Applied to the Bullet and Musket Ball Clusters*, ApJ, 772, 131

Dawson, W. A., Jee, M. J., Stroe, A., et al. (2015), *MC^2: Galaxy Imaging and Redshift Analysis of the Merging Cluster CIZA J2242.8+5301*, ApJ, 805(2), 143

Dawson, W. A., Wittman, D., Jee, M. J., et al. (2012), *Discovery of a Dissociative Galaxy Cluster Merger with Large Physical Separation*, ApJ, 747(2), L42

de Jong, J. T. A., Verdoes Kleijn, G. A., Kuijken, K. H., and Valentijn, E. A. (2013), *The Kilo-Degree Survey*, Experimental Astronomy, 35(1-2), 25

De Propris, R., Bremer, M. N., and Phillipps, S. (2016), *Morphological evolution of cluster red sequence galaxies in the past 9 Gyr*, MNRAS, 461(4), 4517

Delabrouille, J., Betoule, M., Melin, J. B., et al. (2013), *The pre-launch Planck Sky Model: a model of sky emission at submillimetre to centimetre wavelengths*, A&A, 553, A96

DeRose, J., Wechsler, R. H., Becker, M. R., et al. (2019), *The Buzzard Flock: Dark Energy Survey Synthetic Sky Catalogs*, arXiv e-prints, arXiv:1901.02401

Desai, S., Armstrong, R., Mohr, J. J., et al. (2012), *The Blanco Cosmology Survey: Data Acquisition, Processing, Calibration, Quality Diagnostics, and Data Release*, ApJ, 757(1), 83

Desai, V., Dalcanton, J. J., Aragón-Salamanca, A., et al. (2007), *The Morphological Content of 10 EDisCS Clusters at $0.5 < z < 0.8$*, ApJ, 660, 1151

Dey, A., Schlegel, D. J., Lang, D., et al. (2019), *Overview of the DESI Legacy Imaging Surveys*, AJ, 157(5), 168

Dietrich, J. P., Bocquet, S., Schrabback, T., et al. (2019), *Sunyaev-Zel'dovich effect and X-ray scaling relations from weak lensing mass calibration of 32 South Pole Telescope selected galaxy clusters*, MNRAS, 483(3), 2871

Doubrawa, L., Machado, R. E. G., Laganá, T. F., et al. (2020), *Simulations of gas sloshing induced by a newly discovered gas poor substructure in galaxy cluster Abell 1644*, MNRAS, 495(2),

2022

Dressler, A. (1980), *Galaxy morphology in rich clusters - Implications for the formation and evolution of galaxies*, ApJ, 236, 351

Dressler, A. and Gunn, J. E. (1983), *Spectroscopy of galaxies in distant clusters. II. The population of the 3C 295 cluster.*, ApJ, 270, 7

Dressler, A. and Shectman, S. A. (1988), *Evidence for substructure in rich clusters of galaxies from radial-velocity measurements*, AJ, 95, 985

Drlica-Wagner, A., Sevilla-Noarbe, I., Rykoff, E. S., et al. (2018), *Dark Energy Survey Year 1 Results: The Photometric Data Set for Cosmology*, ApJS, 235(2), 33

Duncan, K. J. (2022), *All-purpose, all-sky photometric redshifts for the Legacy Imaging Surveys Data Release 8*, Monthly Notices of the Royal Astronomical Society, 512(3), 3662

Ebeling, H., Edge, A. C., Bohringer, H., et al. (1998), *The ROSAT Brightest Cluster Sample - I. The compilation of the sample and the cluster log N-log S distribution*, MNRAS, 301(4), 881

Ebeling, H., Edge, A. C., and Henry, J. P. (2001), *MACS: A Quest for the Most Massive Galaxy Clusters in the Universe*, ApJ, 553(2), 668

Einasto, M., Deshev, B., Lietzen, H., et al. (2018), *Infalling groups and galaxy transformations in the cluster A2142*, A&A, 610, A82

Einstein, A. (1917), *Kosmologische Betrachtungen zur allgemeinen Relativitätstheorie*, Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften, 142–152

Einstein, A. and de Sitter, W. (1932), *On the Relation between the Expansion and the Mean Density of the Universe*, Proceedings of the National Academy of Science, 18(3), 213

Ester, M., Kriegel, H.-P., Sander, J., and Xu, X. (1996), *A Density-based Algorithm for Discovering Clusters a Density-based Algorithm for Discovering Clusters in Large Spatial Databases with Noise*, in *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, KDD'96, 226–231, AAAI Press

Evrard, A. E. (1989), *Biased Cold Dark Matter Theory: Trouble from Rich Clusters?*, ApJ, 341, L71

Evrard, A. E., Bialek, J., Busha, M., et al. (2008), *Virial Scaling of Massive Dark Matter Halos: Why Clusters Prefer a High Normalization Cosmology*, ApJ, 672(1), 122

Evrard, A. E., Metzler, C. A., and Navarro, J. F. (1996), *Mass Estimates of X-Ray Clusters*, ApJ, 469, 494

Fabricant, D., Cheimets, P., Caldwell, N., and Geary, J. (1998), *The FAST Spectrograph for the Tillinghast Telescope*, PASP, 110(743), 79

Ferragamo, A., Rubiño-Martín, J. A., Betancort-Rijo, J., et al. (2020), *Biases in galaxy cluster velocity dispersion and mass estimates in the small N_{gal} regime*, A&A, 641, A41

Ferrari, C., Maurogordato, S., Cappi, A., and Benoist, C. (2003), *Multiple merging events in Abell 521*, A&A, 399, 813

Finoguenov, A., Guzzo, L., Hasinger, G., et al. (2007), *The XMM-Newton Wide-Field Survey in the COSMOS Field: Statistical Properties of Clusters of Galaxies*, ApJS, 172(1), 182

Finoguenov, A., Rykoff, E., Clerc, N., et al. (2020), *CODEX clusters. Survey, catalog, and*

cosmology of the X-ray luminosity function, A&A, 638, A114

Flaugher, B., Diehl, H. T., Honscheid, K., et al. (2015), *The Dark Energy Camera*, AJ, 150(5), 150

Flesch, E. W. (2021), *The Million Quasars (Milliquas) v7.2 Catalogue, now with VLASS associations. The inclusion of SDSS-DR16Q quasars is detailed*, arXiv e-prints, arXiv:2105.12985

Fowler, J. W., Niemack, M. D., Dicker, S. R., et al. (2007), *Optical design of the Atacama Cosmology Telescope and the Millimeter Bolometric Array Camera*, Appl. Opt., 46(17), 3444

Frenk, C. S., White, S. D. M., Efstathiou, G., and Davis, M. (1990), *Galaxy Clusters and the Amplitude of Primordial Fluctuations*, ApJ, 351, 10

Freund, S., Czesla, S., Robrade, J., et al. (2022), *The stellar content of the ROSAT all-sky survey*, A&A, 664, A105

Friedmann, A. (1922), *Über die Krümmung des Raumes*, Zeitschrift fur Physik, 10, 377

Fritz, J., Poggianti, B. M., Cava, A., et al. (2014), *WINGS-SPE. III. Equivalent width measurements, spectral properties, and evolution of local cluster galaxies*, A&A, 566, A32

Fumagalli, A., Costanzi, M., Saro, A., et al. (2024), *Cosmological constraints from the abundance, weak lensing, and clustering of galaxy clusters: Application to the SDSS*, A&A, 682, A148

Gatti, M., Sheldon, E., Amon, A., et al. (2021), *Dark energy survey year 3 results: weak lensing shape catalogue*, MNRAS, 504(3), 4312

Ghirardini, V., Bulbul, E., Artis, E., et al. (2024), *The SRG/eROSITA all-sky survey: Cosmology constraints from cluster abundances in the western Galactic hemisphere*, arXiv e-prints, arXiv:2402.08458

Giacconi, R., Rosati, P., Tozzi, P., et al. (2001), *First Results from the X-Ray and Optical Survey of the Chandra Deep Field South**, ApJ, 551(2), 624

Gladders, M. D. and Yee, H. K. C. (2000), *A New Method For Galaxy Cluster Detection. I. The Algorithm*, AJ, 120, 2148

Gladders, M. D. and Yee, H. K. C. (2005), *The Red-Sequence Cluster Survey. I. The Survey and Cluster Catalogs for Patches RCS 0926+37 and RCS 1327+29*, ApJS, 157(1), 1

Gladders, M. D., Yee, H. K. C., Majumdar, S., et al. (2007), *Cosmological Constraints from the Red-Sequence Cluster Survey*, ApJ, 655(1), 128

Gonzalez, E. J., de los Rios, M., Oio, G. A., et al. (2018), *Analysis of candidates for interacting galaxy clusters. I. A1204 and A2029/A2033*, A&A, 611, A78

Goto, T., Okamura, S., Sekiguchi, M., et al. (2003), *The Environment of Passive Spiral Galaxies in the SDSS*, PASJ, 55(4), 757

Grandis, S., Ghirardini, V., Bocquet, S., et al. (2024), *The SRG/eROSITA All-Sky Survey. Dark Energy Survey year 3 weak gravitational lensing by eRASS1 selected galaxy clusters*, A&A, 687, A178

Grandis, S., Klein, M., Mohr, J. J., et al. (2020), *Validation of selection function, sample contamination and mass calibration in galaxy cluster samples*, MNRAS, 498(1), 771

Grandis, S., Mohr, J. J., Costanzi, M., et al. (2021), *Exploring the contamination of the DES-Y1 cluster sample with SPT-SZ selected clusters*, MNRAS, 504(1), 1253

Grandis, S., Mohr, J. J., Dietrich, J. P., et al. (2019), *Impact of weak lensing mass calibration on eROSITA galaxy cluster cosmological studies - a forecast*, MNRAS, 488(2), 2041

Gruen, D., Seitz, S., Brimioulle, F., et al. (2014), *Weak lensing analysis of SZ-selected clusters of galaxies from the SPT and Planck surveys*, MNRAS, 442(2), 1507

Güdel, M. (2004), *X-ray astronomy of stellar coronae*, A&ARv, 12(2-3), 71

Gunn, J. E. and Gott, III, J. R. (1972), *On the Infall of Matter Into Clusters of Galaxies and Some Effects on Their Evolution*, ApJ, 176, 1

Guth, A. H. (1981), *Inflationary universe: A possible solution to the horizon and flatness problems*, Phys. Rev. D, 23(2), 347

Guth, A. H. and Tye, S. H. H. (1980), *Phase Transitions and Magnetic Monopole Production in the Very Early Universe*, Phys. Rev. Lett., 44(10), 631

Ha, J.-H., Ryu, D., and Kang, H. (2018), *Properties of Merger Shocks in Merging Galaxy Clusters*, ApJ, 857(1), 26

Haiman, Z., Mohr, J. J., and Holder, G. P. (2001), *Constraints on Cosmological Parameters from Future Galaxy Cluster Surveys*, ApJ, 553(2), 545

Haines, C. P., Pereira, M. J., Smith, G. P., et al. (2015), *LoCuSS: The Slow Quenching of Star Formation in Cluster Galaxies and the Need for Pre-processing*, ApJ, 806, 101

Harnois-Déraps, J., Heydenreich, S., Giblin, B., et al. (2024), *KiDS-1000 and DES-Y1 combined: cosmology from peak count statistics*, MNRAS, 534(4), 3305

Harvey, D., Massey, R., Kitching, T., et al. (2015), *The nongravitational interactions of dark matter in colliding galaxy clusters*, Science, 347(6229), 1462

Hasinger, G. (1996), *The extragalactic X-ray and gamma-ray background.*, A&AS, 120, 607

Hasinger, G., Freyberg, M., Hu, E. M., et al. (2021), *The ROSAT Raster survey in the north ecliptic pole field - X-ray catalogue and optical identifications*, A&A, 645, A95

Henke, B. L., Lee, P., Tanaka, T. J., et al. (1982), *Low-Energy X-Ray Interaction Coefficients: Photoabsorption, Scattering and Reflection, E = 100-2000 eV, Z = 1-94*, Atomic Data and Nuclear Data Tables, 27, 1

Hennig, C., Mohr, J. J., Zenteno, A., et al. (2017), *Galaxy populations in massive galaxy clusters to z = 1.1: colour distribution, concentration, halo occupation number and red sequence fraction*, MNRAS, 467(4), 4015

Henry, J. P., Mullis, C. R., Voges, W., et al. (2006), *The ROSAT North Ecliptic Pole Survey: The X-Ray Catalog*, The Astrophysical Journal Supplement Series, 162(2), 304

Hernández-Lang, D., Mohr, J. J., Klein, M., et al. (2022), *The MADPSZ catalogue of Planck clusters over the DES region: extending to lower mass and higher redshift*, arXiv e-prints, arXiv:2210.04666

Herranz, D., Sanz, J. L., Hobson, M. P., et al. (2002), *Filtering techniques for the detection of Sunyaev-Zel'dovich clusters <br/> in multifrequency maps*, MNRAS, 336(4), 1057

Herschel, W. (1785), *On the Construction of the Heavens.*, Philosophical Transactions of the Royal Society of London Series I, 75, 213

Heymans, C., Tröster, T., Asgari, M., et al. (2021), *KiDS-1000 Cosmology: Multi-probe weak*

gravitational lensing and spectroscopic galaxy clustering constraints, A&A, 646, A140

High, F. W., Hoekstra, H., Leethochawalit, N., et al. (2012), *Weak-lensing Mass Measurements of Five Galaxy Clusters in the South Pole Telescope Survey Using Magellan/Megacam*, ApJ, 758(1), 68

High, F. W., Stalder, B., Song, J., et al. (2010), *Optical Redshift and Richness Estimates for Galaxy Clusters Selected with the Sunyaev-Zel'dovich Effect from 2008 South Pole Telescope Observations*, ApJ, 723(2), 1736

High, F. W., Stubbs, C. W., Rest, A., et al. (2009), *Stellar Locus Regression: Accurate Color Calibration and the Real-Time Determination of Galaxy Cluster Photometric Redshifts*, AJ, 138(1), 110

Hilton, M., Sifón, C., Naess, S., et al. (2021), *The Atacama Cosmology Telescope: A Catalog of >4000 Sunyaev–Zel'dovich Galaxy Clusters*, ApJS, 253(1), 3

Hinton, S. R., Davis, T. M., Lidman, C., et al. (2016), *MARZ: Manual and automatic redshifting software*, Astronomy and Computing, 15, 61

Hoekstra, H., Herbonnet, R., Muzzin, A., et al. (2015), *The Canadian Cluster Comparison Project: detailed study of systematics and updated weak lensing masses*, MNRAS, 449(1), 685

Høg, E., Fabricius, C., Makarov, V. V., et al. (2000), *The Tycho-2 catalogue of the 2.5 million brightest stars*, A&A, 355, L27

Holden, B. P., Illingworth, G. D., Franx, M., et al. (2007), *Mass Selection and the Evolution of the Morphology-Density Relation from $z = 0.8$ to 0*, ApJ, 670, 190

Hornschemeier, A. E., Brandt, W. N., Garmire, G. P., et al. (2001), *The Chandra Deep Survey of the Hubble Deep Field-North Area. II. Results from the Caltech Faint Field Galaxy Redshift Survey Area**, ApJ, 554(2), 742

Hou, A., Parker, L. C., Harris, W. E., and Wilman, D. J. (2009), *Statistical Tools for Classifying Galaxy Group Dynamics*, ApJ, 702, 1199

Hou, A., Parker, L. C., Wilman, D. J., et al. (2012), *Substructure in the most massive GEEC groups: field-like populations in dynamically active groups*, MNRAS, 421(4), 3594

Hu, W. (2003), *Self-consistency and calibration of cluster number count surveys for dark energy*, Phys. Rev. D, 67(8), 081304

Hubble, E. (1929), *A Relation between Distance and Radial Velocity among Extra-Galactic Nebulae*, Proceedings of the National Academy of Science, 15(3), 168

Huchra, J. P., Macri, L. M., Masters, K. L., et al. (2012), *The 2MASS Redshift Survey—Description and Data Release*, ApJS, 199(2), 26

Huterer, D. and White, M. (2002), *Weak Lensing as a Calibrator of the Cluster Mass-Temperature Relation*, ApJ, 578(2), L95

Ider Chitham, J., Comparat, J., Finoguenov, A., et al. (2020), *Cosmological constraints from CODEX galaxy clusters spectroscopically confirmed by SDSS-IV/SPIDERS DR16*, MNRAS, 499(4), 4768

Jansen, F., Lumb, D., Altieri, B., et al. (2001), *XMM-Newton observatory* - I. The spacecraft and operations*, A&A, 365(1), L1

Jenkins, A., Frenk, C. S., White, S. D. M., et al. (2001), *The mass function of dark matter haloes*, MNRAS, 321(2), 372

Jones, D. H., Read, M. A., Saunders, W., et al. (2009), *The 6dF Galaxy Survey: final redshift release (DR3) and southern large-scale structures*, MNRAS, 399(2), 683

Kalita, B. S. and Ebeling, H. (2019), *Jellyfish: Resolving the Kinematics of Extreme Ram-pressure Stripping at z ∼ 0.3*, ApJ, 887(2), 158

Kapferer, W., Sluka, C., Schindler, S., et al. (2009), *The effect of ram pressure on the star formation, mass distribution and morphology of galaxies*, A&A, 499(1), 87

Kelkar, K., Dwarakanath, K. S., Poggianti, B. M., et al. (2020), *Passive spirals and shock influenced star formation in the merging cluster A3376*, MNRAS, 496(1), 442

Kim, S. Y., Peter, A. H. G., and Wittman, D. (2017), *In the wake of dark giants: new signatures of dark matter self-interactions in equal-mass mergers of galaxy clusters*, MNRAS, 469(2), 1414

Klein, M., Grandis, S., Mohr, J. J., et al. (2019), *A new RASS galaxy cluster catalogue with low contamination extending to z ∼ 1 in the DES overlap region*, MNRAS, 488(1), 739

Klein, M., Hernández-Lang, D., Mohr, J. J., et al. (2023), *RASS-MCMF: a full-sky X-ray selected galaxy cluster catalogue*, MNRAS, 526(3), 3757

Klein, M., Mohr, J. J., Bocquet, S., et al. (2024a), *SPT-SZ MCMF: an extension of the SPT-SZ catalogue over the DES region*, MNRAS, 531(4), 3973

Klein, M., Mohr, J. J., and Davies, C. T. (2024b), *The ACT-DR5 MCMF Galaxy Cluster Catalog*, arXiv e-prints, arXiv:2406.14754

Klein, M., Mohr, J. J., Desai, S., et al. (2018), *A multicomponent matched filter cluster confirmation tool for eROSITA: initial application to the RASS and DES-SV data sets*, MNRAS, 474(3), 3324

Klein, M., Oguri, M., Mohr, J. J., et al. (2022), *The eROSITA Final Equatorial-Depth Survey (eFEDS). Optical confirmation, redshifts, and properties of the cluster and group catalog*, A&A, 661, A4

Kluge, M., Comparat, J., Liu, A., et al. (2024), *The SRG/eROSITA All-Sky Survey: Optical identification and properties of galaxy clusters and groups in the western galactic hemisphere*, A&A, 688, A210

Kodama, T. and Arimoto, N. (1997), *Origin of the colour-magnitude relation of elliptical galaxies.*, A&A, 320, 41

Koester, B. P., McKay, T. A., Annis, J., et al. (2007a), *A MaxBCG Catalog of 13,823 Galaxy Clusters from the Sloan Digital Sky Survey*, ApJ, 660(1), 239

Koester, B. P., McKay, T. A., Annis, J., et al. (2007b), *MaxBCG: A Red-Sequence Galaxy Cluster Finder*, ApJ, 660(1), 221

Komatsu, E., Smith, K. M., Dunkley, J., et al. (2011), *Seven-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation*, ApJS, 192(2), 18

Koulouridis, E., Clerc, N., Sadibekova, T., et al. (2021), *The X-CLASS survey: A catalogue of 1646 X-ray-selected galaxy clusters up to z ∼ 1.5*, A&A, 652, A12

Koyama, Y., Kodama, T., Nakata, F., et al. (2011), *Red Star-forming Galaxies and Their Environment at z = 0.4 Revealed by Panoramic Hα Imaging*, ApJ, 734(1), 66

Kravtsov, A. V. and Borgani, S. (2012), *Formation of Galaxy Clusters*, ARA&A, 50, 353

Leauthaud, A., Finoguenov, A., Kneib, J.-P., et al. (2010), *A Weak Lensing Study of X-ray Groups in the Cosmos Survey: Form and Evolution of the Mass-Luminosity Relation*, ApJ, 709(1), 97

Lemaître, G. (1927), *Un Univers homogène de masse constante et de rayon croissant rendant compte de la vitesse radiale des nébuleuses extra-galactiques*, Annales de la Société Scientifique de Bruxelles, 47, 49

Lesci, G. F., Marulli, F., Moscardini, L., et al. (2022), *AMICO galaxy clusters in KiDS-DR3: Cosmological constraints from counts and stacked weak lensing*, A&A, 659, A88

Levine, E. S., Schulz, A. E., and White, M. (2002), *Future Galaxy Cluster Surveys: The Effect of Theory Uncertainty on Constraining Cosmological Parameters*, ApJ, 577(2), 569

Lima, M. and Hu, W. (2005), *Self-calibration of cluster dark energy studies: Observable-mass distribution*, Phys. Rev. D, 72(4), 043006

Lin, Y., Mohr, J. J., and Stanford, S. A. (2004), *K-Band Properties of Galaxy Clusters and Groups: Luminosity Function, Radial Distribution, and Halo Occupation Number*, ApJ, 610, 745

Liu, A., Bulbul, E., Ghirardini, V., et al. (2022), *The eROSITA Final Equatorial-Depth Survey (eFEDS). Catalog of galaxy clusters and groups*, A&A, 661, A2

Liu, J., Hennig, C., Desai, S., et al. (2015), *Optical confirmation and redshift estimation of the Planck cluster candidates overlapping the Pan-STARRS Survey*, MNRAS, 449(4), 3370

López-Cruz, O., Barkhouse, W. A., and Yee, H. K. C. (2004), *The Color-Magnitude Effect in Early-Type Cluster Galaxies*, ApJ, 614, 679

Ma, C.-J., Ebeling, H., Marshall, P., and Schrabback, T. (2010), *The impact of a major cluster merger on galaxy evolution in MACSJ0025.4-1225*, MNRAS, 406, 121

Machado, R. E. G. and Lima Neto, G. B. (2013), *Simulations of the merging galaxy cluster Abell 3376*, MNRAS, 430(4), 3249

Machado, R. E. G., Monteiro-Oliveira, R., Lima Neto, G. B., and Cypriano, E. S. (2015), *Simulating the shocks in the dissociative galaxy cluster Abell 1758N*, MNRAS, 451, 3309

Mahler, G., Sharon, K., Gladders, M. D., et al. (2020), *Strong Lensing Model of SPT-CL J0356-5337, a Major Merger Candidate at Redshift 1.0359*, ApJ, 894(2), 150

Majumdar, S. and Mohr, J. J. (2003), *Importance of Cluster Structural Evolution in Using X-Ray and Sunyaev-Zeldovich Effect Galaxy Cluster Surveys to Study Dark Energy*, ApJ, 585(2), 603

Majumdar, S. and Mohr, J. J. (2004), *Self-Calibration in Cluster Studies of Dark Energy: Combining the Cluster Redshift Distribution, the Power Spectrum, and Mass Measurements*, ApJ, 613(1), 41

Mantz, A., Allen, S. W., Rapetti, D., and Ebeling, H. (2010), *The observed growth of massive galaxy clusters - I. Statistical methods and cosmological constraints*, MNRAS, 406(3), 1759

Mantz, A. B., von der Linden, A., Allen, S. W., et al. (2015), *Weighing the giants - IV. Cosmology and neutrino mass*, MNRAS, 446(3), 2205

Marinacci, F., Vogelsberger, M., Pakmor, R., et al. (2018), *First results from the IllustrisTNG simulations: radio haloes and magnetic fields*, MNRAS, 480(4), 5113

Markevitch, M., Gonzalez, A. H., Clowe, D., et al. (2004), *Direct Constraints on the Dark Matter Self-Interaction Cross Section from the Merging Galaxy Cluster 1E 0657-56*, ApJ, 606, 819

Markevitch, M. and Vikhlinin, A. (2007), *Shocks and cold fronts in galaxy clusters*, Phys. Rep., 443(1), 1

Marriage, T. A., Acquaviva, V., Ade, P. A. R., et al. (2011), *The Atacama Cosmology Telescope: Sunyaev-Zel'dovich-Selected Galaxy Clusters at 148 GHz in the 2008 Survey*, ApJ, 737(2), 61

Massey, R., Williams, L., Smit, R., et al. (2015), *The behaviour of dark matter associated with four bright cluster galaxies in the 10 kpc core of Abell 3827*, MNRAS, 449(4), 3393

Mastropietro, C. and Burkert, A. (2008), *Simulating the Bullet Cluster*, MNRAS, 389(2), 967

Maturi, M., Bellagamba, F., Radovich, M., et al. (2019), *AMICO galaxy clusters in KiDS-DR3: sample properties and selection function*, MNRAS, 485(1), 498

Mazoun, A., Bocquet, S., Garny, M., et al. (2024), *Probing interacting dark sector models with future weak lensing-informed galaxy cluster abundance constraints from SPT-3G and CMB-S4*, Phys. Rev. D, 109(6), 063536

McClintock, T., Varga, T. N., Gruen, D., et al. (2019), *Dark Energy Survey Year 1 results: weak lensing mass calibration of redMaPPer galaxy clusters*, MNRAS, 482(1), 1352

McDonald, M., Allen, S. W., Bayliss, M., et al. (2017), *The Remarkable Similarity of Massive Galaxy Clusters from z 0 to z 1.9*, ApJ, 843(1), 28

McDonald, M., Benson, B. A., Vikhlinin, A., et al. (2013), *The Growth of Cool Cores and Evolution of Cooling Properties in a Sample of 83 Galaxy Clusters at $0.3 < z < 1.2$ Selected from the SPT-SZ Survey*, ApJ, 774, 23

McDonald, M., McNamara, B. R., van Weeren, R. J., et al. (2015), *Deep Chandra, HST-COS, and Megacam Observations of the Phoenix Cluster: Extreme Star Formation and AGN Feedback on Hundred Kiloparsec Scales*, ApJ, 811(2), 111

McPartland, C., Ebeling, H., Roediger, E., and Blumenthal, K. (2016), *Jellyfish: the origin and distribution of extreme ram-pressure stripping events in massive galaxy clusters*, MNRAS, 455, 2994

Mei, S., Stanford, S. A., Holden, B. P., et al. (2012), *Early-type Galaxies at $z = 1.3$. I. The Lynx Supercluster: Cluster and Groups at $z = 1.3$. Morphology and Color-Magnitude Relation*, ApJ, 754, 141

Melchior, P., Gruen, D., McClintock, T., et al. (2017), *Weak-lensing mass calibration of redMaPPer galaxy clusters in Dark Energy Survey Science Verification data*, MNRAS, 469(4), 4899

Melin, J. B., Bartlett, J. G., and Delabrouille, J. (2006), *Catalog extraction in SZ cluster surveys: a matched filter approach*, A&A, 459(2), 341

Melin, J. B., Bartlett, J. G., Tarrío, P., and Pratt, G. W. (2021), *PSZSPT: A joint Planck and SPT-SZ cluster catalog*, A&A, 647, A106

Mellier, Y. (1999), *Probing the Universe with Weak Lensing*, ARA&A, 37, 127

Menanteau, F., Hughes, J. P., Sifón, C., et al. (2012), *The Atacama Cosmology Telescope:*

ACT-CL J0102-4915 "El Gordo," a Massive Merging Cluster at Redshift 0.87, ApJ, 748(1), 7

Menci, N. and Fusco-Femiano, R. (1996), *Galaxy Velocity Dispersion Profiles from Merging in Clusters*, ApJ, 472, 46

Merloni, A., Lamer, G., Liu, T., et al. (2024), *The SRG/eROSITA all-sky survey. First X-ray catalogues and data release of the western Galactic hemisphere*, A&A, 682, A34

Merloni, A., Predehl, P., Becker, W., et al. (2012), *eROSITA Science Book: Mapping the Structure of the Energetic Universe*, arXiv e-prints, arXiv:1209.3114

Miralda-Escude, J. (1991), *Gravitational Lensing by Clusters of Galaxies: Constraining the Mass Distribution*, ApJ, 370, 1

Misner, C. W., Thorne, K. S., and Wheeler, J. A. (1973), *Gravitation*

Mohr, J. J. (2005), *Cluster Survey Studies of the Dark Energy*, in *Observing Dark Energy*, edited by S. C. Wolff, T. R. Lauer, volume 339 of *Astronomical Society of the Pacific Conference Series*, 140

Monteiro-Oliveira, R., Cypriano, E. S., Vitorelli, A. Z., et al. (2018), *New insights on the dissociative merging galaxy cluster Abell 2034*, MNRAS, 481, 1097

Monteiro-Oliveira, R., Doubrawa, L., Machado, R. E. G., et al. (2020), *Revising the merger scenario of the galaxy cluster Abell 1644: a new gas poor structure discovered by weak gravitational lensing*, MNRAS, 495(2), 2007

Monteiro-Oliveira, R., Lima Neto, G. B., Cypriano, E. S., et al. (2017), *Weak lensing and spectroscopic analysis of the nearby dissociative merging galaxy cluster Abell 3376*, MNRAS, 468(4), 4566

Monteiro-Oliveira, R., Soja, A. C., Ribeiro, A. L. B., et al. (2021), *Probing Saraswati's heart: evaluating the dynamical state of the massive galaxy cluster A2631 through a comprehensive weak-lensing and dynamical analysis*, MNRAS, 501(1), 756

Moran, S. M., Ellis, R. S., Treu, T., et al. (2007), *A Wide-Field Survey of Two $z \sim 0.5$ Galaxy Clusters: Identifying the Physical Processes Responsible for the Observed Transformation of Spirals into S0s*, ApJ, 671(2), 1503

Morganson, E., Gruendl, R. A., Menanteau, F., et al. (2018), *The Dark Energy Survey Image Processing Pipeline*, PASP, 130(989), 074501

Morrison, R. and McCammon, D. (1983), *Interstellar photoelectric absorption cross sections, 0.03-10 keV.*, ApJ, 270, 119

Motl, P. M., Hallman, E. J., Burns, J. O., and Norman, M. L. (2005), *The Integrated Sunyaev-Zeldovich Effect as a Superior Method for Measuring the Mass of Clusters of Galaxies*, ApJ, 623(2), L63

Moura, M. T., Machado, R. E. G., and Monteiro-Oliveira, R. (2021), *Simulations of the merging galaxy cluster Abell 2034: what determines the level of separation between gas and dark matter*, MNRAS, 500(2), 1858

Munari, E., Biviano, A., Borgani, S., et al. (2013), *The relation between velocity dispersion and mass in simulated clusters of galaxies: dependence on the tracer and the baryonic physics*, MNRAS, 430(4), 2638

Murata, R., Oguri, M., Nishimichi, T., et al. (2019), *The mass-richness relation of optically selected clusters from weak gravitational lensing and abundance with Subaru HSC first-year data*, PASJ, 71(5), 107

Myles, J., Alarcon, A., Amon, A., et al. (2021), *Dark Energy Survey Year 3 results: redshift calibration of the weak lensing source galaxies*, MNRAS, 505(3), 4249

Nagai, D., Kravtsov, A. V., and Vikhlinin, A. (2007), *Effects of Galaxy Formation on Thermodynamics of the Intracluster Medium*, ApJ, 668(1), 1

Naiman, J. P., Pillepich, A., Springel, V., et al. (2018), *First results from the IllustrisTNG simulations: a tale of two elements - chemical evolution of magnesium and europium*, MNRAS, 477(1), 1206

Navarro, J. F., Frenk, C. S., and White, S. D. M. (1996), *The Structure of Cold Dark Matter Halos*, ApJ, 462, 563

Navarro, J. F., Frenk, C. S., and White, S. D. M. (1997), *A Universal Density Profile from Hierarchical Clustering*, ApJ, 490(2), 493

Nelson, D., Pillepich, A., Springel, V., et al. (2018), *First results from the IllustrisTNG simulations: the galaxy colour bimodality*, MNRAS, 475(1), 624

Nelson, K., Lau, E. T., Nagai, D., et al. (2014), *Weighing Galaxy Clusters with Gas. II. On the Origin of Hydrostatic Mass Bias in ΛCDM Galaxy Clusters*, ApJ, 782, 107

Nelson, K., Rudd, D. H., Shaw, L., and Nagai, D. (2012), *Evolution of the Merger-induced Hydrostatic Mass Bias in Galaxy Clusters*, ApJ, 751, 121

Ng, K. Y., Dawson, W. A., Wittman, D., et al. (2015), *The return of the merging galaxy subclusters of El Gordo?*, MNRAS, 453(2), 1531

Nurgaliev, D., McDonald, M., Benson, B. A., et al. (2017), *Testing for X-Ray-SZ Differences and Redshift Evolution in the X-Ray Morphology of Galaxy Clusters*, ApJ, 841(1), 5

Ochsenbein, F., Bauer, P., and Marcout, J. (2000), *The VizieR database of astronomical catalogues*, A&AS, 143, 23

Oemler, Augustus, J. (1974), *The Systematic Properties of Clusters of Galaxies. Photometry of 15 Clusters*, ApJ, 194, 1

Olave-Rojas, D., Cerulo, P., Demarco, R., et al. (2018), *Galaxy pre-processing in substructures around z 0.4 galaxy clusters*, MNRAS, 479(2), 2328

Owers, M. S., Couch, W. J., Nulsen, P. E. J., and Randall, S. W. (2012), *Shocking Tails in the Major Merger Abell 2744*, ApJ, 750(1), L23

Paccagnella, A., Vulcani, B., Poggianti, B. M., et al. (2016), *Slow Quenching of Star Formation in OMEGAWINGS Clusters: Galaxies in Transition in the Local Universe*, ApJ, 816, L25

Paccagnella, A., Vulcani, B., Poggianti, B. M., et al. (2019), *The strong correlation between post-starburst fraction and environment*, MNRAS, 482(1), 881

Pallero, D., Gómez, F. A., Padilla, N. D., et al. (2022), *Too dense to go through: the role of low-mass clusters in the pre-processing of satellite galaxies*, MNRAS, 511(3), 3210

Paulino-Afonso, A., Sobral, D., Darvish, B., et al. (2020), *VIS^{3}COS. III. Environmental effects on the star formation histories of galaxies at $z \sim 0.8$ seen in [O II], Hδ, and D_n4000*, A&A,

633, A70

Paulus, M. (2021), *Toward accurate galaxy cluster masses with gravitational weak lensing*, Ph.D. thesis, Ludwig-Maximilians University of Munich, Germany

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

Penzias, A. A. and Wilson, R. W. (1965), *A Measurement of Excess Antenna Temperature at 4080 Mc/s.*, ApJ, 142, 419

Perlmutter, S., Aldering, G., Goldhaber, G., et al. (1999), *Measurements of* Ω *and* Λ *from 42 High-Redshift Supernovae*, ApJ, 517(2), 565

Philcox, O. H. E. and Ivanov, M. M. (2022), *BOSS DR12 full-shape cosmology:* Λ *CDM constraints from the large-scale galaxy power spectrum and bispectrum monopole*, Phys. Rev. D, 105(4), 043517

Piffaretti, R., Arnaud, M., Pratt, G. W., et al. (2011), *The MCXC: a meta-catalogue of x-ray detected clusters of galaxies*, A&A, 534, A109

Pillepich, A., Nelson, D., Hernquist, L., et al. (2018), *First results from the IllustrisTNG simulations: the stellar mass content of groups and clusters of galaxies*, MNRAS, 475(1), 648

Pinkney, J., Roettiger, K., Burns, J. O., and Bird, C. M. (1996), *Evaluation of Statistical Tests for Substructure in Clusters of Galaxies*, ApJS, 104, 1

Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. (2014a), *Planck 2013 results. I. Overview of products and scientific results*, A&A, 571, A1

Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. (2014b), *Planck 2013 results. XXIX. The Planck catalogue of Sunyaev-Zeldovich sources*, A&A, 571, A29

Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. (2011), *Planck early results. I. The Planck mission*, A&A, 536, A1

Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. (2016a), *Planck 2015 results. XXVII. The second Planck catalogue of Sunyaev-Zeldovich sources*, A&A, 594, A27

Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. (2016b), *Planck 2015 results. XXIV. Cosmology from Sunyaev-Zeldovich cluster counts*, A&A, 594, A24

Planck Collaboration, Aghanim, N., Akrami, Y., et al. (2020a), *Planck 2018 results. I. Overview and the cosmological legacy of Planck*, A&A, 641, A1

Planck Collaboration, Aghanim, N., Akrami, Y., et al. (2020b), *Planck 2018 results. VI. Cosmological parameters*, A&A, 641, A6

Poggianti, B. M., Bridges, T. J., Komiyama, Y., et al. (2004), *A Comparison of the Galaxy Populations in the Coma and Distant Clusters: The Evolution of k+a Galaxies and the Role of the Intracluster Medium*, ApJ, 601(1), 197

Poggianti, B. M., Fasano, G., Omizzolo, A., et al. (2016), *Jellyfish Galaxy Candidates at Low Redshift*, AJ, 151, 78

Poole, G. B., Fardal, M. A., Babul, A., et al. (2006), *The impact of mergers on relaxed X-ray clusters - I. Dynamical evolution and emergent transient structures*, MNRAS, 373(3), 881

Postman, M., Franx, M., Cross, N. J. G., et al. (2005), *The Morphology-Density Relation in z ˜ 1 Clusters*, ApJ, 623, 721

Pranger, F., Böhm, A., Ferrari, C., et al. (2013), *The galaxy population of the complex cluster system Abell 3921*, A&A, 557, A62

Pranger, F., Böhm, A., Ferrari, C., et al. (2014), *Abell 2384: the galaxy population of a cluster post-merger*, A&A, 570, A40

Predehl, P., Andritschke, R., Arefiev, V., et al. (2021), *The eROSITA X-ray telescope on SRG*, A&A, 647, A1

Press, W. H. and Schechter, P. (1974), *Formation of Galaxies and Clusters of Galaxies by Self-Similar Gravitational Condensation*, ApJ, 187, 425

Prusti, T., De Bruijne, J., Brown, A. G., et al. (2016), *The gaia mission*, Astronomy & astrophysics, 595, A1

Quintana, H., Carrasco, E. R., and Reisenegger, A. (2000), *The Shapley Supercluster. II. Spectroscopic Observations in a Wide Area and General Morphology*, AJ, 120(2), 511

Raouf, M., Smith, R., Khosroshahi, H. G., et al. (2019), *The Impact of the Dynamical State of Galaxy Groups on the Stellar Populations of Central Galaxies*, ApJ, 887(2), 264

Reiprich, T. H. and Böhringer, H. (2002), *The Mass Function of an X-Ray Flux-limited Sample of Galaxy Clusters*, ApJ, 567(2), 716

Ribeiro, A. L. B., Lopes, P. A. A., and Rembold, S. B. (2013), *NoSOCS in SDSS. III. The interplay between galaxy evolution and the dynamical state of galaxy clusters*, A&A, 556, A74

Riess, A. G., Casertano, S., Yuan, W., et al. (2019), *Large Magellanic Cloud Cepheid Standards Provide a 1% Foundation for the Determination of the Hubble Constant and Stronger Evidence for Physics beyond ΛCDM*, ApJ, 876(1), 85

Riess, A. G., Filippenko, A. V., Challis, P., et al. (1998), *Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant*, AJ, 116(3), 1009

Rindler, W. (1956), *Visual horizons in world models*, MNRAS, 116, 662

Rines, K., Geller, M. J., Diaferio, A., and Kurtz, M. J. (2013), *Measuring the Ultimate Halo Mass of Galaxy Clusters: Redshifts and Mass Profiles from the Hectospec Cluster Survey (HeCS)*, ApJ, 767(1), 15

Rizza, E., Burns, J. O., Ledlow, M. J., et al. (1998), *X-ray observations of distant Abell clusters*, MNRAS, 301(2), 328

Rosen, S. R., Webb, N. A., Watson, M. G., et al. (2016), *The XMM-Newton serendipitous survey - VII. The third XMM-Newton serendipitous source catalogue*, A&A, 590, A1

Rozo, E. and Rykoff, E. S. (2014), *redMaPPer II: X-Ray and SZ Performance Benchmarks for the SDSS Catalog*, ApJ, 783(2), 80

Rozo, E., Rykoff, E. S., Evrard, A., et al. (2009a), *Constraining the Scatter in the Mass-richness Relation of maxBCG Clusters with Weak Lensing and X-ray Data*, ApJ, 699(1), 768

Rozo, E., Rykoff, E. S., Koester, B. P., et al. (2009b), *Improvement of the Richness Estimates of maxBCG Clusters*, ApJ, 703(1), 601

Ruel, J., Bazin, G., Bayliss, M., et al. (2014), *Optical Spectroscopy and Velocity Dispersions of Galaxy Clusters from the SPT-SZ Survey*, ApJ, 792(1), 45

Rykoff, E. S., McKay, T. A., Becker, M. R., et al. (2008), *Measuring the Mean and Scatter of the*

X-Ray Luminosity-Optical Richness Relation for maxBCG Galaxy Clusters, ApJ, 675(2), 1106

Rykoff, E. S., Rozo, E., Busha, M. T., et al. (2014), *redMaPPer. I. Algorithm and SDSS DR8 Catalog*, ApJ, 785(2), 104

Rykoff, E. S., Rozo, E., Hollowood, D., et al. (2016), *The RedMaPPer Galaxy Cluster Catalog From DES Science Verification Data*, ApJS, 224(1), 1

Salvati, L., Saro, A., Bocquet, S., et al. (2022), *Combining Planck and SPT Cluster Catalogs: Cosmological Analysis and Impact on the Planck Scaling Relation Calibration*, The Astrophysical Journal, 934(2), 129

Salvato, M., Buchner, J., Budavári, T., et al. (2018), *Finding counterparts for all-sky X-ray surveys with NWAY: a Bayesian algorithm for cross-matching multiple catalogues*, MNRAS, 473, 4937

Salvato, M., Wolf, J., Dwelly, T., et al. (2021), *The eROSITA Final Equatorial-Depth Survey (eFEDS): Identification and characterization of the counterparts to the point-like sources*, arXiv e-prints, arXiv:2106.14520

Sarazin, C. L. (1986), *X-ray emission from clusters of galaxies*, Reviews of Modern Physics, 58(1), 1

Sarazin, C. L. (1988), *X-ray emission from clusters of galaxies*

Sarazin, C. L. (2002), *The Physics of Cluster Mergers*, in *Merging Processes in Galaxy Clusters*, edited by L. Feretti, I. M. Gioia, G. Giovannini, volume 272 of *Astrophysics and Space Science Library*, 1–38

Sarazin, C. L. (2004), *Mergers, Cosmic Rays, and Nonthermal Processes in Clusters of Galaxies*, Journal of Korean Astronomical Society, 37(5), 433

Saro, A., Mohr, J. J., Bazin, G., and Dolag, K. (2013), *Toward Unbiased Galaxy Cluster Masses from Line-of-sight Velocity Dispersions*, ApJ, 772(1), 47

Schindler, S. and Muller, E. (1993), *Simulations of the evolution of galaxy clusters. II. Dynamics of the intra-cluster gas.*, A&A, 272, 137

Schneider, P. (2006), *Extragalactic Astronomy and Cosmology: An Introduction*, Springer

Schneider, P., Ehlers, J., and Falco, E. E. (1992), *Gravitational Lenses*

Sevilla-Noarbe, I., Bechtol, K., Carrasco Kind, M., et al. (2021), *Dark Energy Survey Year 3 Results: Photometric Data Set for Cosmology*, ApJS, 254(2), 24

Sheen, Y.-K., Yi, S. K., Ree, C. H., et al. (2016), *Recent Galaxy Mergers and Residual Star Formation of Red Sequence Galaxies in Galaxy Clusters*, ApJ, 827(1), 32

Sheldon, E. S. (2014), *An implementation of Bayesian lensing shear measurement.*, MNRAS, 444, L25

Singh, A., Mohr, J. J., Davies, C. T., et al. (2024), *Galaxy cluster matter profiles: I. Self-similarity and mass calibration*, arXiv e-prints, arXiv:2407.10961

Skrutskie, M. F., Cutri, R. M., Stiening, R., et al. (2006), *The Two Micron All Sky Survey (2MASS)*, AJ, 131(2), 1163

Smit, R., Swinbank, A. M., Massey, R., et al. (2017), *A gravitationally boosted MUSE survey for emission-line galaxies at $z \gtrsim 5$ behind the massive cluster RCS 0224*, MNRAS, 467, 3306

Smith, R. K., Brickhouse, N. S., Liedahl, D. A., and Raymond, J. C. (2001), *Collisional Plasma Models with APEC/APED: Emission-Line Diagnostics of Hydrogen-like and Helium-like Ions*, ApJ, 556(2), L91

Smoot, G. F., Bennett, C. L., Kogut, A., et al. (1992), *Structure in the COBE Differential Microwave Radiometer First-Year Maps*, ApJ, 396, L1

Sobral, D., Stroe, A., Koyama, Y., et al. (2016), *The nature of Hα star-forming galaxies at $z \sim 0.4$ in and around Cl 0939+4713: the environment matters*, MNRAS, 458(4), 3443

Somboonpanyakul, T., McDonald, M., Gaspari, M., et al. (2021), *The Clusters Hiding in Plain Sight (CHiPS) Survey: Complete Sample of Extreme BCG Clusters*, ApJ, 910(1), 60

Song, J., Mohr, J. J., Barkhouse, W. A., et al. (2012a), *A Parameterized Galaxy Catalog Simulator for Testing Cluster Finding, Mass Estimation, and Photometric Redshift Estimation in Optical and Near-infrared Surveys*, ApJ, 747(1), 58

Song, J., Zenteno, A., Stalder, B., et al. (2012b), *Redshifts, Sample Purity, and BCG Positions for the Galaxy Cluster Catalog from the First 720 Square Degrees of the South Pole Telescope Survey*, ApJ, 761(1), 22

Soucail, G., Mellier, Y., Fort, B., et al. (1987), *Further data on the blue ring-like structure in A 370.*, A&A, 184, L7

Spergel, D. N., Verde, L., Peiris, H. V., et al. (2003), *First-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Determination of Cosmological Parameters*, ApJS, 148(1), 175

Springel, V., Pakmor, R., Pillepich, A., et al. (2018), *First results from the IllustrisTNG simulations: matter and galaxy clustering*, MNRAS, 475(1), 676

Springel, V., White, S. D. M., Jenkins, A., et al. (2005), *Simulations of the formation, evolution and clustering of galaxies and quasars*, Nature, 435(7042), 629

Stanford, S. A., Eisenhardt, P. R., and Dickinson, M. (1998), *The Evolution of Early-Type Galaxies in Distant Clusters*, ApJ, 492(2), 461

Staniszewski, Z., Ade, P. A. R., Aird, K. A., et al. (2009), *Galaxy Clusters Discovered with a Sunyaev-Zel'dovich Effect Survey*, ApJ, 701(1), 32

Stoehr, F., Fraquelli, D., Kamp, I., et al. (2007), *DER_SNR: A Simple & General Spectroscopic Signal-to-Noise Measurement Algorithm*, Space Telescope European Coordinating Facility Newsletter, 42, 4

Strazzullo, V., Pannella, M., Mohr, J. J., et al. (2019), *Galaxy populations in the most distant SPT-SZ clusters. I. Environmental quenching in massive clusters at $1.4 \lesssim z \lesssim 1.7$*, A&A, 622, A117

Stroe, A. and Sobral, D. (2021), *ENISALA. II. Distinct Star Formation and Active Galactic Nucleus Activity in Merging and Relaxed Galaxy Clusters*, ApJ, 912(1), 55

Stroe, A., Sobral, D., Dawson, W., et al. (2015), *The rise and fall of star formation in z 0.2 merging galaxy clusters*, MNRAS, 450(1), 646

Stroe, A., Sobral, D., Paulino-Afonso, A., et al. (2017), *A large Hα survey of star formation in relaxed and merging galaxy cluster environments at z 0.15-0.3*, MNRAS, 465, 2916

Stroe, A., Sobral, D., Röttgering, H. J. A., and van Weeren, R. J. (2014), *The role of cluster*

mergers and travelling shocks in shaping the Hα luminosity function at z 0.2: 'sausage' and 'toothbrush' clusters, MNRAS, 438(2), 1377

Sunyaev, R. A. and Zeldovich, Y. B. (1972), *The Observations of Relic Radiation as a Test of the Nature of X-Ray Radiation from the Clusters of Galaxies*, Comments on Astrophysics and Space Physics, 4, 173

Swetz, D. S., Ade, P. A. R., Amiri, M., et al. (2011), *Overview of the Atacama Cosmology Telescope: Receiver, Instrumentation, and Telescope Systems*, ApJS, 194(2), 41

Takizawa, M., Nagino, R., and Matsushita, K. (2010), *Mass Estimation of Merging Galaxy Clusters*, PASJ, 62, 951

Testa, P., Saar, S. H., and Drake, J. J. (2015), *Stellar activity and coronal heating: an overview of recent results*, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 373(2042), 20140259

Thompson, R., Davé, R., and Nagamine, K. (2015), *The rise and fall of a challenger: the Bullet Cluster in Λ cold dark matter simulations*, MNRAS, 452, 3030

Tinker, J., Kravtsov, A. V., Klypin, A., et al. (2008), *Toward a Halo Mass Function for Precision Cosmology: The Limits of Universality*, ApJ, 688(2), 709

Tonry, J. and Davis, M. (1979), *A survey of galaxy redshifts. I. Data reduction techniques.*, AJ, 84, 1511

Toomre, A. and Toomre, J. (1972), *Galactic Bridges and Tails*, ApJ, 178, 623

Treu, T., Ellis, R. S., Kneib, J.-P., et al. (2003), *A Wide-Field Hubble Space Telescope Study of the Cluster Cl 0024+16 at z = 0.4. I. Morphological Distributions to 5 Mpc Radius*, ApJ, 591, 53

Truemper, J. (1982), *The ROSAT mission*, Advances in Space Research, 2, 241

Trümper, J. (1993), *ROSAT—A New Look at the X-ray Sky*, Science, 260(5115), 1769

Tyson, J. A., Valdes, F., and Wenk, R. A. (1990), *Detection of Systematic Gravitational Lens Galaxy Image Alignments: Mapping Dark Matter in Galaxy Clusters*, ApJ, 349, L1

van der Wel, A., Holden, B. P., Franx, M., et al. (2007), *The Evolution of the Field and Cluster Morphology-Density Relation for Mass-Selected Samples of Galaxies*, ApJ, 670, 206

van Weeren, R. J., Brüggen, M., Röttgering, H. J. A., and Hoeft, M. (2011), *Using double radio relics to constrain galaxy cluster mergers: a model of double radio relics in CIZA J2242.8+5301*, MNRAS, 418(1), 230

Viana, P. T. P. and Liddle, A. R. (1996), *The cluster abundance in flat and open cosmologies*, MNRAS, 281, 323

Vikhlinin, A., McNamara, B. R., Forman, W., et al. (1998), *A Catalog of 200 Galaxy Clusters Serendipitously Detected in the ROSAT PSPC Pointed Observations*, ApJ, 502(2), 558

Vilhu, O. (1984), *The nature of magnetic activity in lower main sequence stars.*, A&A, 133, 117

Visvanathan, N. and Sandage, A. (1977), *The color - absolute magnitude relation for E and S0 galaxies. I. Calibration and tests for universality using Virgo and eight other nearby clusters.*, ApJ, 216, 214

Voges, W., Aschenbach, B., Boller, T., et al. (1999), *The ROSAT all-sky survey bright source*

catalogue, A&A, 349, 389

Voges, W., Aschenbach, B., Boller, T., et al. (2000), *ROSAT All-Sky Survey Faint Source Catalog (Voges+ 2000)*, VizieR Online Data Catalog, 9029, 0

Vogt, N. P., Haynes, M. P., Giovanelli, R., and Herter, T. (2004), *M/L, Hα Rotation Curves, and H I Gas Measurements for 329 Nearby Cluster and Field Spirals. II. Evidence for Galaxy Infall*, AJ, 127(6), 3300

Vogt, S. M. L., Bocquet, S., Davies, C. T., et al. (2024), *Constraining f (R) gravity using future galaxy cluster abundance and weak-lensing mass calibration datasets*, Phys. Rev. D, 109(12), 123503

von der Linden, A., Mantz, A., Allen, S. W., et al. (2014), *Robust weak-lensing mass calibration of Planck galaxy clusters*, MNRAS, 443(3), 1973

Weilbacher, P. M., Streicher, O., and Palsa, R. (2016), *MUSE-DRP: MUSE Data Reduction Pipeline*, Astrophysics Source Code Library, record ascl:1610.004

Weilbacher, P. M., Streicher, O., Urrutia, T., et al. (2014), *The MUSE Data Reduction Pipeline: Status after Preliminary Acceptance Europe*, in *Astronomical Data Analysis Software and Systems XXIII*, edited by N. Manset, P. Forshay, volume 485 of *Astronomical Society of the Pacific Conference Series*, 451

Weinberg, S. (1972), *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity*

Weisskopf, M. C., Tananbaum, H. D., Van Speybroeck, L. P., and O'Dell, S. L. (2000), *Chandra X-ray Observatory (CXO): overview*, in *X-Ray Optics, Instruments, and Missions III*, edited by J. E. Truemper, B. Aschenbach, SPIE

Wen, Z. L. and Han, J. L. (2022), *Clusters of galaxies up to z = 1.5 identified from photometric data of the Dark Energy Survey and unWISE*, MNRAS, 513(3), 3946

Wen, Z. L. and Han, J. L. (2024), *A Catalog of 1.58 Million Clusters of Galaxies Identified from the DESI Legacy Imaging Surveys*, ApJS, 272(2), 39

White, J. A., Canning, R. E. A., King, L. J., et al. (2015), *Dynamical analysis of galaxy cluster merger Abell 2146*, MNRAS, 453(3), 2718

White, M. (2001), *The mass of a halo*, A&A, 367, 27

White, M. (2002), *The Mass Function*, ApJS, 143(2), 241

White, S. D. M. (1976), *Dynamical friction in spherical clusters.*, MNRAS, 174, 19

White, S. D. M., Efstathiou, G., and Frenk, C. S. (1993), *The amplitude of mass fluctuations in the universe*, MNRAS, 262(4), 1023

Williamson, R., Benson, B. A., High, F. W., et al. (2011), *A Sunyaev-Zel'dovich-selected Sample of the Most Massive Galaxy Clusters in the 2500 deg^2 South Pole Telescope Survey*, ApJ, 738, 139

Wittman, D. (2019), *Dynamical Properties of Merging Galaxy Clusters from Simulated Analogs*, ApJ, 881(2), 121

Wittman, D., Cornell, B. H., and Nguyen, J. (2018), *Simulated Analogs of Merging Galaxy Clusters Constrain the Viewing Angle*, ApJ, 862(2), 160

Wright, C. O. and Brainerd, T. G. (2000), *Gravitational Lensing by NFW Halos*, ApJ, 534(1), 34

Wright, N. J., Drake, J. J., Mamajek, E. E., and Henry, G. W. (2011), *The Stellar-activity-Rotation Relationship and the Evolution of Stellar Dynamos*, ApJ, 743(1), 48

Ye, J.-N., Guo, H., Zheng, Z., and Zehavi, I. (2017), *Properties and Origin of Galaxy Velocity Bias in the Illustris Simulation*, ApJ, 841(1), 45

Yee, H. K. C. and Ellingson, E. (2003), *Correlations of Richness and Global Properties in Galaxy Clusters*, ApJ, 585(1), 215

Yoon, Y. and Im, M. (2020), *Star Formation Enhancement in Barred Disk Galaxies in Interacting Galaxy Clusters*, ApJ, 893(2), 117

Yoon, Y., Im, M., Lee, G.-H., et al. (2019), *Observational evidence for bar formation in disk galaxies via cluster-cluster interaction*, Nature Astronomy, 3, 844

York, D. G., Adelman, J., Anderson, John E., J., et al. (2000), *The Sloan Digital Sky Survey: Technical Summary*, AJ, 120(3), 1579

Yoshikawa, K., Jing, Y. P., and Börner, G. (2003), *Spatial and Dynamical Biases in Velocity Statistics of Galaxies*, ApJ, 590(2), 654

Yuan, S., Garrison, L. H., Eisenstein, D. J., and Wechsler, R. H. (2022), *Stringent σ_8 constraints from small-scale galaxy clustering using a hybrid MCMC + emulator framework*, MNRAS, 515(1), 871

Zeldovich, Y. B. and Khlopov, M. Y. (1978), *On the concentration of relic magnetic monopoles in the universe*, Physics Letters B, 79(3), 239

Zenteno, A., Hernández-Lang, D., Klein, M., et al. (2020), *A joint SZ-X-ray-optical analysis of the dynamical state of 288 massive galaxy clusters*, MNRAS, 495(1), 705

Zenteno, A., Mohr, J. J., Desai, S., et al. (2016), *Galaxy populations in the 26 most massive galaxy clusters in the South Pole Telescope SPT-SZ survey*, MNRAS, 462, 830

Zenteno, A., Song, J., Desai, S., et al. (2011), *A Multiband Study of the Galaxy Populations of the First Four Sunyaev-Zel'dovich Effect Selected Galaxy Clusters*, ApJ, 734, 3

Zhang, Y.-Y., Verdugo, M., Klein, M., and Schneider, P. (2012), *Probing cluster dynamics in RXC J1504.1-0248 via radial and two-dimensional gas and galaxy properties*, A&A, 542, A106

ZuHone, J. A. (2011), *A Parameter Space Exploration of Galaxy Cluster Mergers. I. Gas Mixing and the Generation of Cluster Entropy*, ApJ, 728(1), 54

ZuHone, J. A., Kowalik, K., Öhman, E., et al. (2018), *The Galaxy Cluster Merger Catalog: An Online Repository of Mock Observations from Simulated Galaxy Cluster Mergers*, ApJS, 234(1), 4

Zuntz, J., Paterno, M., Jennings, E., et al. (2015), *CosmoSIS: Modular cosmological parameter estimation*, Astronomy and Computing, 12, 45

Zwicky, F., Herzog, E., Wild, P., et al. (1961), *Catalogue of galaxies and of clusters of galaxies, Vol. I*

ACKNOWLEDGMENTS/AGRADECIMIENTOS

First of all I want to thank my supervisors Joe Mohr and Matthias Klein for having me as a student. Thanks a lot for your patience, guidance and advice throughout my research. I will always be grateful for the opportunity you gave me as your student and for helping me figure out my life even after the PhD.

Ok, lets get this over with. There are so many people that I have/want to say thank you for these few years. First, USM people, without all of you these years wouldn't have been the same, even with the pandemic and stuff. For those of you still at USM, don't forget about me when I become container person, or I'll hunt you down. I'll thank you all in alphabetical order (of the real names) because I'm a coward, you can try to order it yourselves afterwards: Ex office mate #2, Ale(x)andra, Christian Bale, Tine (never once said), Elenna, Elisa Catgirl, Kristof, Maria, Peter, Sarvesh, Samantha, Stefan, Tirso, Thomasinho. Thank you for the good times, trips, gatherings and memories :). Ahora pa los K, misma wea, en orden alfabético por simpleza, pero cada uno de ustedes me ayudó y apoyó en su momento y estoy siempre agradecido de haberlos conocido y haber pasado tiempo con ustedes durante estos años: Ale, Catarina, Chichi, Pipeño, Jay, Joako, Luci, Nico, Paula, Simón, Tiara, Yru. From MPA I also want to thank Miha and Oliver for all the good times. And then for the outsiders, of couse I also want to thank Camila, Eli, Eva and Konstantina for all the karaoke nights and more. Y bueno, por ultimo quiero igual agradecer a los eK, con los que jugué tantas noches y me ayudaron a sobrevivir durante la pandemia y mas: Alex, Andrés, Andru, Black, Bruno, Jeremías, Jonny, Kevin, Kong, Pallero, Paulo, Point, Seba. Las noches de juego y conversaciones se extrañan mas que nunca.

I want to highlight a few names from those alphabetic lists, because you were also helping me a lot during though times and you hold a special place in my heart, so I think its unfair to not highlight you as you did with me (did that make sense?). So, Ale(j)andra, Camila, Christiane, Eli, Maria and Nico, you guys were very important to me in different stages of my PhD and of my life in general. I don't think I have the words (and time) to write how important you are to me (but hopefully you know it), so thank you for everything and I'm wishing that life always keeps us connected, hopefully I wasn't too much ♡. Finalmente, también quiero agradecer al Alex, el Pablito mechón y el Pallero porque, aún con la distancia y todo, ustedes siempre están ahí para apoyarme. Lamento que la vida haya querido otra cosa y no nos veamos tanto como me gustaría,

pero me tranquiliza saber que, sin importar el momento de la vida en el que estemos, siempre tengo un lugar donde caer con ustedes. Especialmente tu Pallero, mi bro, que sin importar si quiero que esté ahí o no, siempre vuelves ctm. Te amo hermano.

Quiero agradecer a mi familia por apoyarme y estar siempre ahí. Obviamente a mi mamá, que sin su apoyo y amor incondicional quizás me hubiera costado un poquito más. Te amo mamá y ojalá puedas volver a visitarnos pronto ♡. Nacho, Vale, tía Sandra, Meme y Tata, gracias por todo, porque sin ustedes nada de esto hubiera sido posible. También a mi suegra por supuesto, que siempre me recibe con los brazos abiertos. Y puta Bruno, esta tesis es para ti, solo porque así te puedo sacar en cara que tengo un doctorado. Espero que sepas que siempre estaré ahí para tí, asi como yo sé que siempre puedo contar contigo. Te amo hermano y yo sé que lo que sea que te propongas en la vida lo puedes lograr ♡.

Finalmente, Nahir, babe, te amo mucho. Gracias por todo el apoyo que me diste durante estos años. Literal hubiera sido imposible para mí llegar hasta este momento si no te tenía a mi lado. Gracias por soportarme y apoyarme siempre. Los viajes, las salidas y descubrir cosas nuevas juntos hicieron que todo este proceso fuera mucho mas sencillo y, hasta, disfrutable. Por fin seremos una pareja de doctores! Estoy esperando con ansias (como siempre) lo que se venga mas adelante para nosotros ♡.