Cosmology with clusters of galaxies in the eROSITA era

Riccardo Seppi

München 2023

Cosmology with clusters of galaxies in the eROSITA era

Riccardo Seppi

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

vorgelegt von
Riccardo Seppi
aus Cles (TN), Italy

München, den 09/11/2022

Erstgutachter: Prof. Dr. Kirpal Nandra
Zweitgutachter: Dr. habil. Ariel Sanchez
Tag der mündlichen Prüfung: 01/02/2023

A Federica,
Rosella e Angelo,
Marisa e Roberto, Angelina e Adriano.

Inhaltsverzeichnis

Abbildungsverzeichnis

Tabellenverzeichnis

Zusammenfassung

Die Natur Dunkler Materie und Dunkler Energie ist eine der faszinierendsten wissenschaftlichen Fragen dieses Jahrhunderts. Winzige Störungen im frühen Universum entstanden unter der Einwirkung der Gravitation und entwickelten sich zu heute sichtbaren Strukturen wie Galaxien und Haufen. Ein solches Wachstum über kosmische Zeiten hinweg hängt vom Wesen und der Menge dieser mysteriösen dunklen Komponenten ab. Galaxienhaufen bieten ein wichtiges Werkzeug zum Studium der kosmologischen Entwicklung des Universums. Galaxienhaufen sind die massereichsten virialisierten Objekte im Universum, sie befinden sich in den Knoten des kosmischen Netzes und sind ein direkter Tracer der großen Skalenstruktur des Universums. Ihre Häufigkeit als Funktion von Masse und Rotverschiebung (die Halo-Massenfunktion) bietet einen strengen Test für verschiedene kosmologische Modelle und liefert strenge Randbedingungen für die Gesamtmenge an Materie und die Amplitude der Dichteschwankungen im Universum. Neben Informationen zur Dunklen Materie liefern Clusterbeobachtungen auch Hinweise zur Natur der Dunklen Energie.
Das eROSITA Röntgenteleskop wird die größte röntgenselektierte Probe von Galaxienhaufen aller Zeiten liefern, mit dem Potenzial, Dunkle Materie und Dunkle Energie mit beispielloser Präzision zu untersuchen. Um dieses Ziel zu erreichen, ist ein detailliertes Verständnis der von eROSITA erfassten Quellen erforderlich und die Unsicherheiten der Daten sind von grundlegender Bedeutung. Das erste Ziel dieser Arbeit ist es, einen digitalen Zwilling der ersten eROSITA All-Sky Survey (eRASS1) zu produzieren und zu untersuchen, um den Quellkatalog zu charakterisieren und die wissenschaftlichen Analysespipelines zu testen. Wir verwenden neueste Modelle zur Vorhersage der Röntgenemission von Galaxienhaufen und aktiven Galaxiekernen (AGN) und modellieren den Röntgenhintergrund mit realen Daten als Ausgangspunkt. Mit diesem Setup simulieren wir eine Röntgenaufnahme unter Berücksichtigung technischer Details wie der Instrumentenantwort und der Beobachtungsstrategie. Wir verwenden das eROSITA Standard Analyse-Software-System (eSASS), welches einen Quellenkatalog mit sehr ähnlichen Eigenschaften wie der aus realen Daten gewonnene erzeugt. Wir vergleichen die Input und Output Kataloge mit einem Algorithmus basierend auf dem Ursprung jedes Photons. Wir weisen alle hellsten Haufen und AGN nach. Der Anteil der detektierten Quellen hängt in erster Linie von Fluss und Belichtungszeit ab. Zusätzlich sind sekundäre Effekte, wie die Quellengröße und der zentrale Emissionsgrad, für Galaxienhaufen relevant. Wir bieten eine detaillierte Studie über den Kompromiss zwischen Vollständigkeit und Reinheit. Wir stellen fest, dass progressive Grenzwerte für die Detection Likelihood die Hintergrundschwankungen im Quellenkata-

log beseitigen, während Grenzwerte der Extension Likelihood erforderlich sind, um helle Punktquellen zu entfernen, die die Clusterprobe kontaminieren. Wir besprechen verschiedene Wahlmöglichkeiten entsprechend bestimmter wissenschaftlicher Ziele. Zum Beispiel sind zur Studie schwacher Haufen konservative Grenzwerte am besten, wohingegen für ein kosmologisches Experiment eine sicherere Probe mit geringer Kontamination benötigt wird.

Sobald die Auswahl einer Cluster-Probe klar ist, sind genaue Modelle der Massenfunktion Schlüssel, um genaue Einschränkungen kosmologischer Modelle zu erreichen. Das zweite Ziel der Dissertation ist es, ein Modell der Massenfunktion zu kalibrieren, das auch Variablen enthält, die den dynamischen Zustand der Dunklen Materie Halos beschreiben. Wir verwenden die reine Dunkle Materie MultiDark Suite von Simulationen und die darin vorhandenen massereichen Objekte. Wir messen die mittleren Beziehungen von Konzentration, Offset-Parameter (X_{off}) und Spin als Funktion von Halomasse und Rotverschiebung. Wir untersuchen die Verteilungen um die gemittelten Zusammenhänge, und bestätigen den Konzentrationsanstieg bei hohen Massen und stellen ein Modell bereit, das die Konzentration für verschiedene Werte von Masse und Rotverschiebung mit einer einzigen Gleichung vorhersagt. Unsere Untersuchungen zeigen, dass die Konzentration von Halos mit geringer Masse schneller mit der Rotverschiebung abnimmt als bei massereichen Halos. Dies ist im Hochkonzentrationsbereich am stärksten ausgeprägt. Wir finden, dass der Offset-Parameter bei niedriger Rotverschiebung systematisch kleiner ist, was mit der beobachteten Lockerung der Strukturen übereinstimmt. Wir messen die Halomassen-Funktion als Funktion von Offset, Spin und Rotverschiebung. Die einzelnen Modelle werden in ein umfassendes Rahmenwerk kombiniert, das die Massenfunktion als Funktion des Spins und Offset vorhersagt. Unser Modell stellt die Referenzmassenfunktion mit großer Genauigkeit bei Rotverschiebung 0 und berücksichtigt die Rotverschiebungsentwicklung bis zu $z \sim 1.5$.

Der verallgemeinerte Rahmen für die Massenfunktion ermöglicht eine Marginalisierung gegenüber Selektionseffekten im Zusammenhang mit dem dynamischen Zustand von Halos aus Dunkler Materie in einem Cluster-Count-Experiment. Jedoch fehlt eine Verbindung zwischen Beobachtungen und theoretischen Modellen. Das dritte Ziel dieser Arbeit ist die Untersuchung des dynamischen Zustands von Clustern, die von eROSITA beobachtet wurden. Dabei nutzen wir den Offset zwischen dem Röntgen- und dem optischen Zentrum. Wir zielen darauf ab, den in eROSITA Beobachtungen gemessenen Offset mit Vorhersagen aus hydrodynamische Simulationen und N-body-Modelle zu verbinden, und die astrophysikalischen Effekte zu bewerten, die die Verschiebungen beeinflussen. Als Vergleichprobe messen wir den Offset für Cluster, die im eROSITA Final Equatorial-Depth Survey (eFEDS) und eRASS1 beobachten wurden. Wir fokusieren uns auf eine Stichprobe von 87 massiven eFEDS-Clustern bei niedriger Rotverschiebung, und vergleichen die Verschiebungen in dieser Probe zu denen, die von den TNG- und Magneticum-Simulationen vorhergesagt wurden. Dabei verknüpfen wir die Beobachtungen zusätzlich mit dem gemessenen Offset-Parameter in N-body-Simulationen unter Verwendung der hydrodynamischen Simulationen als Verbindungsglied. Die eFEDS Haufen weisen im Durchschnitt einen kleineren Versatz im Vergleich zu eRASS1 auf, da letzteres einen größeren Anteil an massiven und gestörten Strukturen enthält. Der gemessene Versatz auf der eFEDS-Stichprobe stimmt sowohl

mit den Vorhersagen von TNG und Magneticum als auch mit der Verteilung des Offset-Parameter aus Dunkle Materie Simulationen überein. Allerdings sind die Ausläufer der Verteilungen unterschiedlich. Baryonische Effekte bewirken eine Abnahme (Erhöhung) des tiefen (hohen) Offset-Regimes im Vergleich zur X_{off}-Verteilung aus Simulationen mit nur Dunkler Materie. Schlussendlich finden wir eine Korrelation zwischen dem Offset, gemessenen in hydrodynamischen Simulationen, und X_{off}, gemessen in Simulationen mit nur Dunkler Materie, und kalibrieren die Beziehung zwischen ihnen. Diese Beziehung lässt uns die vollständige X_{off}-Verteilung mit hoher Genauigkeit erfassen.

Die in dieser Dissertation entwickelte Arbeit ist wesentlich, um die echte eRASS1-Stichprobe zu charakterisieren, die bald der Öffentlichkeit zugänglich sein wird. Die Entwicklung des innovativen Massefunktionen Frameworks und seine Verknüpfung mit Daten mit dem Offset zwischen verschiedenen Definitionen des Cluster Mittelpunkts ermöglicht die Minimierung von Unsicherheiten in Cluster-Count-Experimenten, die durch Selektionseffekte bezogen auf den dynamischen Zustand des Clusters entstehen.

Abstract

The nature of dark matter and dark energy is one of the most intriguing scientific questions of this century. Tiny density perturbations in the early Universe evolved under the action of gravity, growing into the structures we see nowadays, such as galaxies and clusters. Such growth throughout cosmic time depends on the nature and abundance of these mysterious dark components. Clusters of galaxies provide a great tool for studying the cosmological evolution of the Universe. Galaxy clusters are the most massive virialized objects in the Universe, they reside in the nodes of the cosmic web, and are a direct tracer of the large scale structure of the Universe. Their abundance as a function of mass and redshift (the halo mass function) provides a stringent test for different cosmological models and tight constraints on the total amount of matter and the amplitude of the density perturbations in the Universe, but it also depends on the nature of dark energy.
The eROSITA X-ray telescope will provide the largest X-ray-selected sample of galaxy clusters ever, with the potential of studying dark matter and dark energy with unprecedented precision. A detailed understanding of the sources detected by eROSITA and the uncertainties in the data is fundamental to reach this goal. The first aim of this thesis is to produce and study a digital twin of the first eROSITA all-sky survey (eRASS1), to characterize the source catalog and test the science pipelines. We use state of the art models to predict the X-ray emission from clusters of galaxies and AGN, and model the X-ray background using real data as a starting point. With this setup, we run an X-ray software simulator to produce mock detected photons, accounting for technical details such as the instrument response and the observation strategy. We run the eROSITA Standard Analysis Software System (eSASS), which produces a source catalog with very similar properties to the one obtained from real data. We match the input and output catalogs with an accurate algorithm based on the origin of each photon. We detect all the brightest clusters and AGN in the simulation. The fraction of detected sources primarily depends on flux and exposure time. Secondary effects, such as the source size and the central emissivity, are relevant for clusters. We provide a detailed study of the trade-off between completeness and purity. We find that progressive thresholds of detection likelihood get rid of the background fluctuations in the source catalog, while cuts in extension likelihood are necessary to remove bright point sources contaminating the cluster sample. We discuss different selections according to a given science goal. For example, the inclusion of the majority of sources in the sample is best, if one wants to find interesting objects to study the astrophysics of faint clusters and groups. Instead, a more secure cluster sample with low contamination is required for

a cosmological experiment.

Once the cluster sample is selected, accurate halo mass function models are key to reaching precise constraints on cosmological models. The second aim of this thesis is to calibrate a model of the mass function that also includes variables related to the dynamical state of dark matter halos. We use the dark matter-only MultiDark suite of simulations and the high-mass objects hosting clusters of galaxies therein. We measure the mean relations of concentration, offset parameter (X_{off}), and spin as a function of dark matter halo mass and redshift. We investigate the distributions around the mean relations. We confirm the recent discovery of the concentration upturn at high masses and provide a model that predicts the concentration for different values of mass and redshift with one single equation. We find that the concentration of low-mass halos shows a faster redshift evolution compared to high-mass halos, especially in the high-concentration regime. We find that the offset parameter is systematically smaller at low redshift, in agreement with the relaxation of structures at recent times. The individual models are combined into a comprehensive framework, which predicts the mass function as a function of spin and offset. Our model recovers the fiducial mass function with great accuracy at redshift 0 and accounts for redshift evolution up to $z \sim 1.5$.

The generalized mass function framework allows marginalizing over selection effects related to the dynamical state of dark matter halos in a cluster count experiment. However, a link between observations and theoretical models is lacking. The third goal of this thesis is to study the dynamical state of clusters detected by eROSITA using the offset between the X-ray and the optical centers. We aim to connect the offset measured in eROSITA observations to predictions by hydrodynamical simulations and N-body models, assessing the astrophysical effects affecting the displacements. We measure the offset for clusters observed in the eROSITA Final Equatorial-Depth Survey (eFEDS) and eRASS1. We focus on a subsample of 87 massive eFEDS clusters at low redshift. We compare the displacements in this sample to those predicted by the TNG and the Magneticum simulations. We link the observations to the offset parameter X_{off} measured for dark matter halos in N-body simulations, using the hydrodynamical simulations as a bridge. We find that on average the eFEDS clusters show a smaller offset compared to eRASS1 because the latter contains a larger fraction of massive and disturbed structures. The offset measured in the eFEDS subsample is in agreement with the predictions from TNG and Magneticum, and the distribution of the offset parameter from dark matter only simulations. However, the tails of the distributions are different. Baryonic effects cause a decrement (increment) in the low (high) offset regime compared to the X_{off} distribution from dark matter-only simulations. Finally, we find a correlation between the offset predicted by hydro simulations and X_{off} measured in their parent dark matter-only run and calibrate a relation between them, which allows us to recover the full X_{off} distribution with excellent precision.

The work developed in this thesis is essential to characterize the real eRASS1 sample, that will soon be available to the public. The development of the innovative mass function framework and its link to data with the offset between different definitions of the cluster center will allow the minimization of uncertainties in cluster count experiments due to selection effects related to the cluster dynamical state.

Kapitel 1

Introduction

In this chapter, we introduce the description of the Universe on different scales, including the concepts of dark matter, dark energy, halos, and cosmological models. We summarize the main properties of clusters of galaxies and the focal points of their usage in cosmological experiments. We finally introduce the eROSITA instrument, its cosmological goals, and the complete outline of this thesis.

1.1 Cosmology

The General theory of Relativity (GR) provides the most accurate description of the Universe to date. It was formulated by Einstein (1916). Nowadays we include dark matter and dark energy within this framework (Planck Collaboration et al., 2020a).
Independent experiments point to a standard cosmological model: a flat Universe in a state of accelerated expansion driven by a cosmological constant Λ, where most of the matter is in form of invisible cold dark matter (CDM). This model is named ΛCDM (e.g., Mo and White, 2002). Such experiments include studies of supernovae Type Ia (SNIa, Riess et al., 1998, 2022), that are used as standard candles to study the relation between their distance and their redshift; the cosmic microwave background (CMB, Komatsu et al., 2011; Planck Collaboration et al., 2020a), encoding the imprint of the matter distribution in the early Universe; the clustering of galaxies and baryonic acoustic oscillations (BAO, Eisenstein et al., 2005; Cole et al., 2005; Beutler et al., 2011; Sánchez et al., 2012; Alam et al., 2021), that summarize the distribution of galaxies in the Universe; weak gravitational lensing (Hildebrandt et al., 2017; Asgari et al., 2021), that probes the distribution of matter throughout spacetime; or the abundance of galaxy clusters (Mantz et al., 2015b; Bocquet et al., 2019; Ider Chitham et al., 2020; To et al., 2021), that traces the evolution of the most massive and rarest objects throughout cosmic time. These probes are complementary because their combination constrains both the geometry of the Universe and the growth of structures simultaneously. The time evolution of the ΛCDM model is illustrated in Fig. 1.1. Although dark matter and dark energy respectively contribute to about 25% and 70% of the total mass and energy budget of the Universe, their nature remains unclear.

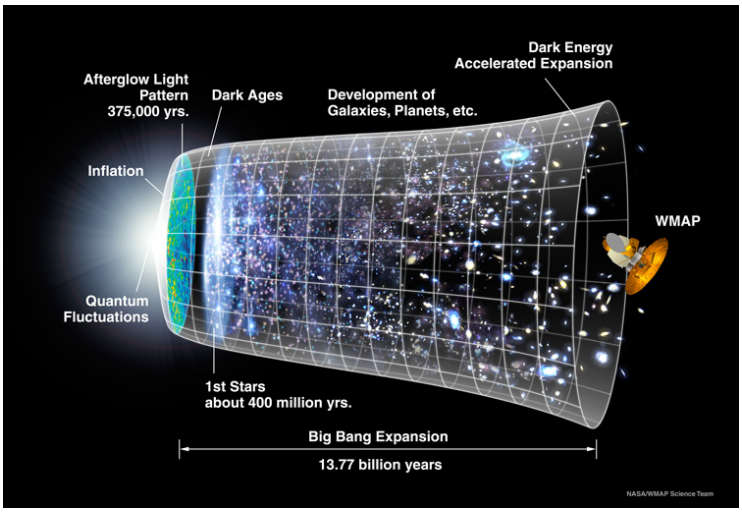

Abbildung 1.1: Evolution of the Universe throughout cosmic time. It began at the Big Bang about 13.7 billion years ago. The cosmic microwave background is a collection of photons left free to propagate uniformly after free electrons coupled with atomic nuclei into neutral hydrogen. Temperature fluctuations of the CMB encode small density perturbations originated in the early phases of the Universe. The initial overdensities grew due to gravity, forming the galaxies we see nowadays.
Credit: NASA/WMAP Science Team.

The term *dark matter* was first introduced by Kapteyn (1922), who studied the oscillation of stars around the galactic plane and found that the amount of invisible matter is negligible around the Solar System. The first evidence for the presence of dark matter was shown by Zwicky (1933), who measured the velocity dispersion in the Coma cluster and estimated that the cluster should be about 400 times more massive than the sum of all the visible matter (Briel et al., 1992). This first estimate did not include the X-ray emission from the hot gas, which makes the total amount of matter about five times larger than the visible components, but that is still not enough. In the 1970s and 1980s, further confirmation came from galaxy rotation curves (Shostak, 1978; Rubin and Ford, 1970; Rubin et al., 1980). The phenomenon of gravitational lensing of distant sources additionally provides evidence of the existence of dark matter, as the observed deflections are not explained by visible matter alone (Soucail et al., 1987). In this context, the smoking gun is the famous bullet cluster (Markevitch et al., 2002; Clowe et al., 2006): the total mass distribution inferred by weak gravitational lensing extends to larger distances compared to the X-ray emission from the hot gas. Current observations favor a cold dark matter scenario, with particle velocities that are negligible compared to the Hubble flow, allowing the growth of perturbations into massive halos (Peebles, 1982; Blumenthal et al., 1984; Viel et al., 2013). Understanding the nature of dark matter is one of the biggest challenges in science, it connects a variety of fields from cosmology to particle physics (see Arbey and Mahmoudi, 2021, for a review).

The concept of dark energy is nowadays related to the cosmological constant Λ. The introduction of the cosmological constant dates back to the beginning of the 1900s. A first analytical solution to the equations of general relativity by Friedmann (1922) described an expanding Universe. At that time, Einstein was convinced that the Universe was static. Therefore, he introduced a cosmological constant Λ in his equations to allow a solution describing a static, matter-dominated Universe (de Sitter, 1917). However, the discovery of the expansion of the Universe by Hubble (1929) and Lemaître (1931) provided evidence that the Universe is not static, which lead Einstein to remove the cosmological constant from GR. Nonetheless, more recent work by Riess et al. (1998); Perlmutter et al. (1999) using SNIa as standard candles proved that the Universe is not only expanding but is also accelerating while doing so. The additional component explaining the cosmic acceleration has been named *dark energy*. One of the most simple descriptions of dark energy is the cosmological constant, which has therefore been reintroduced in the GR framework. Investigating the nature of dark energy and its equation of state is a major challenge of modern cosmology (see Weinberg et al., 2013; Will, 2014; Ishak, 2019, for reviews).

We start by describing the concordance model of the Universe, and how to extract cosmological information about it from the collapse of initial perturbations into the structures we observe nowadays.

1.1.1 The Universe on large scales

The theory of GR describes gravity as the curvature of spacetime due to the presence of matter and energy. This concept is encapsulated in the field equation 1.1:

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = \frac{8\pi G}{c^4}T_{\mu\nu} + \Lambda g_{\mu\nu}, \tag{1.1}$$

where the curvature of spacetime is measured by the Ricci tensor $R_{\mu\nu}$ and the Ricci scalar R, $g_{\mu\nu}$ is the metric describing the geometry of the Universe, $T_{\mu\nu}$ is the mass-energy tensor and Λ is the cosmological constant, often associated with dark energy. Geometrical factors encoding the curvature of spacetime are on the left-hand side. Instead, the sources that shape such curvature, such as mass, energy, and the cosmological constant, are on the right. Eq. 1.1 constitutes a set of non-linear, coupled differential equations, which has an analytical solution only under specific circumstances. For example, the Universe is homogeneous and isotropic on scales larger than hundreds of megaparsec (hereafter Mpc). This is the notion of the cosmological principle. Under this assumption, the metric describing the infinitesimal distance between two points in spacetime is written according to Eq. 1.2:

$$dS^2 \;=\; c^2dt^2 - a^2(t)\left[\frac{dr^2}{1-Kr^2} + r^2(d\theta^2 + \sin^2(\theta)d\phi^2)\right]. \tag{1.2}$$

This is the Friedmann-Lemaître-Robertson-Walker metric (FLRW, Robertson, 1929), where c is the speed of light, and $a(t)$ is the scale factor of the Universe, a time-dependent factor that gives dimensionality to the space and angular coordinates r, θ, ϕ. The geome-

try of the Universe is encoded in the factor k, which can assume three values: +1, 0, -1, respectively describing a spherical, flat, and hyperbolic Universe.

The scale factor $a(t)$ is directly linked to the concept of cosmological redshift. In the late 1920s, Hubble (1929) and Lemaître (1931) provided evidence for the expansion of the Universe by measuring an increasing recession velocity of galaxies as a function of their distance from Earth. They observed spectral lines at redder wavelengths for further galaxies because the photons are stretched by the expansion of the Universe as they propagate through spacetime. For a source emitting light at a wavelength λ_E, that is observed at a longer wavelength λ_O, the redshift is defined as

$$z = \frac{\lambda_O - \lambda_E}{\lambda_E}. \tag{1.3}$$

Within a Universe described by the FLRW metric, the cosmological redshift is directly related to the scale factor at the time of emission $a(t_E)$ and the one at the time of the observation $a(t_O)$ by

$$1 + z = \frac{a(t_O)}{a(t_E)}. \tag{1.4}$$

Starting from Eq. 1.1, assuming the cosmological principle and that the Universe is described by a fluid fully characterized by its mass and pressure (i.e., a perfect fluid), one can derive the Friedmann equations:

$$\begin{aligned} \dot{H} + H^2 &= -\frac{4}{3}\pi G(\rho + \frac{3P}{c^2}) + \frac{\Lambda c^2}{3} \\ H^2 &= \frac{8}{3}\pi G\rho - \frac{Kc^2}{a^2} + \frac{\Lambda c^2}{3}, \end{aligned} \tag{1.5}$$

where $H = \dfrac{\dot{a}}{a}$ is the Hubble parameter and describes the expansion rate of the Universe. Assuming that the Universe is adiabatic and that different components of the Universe are regulated by an equation of state in the form

$$P = w\rho c^2. \tag{1.6}$$

Combining Eq. 1.6 with Eq. 1.5 provides the density evolution of each component as a function of cosmic time:

$$\rho_w(t) = \rho_{w,0}\left(\frac{a(t)}{a_0}\right)^{-3(1+w)}, \tag{1.7}$$

where $\rho_{w,0}$ is the present day density, and a_0 is the present day scale factor. It is usually normalized to 1.

The main components of the Universe are ordinary non relativistic matter (w=0), radiation (w=1/3), and dark energy (w=-1). From Eq. 1.7 it follows that the dark energy density does not evolve with time, which is why it is related to the cosmological constant.

From Eq. 1.5, one can define the critical density of the Universe, the dark energy density, and the curvature density as:

$$\rho_c = \frac{3H^2}{8\pi G} \qquad \rho_\Lambda = \frac{\Lambda c^2}{8\pi G} \qquad \rho_K = -\frac{3Kc^2}{8\pi G}. \tag{1.8}$$

If the density of the Universe is equal to the critical density, its curvature goes to zero. Finally, combining Eqs. 1.5, 1.7, and 1.8, and dividing the density of each i component by ρ_c to obtain the density parameter $\Omega_i = \frac{\rho_i}{\rho_c}$, it is possible to write the redshift evolution of the Hubble parameter:

$$E(z) = \frac{H(z)}{H_0} = \sqrt{\Omega_M(1+z)^3 + \Omega_\Lambda(1+z)^{3(1+w)} + \Omega_R(1+z)^4 + \Omega_K(1+z)^2}, \tag{1.9}$$

where H_0 is the Hubble parameter measured today and is named Hubble constant (or Hubble-Lemaitre constant), and $\Omega_K = 1-\Omega_M-\Omega_R-\Omega_\Lambda$ is the curvature density parameter. The definition of a reference frame at rest with respect to the Hubble flow provides the *comoving distance* between two points by integrating Eq. 1.9 (Weinberg et al., 2013):

$$D_C = D_H \int_0^z \frac{dz'}{E(z')}, \tag{1.10}$$

where D_H=c/H_0 is the Hubble distance. The distance between two objects at fixed redshift is encoded in the *transverse comoving distance* (Hogg, 1999):

$$D_M(z) = \begin{cases} \frac{D_H}{\sqrt{\Omega_K}} \sinh \frac{\sqrt{\Omega_K D_C(z)}}{D_H} & \text{for } \Omega_K > 0 \\ D_C(z) & \text{for } \Omega_K = 0 \\ \frac{D_H}{\sqrt{\Omega_K}} \sin \frac{\sqrt{|\Omega_K| D_C(z)}}{D_H} & \text{for } \Omega_K < 0. \end{cases} \tag{1.11}$$

In the literature, the Hubble constant is often expressed as $H_0 = h \times 100$ km/s/Mpc. Because of the minus sign in the definition of ρ_K, Ω_K assumes positive, negative, or zero values corresponding to k = -1, 1, and 0, describing an open, closed, and flat Universe respectively.

The physical size of an object and its angular extent in the sky are related by the *angular diameter distance*, which is computed from Eq. 1.2 by aligning dt=dr=dϕ=0 and corresponds to:

$$D_A = \frac{D_M}{(1+z)}. \tag{1.12}$$

The definition of the *luminosity distance* preserves the relation between flux F and luminosity L by the distance squared and accounts for intensity dimming due to the cosmological expansion of the Universe, the gravitational redshift of the photons, and time dilation effects. It is related to D_A according to Eq. 1.13.

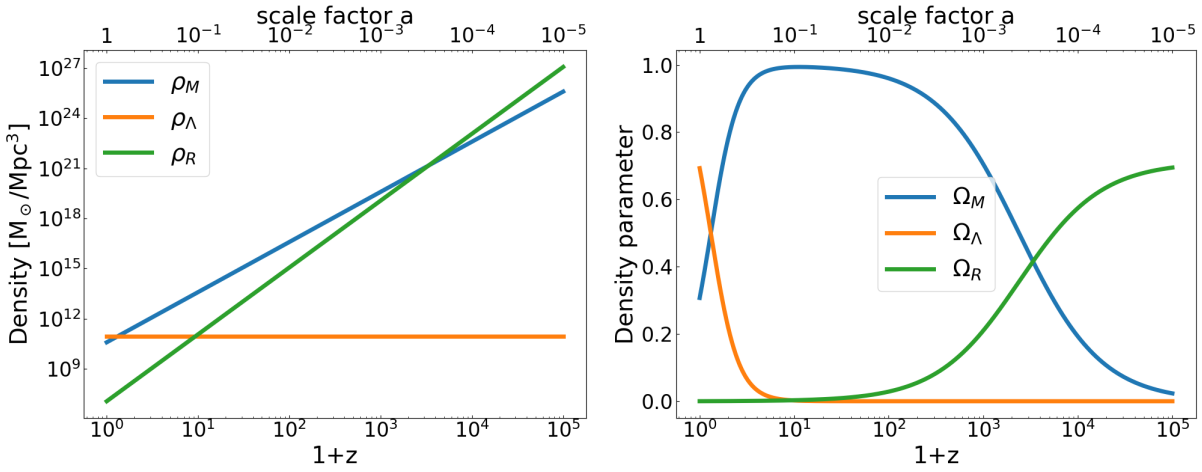

Abbildung 1.2: Density evolution of total matter (in blue), dark energy (in orange), and radiation (in green) throughout cosmic time. **Left-hand panel**: Density evolution in units of $M_\odot/Mpc^3$. **Right-hand panel**: Evolution of the density parameter (i.e., normalized to the critical density ρ_C.) Both figures show the redshift on the lower x-axis and the scale factor on the upper x-axis. The figures have been produced assuming cosmological parameters from Planck Collaboration et al. (2020a).

$$D_L = \sqrt{\frac{L}{4\pi F}}$$
$$D_L = D_A(1+z)^2 \tag{1.13}$$

The density evolution of the total matter, dark energy, and radiation components is shown in Fig. 1.2. The left-hand panel displays the evolution of densities, the right-hand one shows the density parameters. In very early times, the Universe was dominated by radiation. In a second phase, after the matter-radiation equivalence at $z_{EQ} \approx 3000$, the dominant component was matter. Finally, dark energy took over only in recent times at $z \lessapprox 0.7$, when the radiation is negligible. The plots are produced assuming cosmological parameters from Planck Collaboration et al. (2020a). The main ones describing the ΛCDM model assume values of $\Omega_M = 0.3087$, $\Omega_\Lambda = 0.6913$, and $H_0 = 67.66$ km/s/Mpc.

1.1.2 The Universe on small scales

The cosmological principle is a very good assumption only on very large scales $\gtrapprox 250$ Mpc (Laurent et al., 2016). On smaller scales, we see that the Universe is not uniform or smooth, as matter clumps in strong overdensity peaks forming galaxies and clusters of galaxies. We can describe the distribution of matter throughout the Universe as local perturbations compared to the average density defined as:

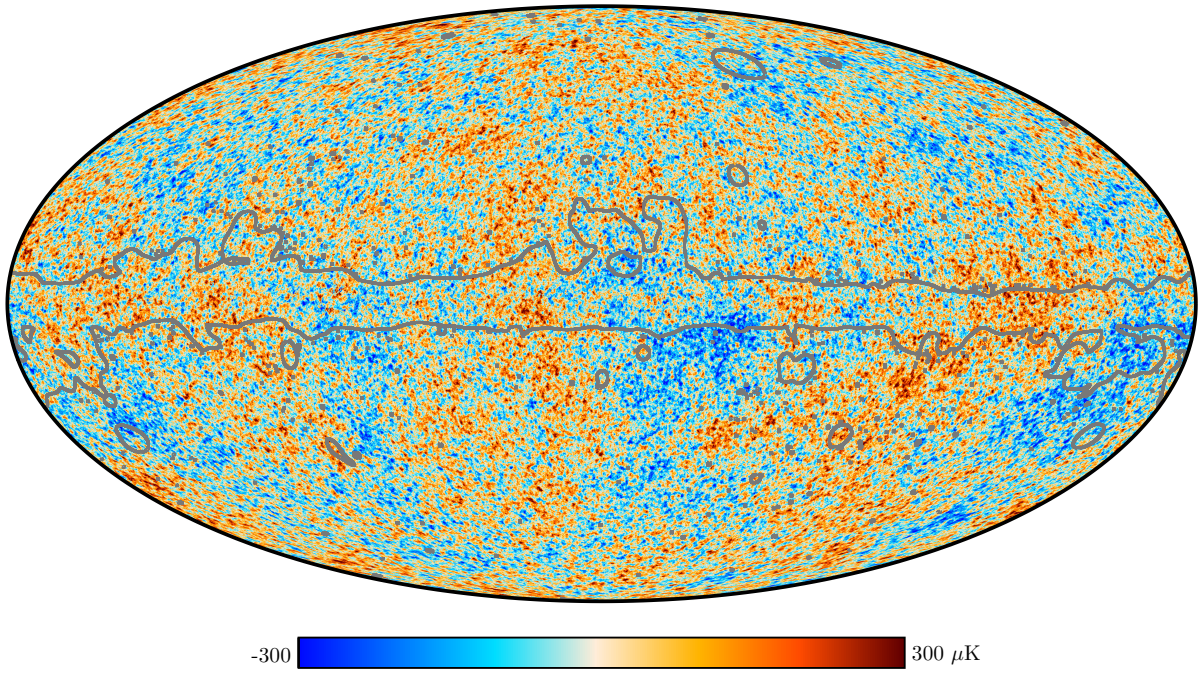

Abbildung 1.3: Temperature fluctuations of the cosmic microwave background measured by Planck Collaboration et al. (2020a). Redder (bluer) regions denote hot (cold) spots.

$$\delta = \frac{\rho - < \rho >}{< \rho >}. \tag{1.14}$$

The temperature fluctuations measured by Planck Collaboration et al. (2020a) yield initial density perturbations of the order $\delta \approx 10^{-5}$. These overdensities collapsed under the action of gravity until reaching values $\delta \approx 10^2$–10^3 that we observe nowadays in collapsed structures.
It is possible to describe the evolution of the overdensities by considering the early Universe as a collisional gas in an expanding background. Its evolution is regulated by the continuity, Euler, and Poisson equations:

$$\begin{aligned} \frac{\partial \rho}{\partial t} + \nabla \cdot \rho \mathbf{v} &= 0 \\ \frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla)\mathbf{v} + \frac{1}{\rho}\nabla P + \nabla \Phi &= 0 \\ \nabla^2 \Phi - 4\pi G \rho &= 0. \end{aligned} \tag{1.15}$$

Assuming small perturbations $\delta \ll 1$, an equation of state relating density and pressure in the form of Eq. 1.6, we can find a solution to Eq. 1.15 in form of $\rho =< \rho > +\delta$, $P =< P > +\delta P$, $\Phi =< \Phi > +\delta\Phi$, $\mathbf{v} =< v > +\delta\mathbf{v}$. Plugging these terms in Eq. 1.15, considering

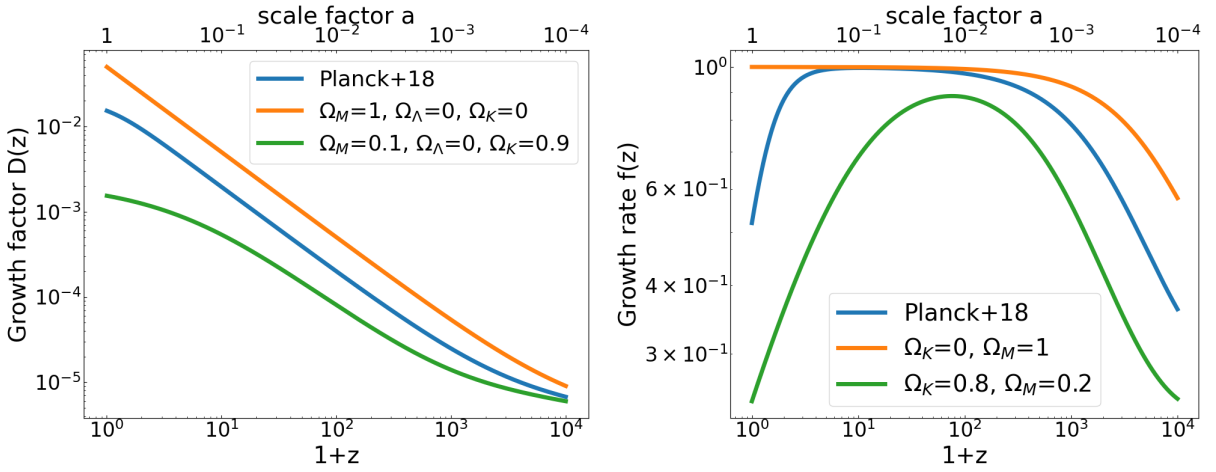

Abbildung 1.4: Growth of perturbations as a function of cosmic time. **Left-hand panel**: growth factor D(z), see Eq. 1.17. **Right-hand panel**: growth rate f(z), see Eq. 1.18. The plots show three different cosmological models: the one constrained by Planck Collaboration et al. (2020a) in blue, a flat EdS Universe full of matter in orange, and an open Universe with no dark energy in green.

a solution in the form of waves $f(x,t) = f_K(t)\ \exp[iKx]$, and ignoring higher order terms yields:

$$\ddot{\delta} + 2H\dot{\delta} + [\frac{k^2}{a^2}wc^2 - \frac{3}{2}H^2\Omega_M]\delta = 0. \tag{1.16}$$

For cold dark matter (w=0), Eq. 1.16 has a growing solution that reads:

$$\delta_+(t) = D(z) = -H(z)\int_\infty^z \frac{1+z'}{H^3(z')}dz', \tag{1.17}$$

where D(z) is named growth factor, and its logarithmic derivative is the growth rate f:

$$f(z) = \frac{d\ln D(z)}{d\ln a} \approx \Omega_M^\gamma + \frac{\Omega_\Lambda}{70}(1+\frac{\Omega_M}{2}), \tag{1.18}$$

where the exponent $\gamma \approx 0.55$ is a stringent test of general relativity (Weinberg et al., 2013).

The evolution of the growth factor and growth rate as a function of cosmic time is shown in Fig. 1.4. It shows three types of Universe: a fiducial ΛCDM in blue, an EdS model in orange, and an open one with no cosmological constant in green. Perturbations grow faster in models with a larger amount of matter and a lower amount of dark energy, which prevents growth due to cosmic acceleration. Such a process is responsible for shaping the Large Scale Structure (LSS) of the Universe.

After the epoch of radiation-matter equivalence and especially after the decoupling between baryons and photons at $z_{DEC} \approx 300$, the assumption that the Universe is a collisional gas breaks down. It is therefore necessary to account for each component in Eq. 1.15. The

result is that in the era dominated by radiation at z $\gtrapprox$ 3000, the perturbations always grow outside the cosmological horizon $R_H = \int \frac{cdt}{a(t)}$. Inside the horizon instead, the growth of dark matter overdensities is limited by the *Meszaros effect* (Meszaros, 1974). Solving Eq. 1.15 for dark matter in the radiation era yields:

$$\delta_{DM}(x) = 1 + \frac{3}{2}x, \tag{1.19}$$

where $x = \frac{a}{a_{eq}}$. Dark matter perturbations can therefore only grow up to a factor of 5/2 until equivalence on small scales.

During the matter-dominated era, baryonic perturbations grow linearly with the scale factor $\delta_B \propto a$. Since the last scattering surface between baryons and radiation is located at the recombination of protons and electrons into neutral hydrogen (z $\approx$ 1000), the initial perturbations should have been of the order $\delta_B/\rho_B \sim 10^{-3}$, to explain the non-linear structures that we observe in the Universe. Nonetheless, the CMB provides evidence for $\delta_B/\rho_B \sim 10^{-5}$. This apparent inconsistency is solved by dark matter. Dark matter does grow between equivalence and decoupling, while baryons are still coupled to the oscillating radiation. The dark matter creates potential wells into which the baryons fall after decoupling. The solution to Eq. 1.15 for baryonic overdensities in the matter-dominated era after decoupling reads:

$$\delta_B = \delta_{DM}(1 - \frac{a_{DEC}}{a}). \tag{1.20}$$

This equation is also known as *baryon catch up*. The evolution of perturbations for dark matter, baryonic matter, and radiation is shown in Fig. 1.5. This figure shows the specific case of a perturbation entering the cosmological horizon at z_{EQ}. The dark matter (solid black line) grows then linearly with the scale factor, while the baryons (dash-dotted line) are coupled with the oscillating radiation (solid grey line). After z_{DEC}, the baryons are not coupled to radiation anymore. The latter keeps oscillating, but loses the baryons' support and rescales to a lower overdensity. The baryons are free to fall into the potential wells created by dark matter and catch up to the dark matter overdensity at later times. This allows them to reach $\delta_B \approx 1$ at recent times when the non-linear evolution takes over, and is responsible for the formation of galaxies and clusters of galaxies.

1.1.3 Power spectrum

Measurements of the CMB spectrum point to a Gaussian isotropic field of primordial density fluctuations (Komatsu et al., 2009; Planck Collaboration et al., 2020b). Under this assumption, one can describe the density contrast by its Fourier transform:

$$\delta(\mathbf{k}) = \frac{1}{(2\pi)^{3/2}} \int d^3x \delta(\mathbf{x}) e^{i\mathbf{k}\cdot\mathbf{x}}, \qquad \delta(\mathbf{x}) = \frac{1}{(2\pi)^{3/2}} \int d^3k \delta(\mathbf{x}) e^{-i\mathbf{k}\cdot\mathbf{x}}, \tag{1.21}$$

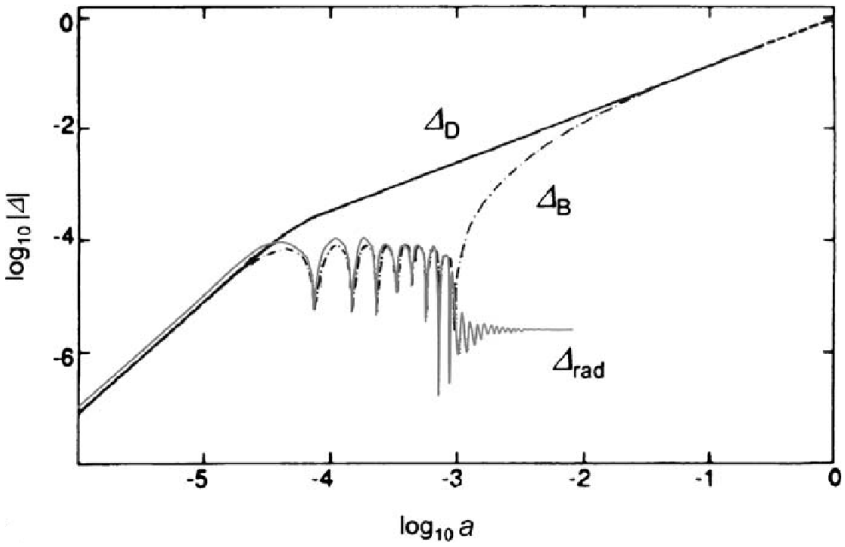

Abbildung 1.5: Evolution of baryonic matter δ_B (in this figure Δ_B, dash-dotted line), dark matter δ_{DM} (in this figure Δ_D, solid black line), and radiation δ_R (in this figure Δ_{rad}, grey solid line) perturbations as a function of cosmic time. This panel is adapted from Longair (2008).

where k is the wavenumber. The power spectrum P(k) provides a complete description of the initial Gaussian field. It is defined by:

$$< \delta(k)\delta(k') >= (2\pi)^3 P(k)\delta^3(k-k'), \tag{1.22}$$

where δ^3 is the three dimensional Dirac delta function. The Fourier analogue of the power spectrum is the two-point correlation function ξ(r):

$$\xi(r) =< \delta(\mathbf{x})\delta(\mathbf{x}+\mathbf{r}) >= \frac{1}{2\pi^2}\int dk k^2 P(k)\frac{\sin(kr)}{kr}, \tag{1.23}$$

which describes the excess probability of finding two overdensities separated by a radius r.

The power spectrum is proportional to the overdensity field squared, therefore the growth factor regulates its evolution. In addition, the shape of the power spectrum is modified by a transfer function T(k,z) encoding the interactions between radiation and baryons. The transfer function is computed by solving Eq. 1.16 for each component. Different physical processes are encoded in the formulation of the transfer function (Eisenstein and Hu, 1998). Assuming that the redshift dependence of the power spectrum is encoded in the growth factor and the transfer function, the evolution of the power spectrum is then described according to:

$$P(k,z) = P_p(k)D^2(z)T^2(k,z), \tag{1.24}$$

where P_p is the primordial power spectrum, describing scale-invariant perturbations in the form $P_p(k) \propto k^{n_s}$, where the index n_s is the slope of the linear matter power spectrum.

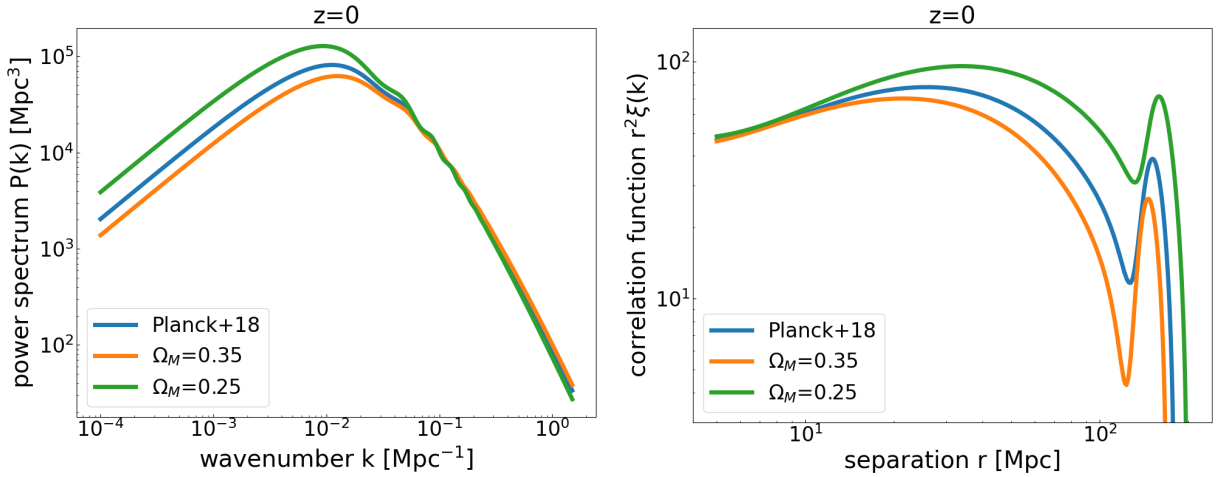

Abbildung 1.6: Linear power spectrum (left-hand panel) and correlation function (right-hand panel) at z=0 for different ΛCDM cosmological models: the fiducial Planck Collaboration et al. (2020a) in blue, and models with larger and lower Ω_M in orange and green.

The large scales are unaffected by baryons and carry information about the original primordial perturbations: the transfer function is equal to 1 at small k. The transfer function instead describes the suppression of power on small scales due to the presence of baryons. The peak of the power spectrum corresponds to the last scale that entered the horizon without suffering from the Meszaros effect, it is therefore located at the horizon scale at matter radiation equality $k_{EQ} \approx 10^{-2}$ Mpc^{-1}. Larger values of Ω_M cause a faster evolution of the Universe to z_{EQ}, which is reflected in a shift of the power spectrum peak. The power spectrum also shows some wiggles at k ≈ 0.1. Such wiggles are the baryon acoustic oscillations (BAO, Peebles and Yu, 1970; Sunyaev and Zeldovich, 1970), caused by the propagation of acoustic waves in the primordial plasma when the baryons were coupled to radiation. In this plasma, the baryons tended to fall in the gravitational wells formed by dark matter, but the radiation pressure was pushing them in the opposite way. After decoupling, the photons are free to stream through the Universe, which leaves a shell of fixed baryonic matter distribution, that subsequently evolved according to the background cosmic expansion. This is visible nowadays as an overdensity of galaxies, which can be measured in the correlation function at the sound horizon scale, and used as a standard ruler to constrain the geometry of the Universe (Eisenstein et al., 2005). The power spectrum and the corresponding correlation function for different cosmological models at z=0 are shown in Fig. 1.6. The BAO wiggles at intermediate scales in the power spectrum are visible as a peak of the correlation function at r ≈ 150 Mpc.

The variance of the smoothed density field is defined as

$$\sigma^2(M,z) = \int \frac{d^3k}{(2\pi)^3} \hat{W}^2(k,R) P(k,z), \tag{1.25}$$

where k is the wavenumber, $\hat{W}$ the Fourier transform of a top-hat filtering function given

by

$$\hat{W} = 3\frac{\sin(kR) - kR\cos(kR)}{k^3R^3}, \tag{1.26}$$

and $P(k,z) = D(z)^2P_p(k)$ the linear matter power spectrum. Its redshift evolution is encoded in the growth factor D(z). The amplitude of the linear matter power spectrum from Eq. 1.25 on scales of 8 Mpc/h is denoted as σ_8. Sánchez et al. (2022) propose an innovative approach, breaking the degeneracy between cosmological parameters that only affect the shape of the power spectrum from the ones that only affect its amplitude. Models with the same shape parameters are then linked by a redefinition of the redshift to match the amplitude of the power spectrum on scales of 12 Mpc, denoted as σ_{12}.
When the growth of perturbations is not linear, the Gaussian field assumption does not hold anymore: the power spectrum and the variance σ(M,z) do not provide a complete description of the distribution of δ. Higher order terms are then necessary to capture non-linear effects on small scales (Smith et al., 2003).

1.1.4 Halo formation

Once the density contrast reaches values of $\delta > 1$, non-linear theory is necessary to describe the formation of halos. The non-linear collapse of perturbations into dark matter halos is modeled as a spherical perturbation viewed as a closed universe evolving in an expanding background. This is a one-dimensional problem and is fully described by the evolution of a top-hat radius (Gunn and Gott, 1972). The overdensity expands together with the background in the initial phase until it reaches a maximum radius at the epoch of turnaround. At this point, the perturbation decouples from the background expansion, it starts collapsing and ultimately virializes into a halo.
This process is described by a set of equations similar to Friedmann equations 1.5 (Liddle and Lyth, 2000). A parametric solution for the scale factor describing the evolution of the perturbation as a function of cosmic time reads:

$$\begin{aligned}\frac{a(t)}{a_0} &= (1-\cos\theta)\frac{\Omega_0}{2(\Omega_0-1)} \\ \frac{\dot{a}}{a_0}t &= (\theta-\sin\theta)\frac{\Omega_0}{2(\Omega_0-1)^{3/2}},\end{aligned} \tag{1.27}$$

where θ is an evolution angle spanning from 0 to 2π, and Ω_0 is the density of the perturbation in its initial stage. Expanding Eq. 1.27 for $\theta \ll 1$ gives:

$$\begin{aligned}\frac{a(t)}{a_{\mathrm{MAX}}} &\simeq \frac{\theta^2}{4} - \frac{\theta^4}{48} \\ \frac{t}{t_{\mathrm{MAX}}} &\simeq \frac{1}{\pi}\left(\frac{\theta^3}{6} - \frac{\theta^5}{120}\right),\end{aligned} \tag{1.28}$$

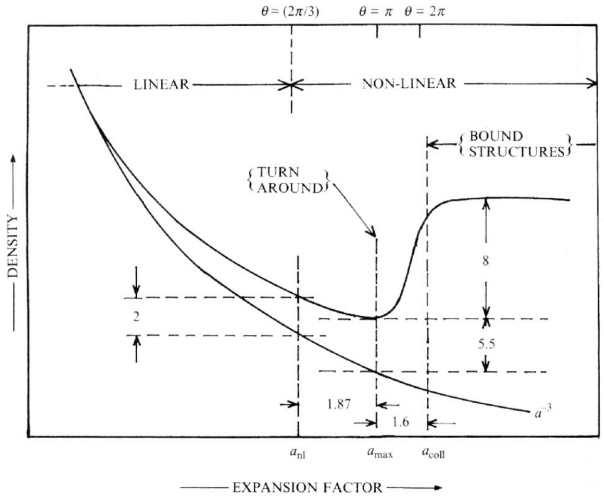

Abbildung 1.7: Non-linear evolution of a density perturbation (top black line) in an expanding background (bottom black line). The figure is taken from Padmanabhan (1993).

where $a_{\rm MAX}$ and $t_{\rm MAX}$ are the value of the scale factor and the time of maximum expansion of the perturbation. These two equations can be combined into:

$$\frac{a_{\rm LIN}(t)}{a_{\rm MAX}} \simeq \frac{1}{4}\left(6\pi\frac{t}{t_{\rm MAX}}\right)^{2/3}\left[1-\frac{1}{20}\left(6\pi\frac{t}{t_{\rm MAX}}\right)^{2/3}\right], \tag{1.29}$$

which describes the evolution of a perturbation in linear approximation.

Combining Eq. 1.29 at the time of maximum expansion with Eq. 1.7 for a matter-dominated Universe gives a solution for the linear density contrast that reads:

$$\delta_{LIN} = \frac{3}{20}\left(6\pi\frac{t}{t_{MAX}}\right)^{2/3}, \tag{1.30}$$

where $t_{MAX} = \frac{\pi}{2H_0}\frac{\Omega_M}{(\Omega_M-1)^{3/2}}$ is the instant of maximum expansion of the perturbation before collapsing. This moment is named turnaround, when the density contrast is then equal to:

$$\delta_{LIN}^{TA} = \frac{3}{20}(6\pi)^{2/3} \approx 1.06. \tag{1.31}$$

At this point, structures depart from the background expansion and the halo formation reaches the end of the linear regime. After the turnaround, the perturbation collapses symmetrically to the background expansion. It is common to define the perturbation as a collapsed halo at $t=2t_{MAX}$, when the density contrast reaches:

$$\delta_{LIN}^{CO} = \frac{3}{20}(12\pi)^{2/3} \approx 1.686. \tag{1.32}$$

This value has very weak cosmological dependence and is historically used to describe structure formation universally. It corresponds to the final stage of the gravitational collapse for a spherically symmetric perturbation.

In reality, including non-linear terms in Eq. 1.27 is necessary for a more accurate description of halo collapse after the turnaround. The non-linear evolution of a density perturbation is shown in Fig. 1.7. The non-linear overdensity at turnaround reads:

$$1 + \delta_{NL}^{TA} = \frac{a_{BKG}^3}{a_{MAX}^3} = \frac{(6\pi)^2}{4^3} \approx 5.55. \tag{1.33}$$

In addition, the virial theorem yields that the radius of a collapsed virialized halo $R_{VIR} = R_{MAX}/2$. Therefore, the density of the halo at virialization is $\rho_{VIR} = 8\rho_{MAX}$. Moreover, the background density decreases by a factor of 4 from the turnaround to virialization, so that:

$$1 + \delta_{NL}^{VIR} \approx 178. \tag{1.34}$$

In contrast to the linear density contrast, the non-linear derivation significantly depends on cosmological parameters. In particular, for a universe with $\Omega_R = 0$, Bryan and Norman (1998) find

$$\begin{aligned} \Delta_{\rm vir} &= (18\pi^2 + 82x - 39x^2)/\Omega(z), \\ \Omega(z) &= \Omega_{\rm M}(1+z)^3/E(z)^2, \end{aligned} \tag{1.35}$$

where $x = \Omega(z) - 1$, Ω_M is the matter density parameter at the present day, and E(z) is the Hubble parameter as a function of redshift in units of H_0 (see Eq. 1.9).

1.1.5 Halo mass function

Since the initial density field is Gaussian and a threshold for the halo collapse is fixed by Eq. 1.32, a prediction for the number of halos arising from the initial perturbations is achievable by studying the behavior of peaks in a random field (BBKS, Bardeen et al., 1986). Considering the density field smoothed on a certain scale:

$$R = \left(\frac{3M}{4\pi}\right)^{1/3}, \tag{1.36}$$

the fraction of space exceeding the linear density contrast ends up in collapsed halos with mass larger than one corresponding to the chosen scale in Eq. 1.36. However, the unique identification of an overdensity in the linear density field with a collapsed halo faces a critical issue: small halos are lost when smoothing the field on large scales. This is the cloud-in-cloud problem. A similar approach was taken by Press and Schechter (1974). They predicted the expected number of dark matter halos in the Universe starting from the simple assumption that the probability for a density perturbation in the primordial field to overcome $\delta > \delta_{LIN}^{CO}$ is equal to the fraction of mass contained in virialized halos.

We refer to δ_{LIN}^{CO} as δ_c in the rest of this thesis. Since the density field is assumed to be Gaussian so that

$$P(\delta) = \frac{1}{\sqrt{2\pi}\sigma_M} e^{-\frac{\delta^2}{2\sigma_M^2}}, \tag{1.37}$$

the probability that a region encompassing a given mass M overcomes the threshold is given by

$$P(M,z) = \int_{\delta_c}^{\infty} P(\delta)d\delta = \frac{1}{2} - \frac{1}{2}\mathrm{erf}\Big(\frac{\delta_c}{\sqrt{2}\sigma}\Big), \tag{1.38}$$

where erf is the error function. The fraction of mass locked up in halos is then obtained by differentiating Eq. 1.38 as a function of mass, so that the total mass per unit volume contained by collapsed halos is equal to

$$n(M)dM = \frac{\rho_M}{M}\frac{\partial F(>M)}{\partial M}dM, \tag{1.39}$$

where F(>M) = 2P(M,z). The factor of two accounts for small regions that do not overcome the δ_c threshold, but are contained in larger overdensities and end up collapsing into halos as well. The number density of dark matter halos as a function of their mass is finally obtained as:

$$n(M)dM = \sqrt{\frac{2}{\pi}}\frac{\rho_M}{M^2}\frac{\delta_c}{\sigma_M}\Big|\frac{d\ln\sigma_M}{d\ln M}\Big| \exp\Big[-\frac{\delta_c^2}{2\sigma_M^2}\Big]. \tag{1.40}$$

Eq. 1.40 provides a theoretical formulation of the halo mass function (HMF). For low mass halos hosting galaxies and groups, the mass function scales as a power law. For massive halos hosting clusters of galaxies ($\gtrapprox 10^{14}\mathrm{M}_\odot$) it decreases with an exponential cutoff, because such large peaks are more rare. The mass variance $\sigma(M,z)$ increases with cosmic time, giving perturbations the chance to collapse into halos in recent times. The result is that the number of structures decreases with redshift. This is in agreement with the bottom up scenario of structure formation, where smaller halos merge to form more massive objects.

Despite the various assumptions, this formalism provides a good description of the abundance of massive structures in large N-body simulations. A comparison to the Millennium simulation (Springel et al., 2005) is shown in Fig. 1.8. The Press-Schechter formalism has been extended by Bond et al. (1991), accounting for the fact that a region that does not exceed δ_c when smoothed on a scale M, could potentially do so when scaled on larger M. If the density field is smoothed with a top-hat function, the Fourier k-modes are independent and the density fluctuations behave as random walks across the δ_c barrier. This approach is known as *excursion set formalism* or *extended Press-Schechter formalism* (EPS). It also naturally produces the ad-hoc factor of 2 in Eq. 1.38, fixing the incorrect assumption that P(>M) = P(δ>δ_c|M) (Mo et al., 2010).

Further progress was made on the Press-Schechter formalism and its extension. Eq. 1.40 can be rewritten as:

$$\frac{dn}{dlnM} = \frac{\rho_M}{M}\Big|\frac{d\ln\sigma}{d\ln M}\Big| f(\sigma), \tag{1.41}$$

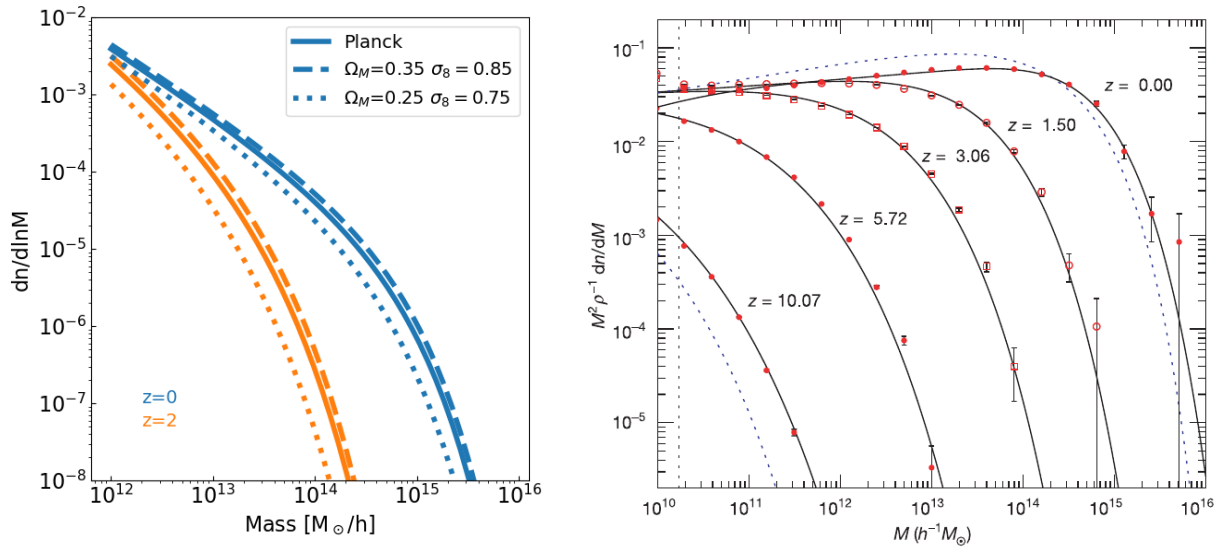

Abbildung 1.8: Halo number density as a function of mass and redshift. **Left-hand panel**: mass function at z=0 and z=2 for different cosmological parameters. The solid lines denote the Planck Collaboration et al. (2020a) cosmology, and the dashed (dotted) lines show larger (smaller) Ω_M and σ_8. The models shown in this panel were computed with the COLOSSUS software (Diemer, 2018). **Right-hand panel**: mass function in the Millennium simulation. The solid and dashed lines show prediction with the models from Jenkins et al. (2001) and Press and Schechter (1974). The panel is adapted from Springel et al. (2005).

where σ is the mass variance and $f(\sigma)$ is known as multiplicity function. Different models of the halo mass function are encoded in various functional forms of the multiplicity function, while the cosmological dependence is mainly hidden in ρ_M and σ. For example, the introduction of ellipsoidal collapse by Bond et al. (1991); Sheth and Tormen (1999, 2002) solved some inconsistencies between the results of N-body simulations and the analytical prediction of halo abundance. The ellipsoidal collapse corresponds to the introduction of a moving barrier δ_c within the EPS framework. Their multiplicity function reads:

$$\nu f_{\mathrm{ST}}(\nu) = A\left(1 + \frac{1}{\nu'^p}\right)\left(\frac{\nu'}{2}\right)^{1/2}\frac{\mathrm{e}^{-\nu'/2}}{\sqrt{\pi}}, \tag{1.42}$$

with $\nu' = 0.707\nu$ and p=0.3. The quantity

$$\nu(M, z) = \delta_c(z)/\sigma(M, z), \tag{1.43}$$

is the peak height, where δ_c is the critical overdensity required for a structure to collapse in a dark matter halo (Eq. 1.32). Nevertheless, the analytical description of the non-linear evolution struggles to account for the complex dynamics of halo formation. The prediction of halo abundance is therefore not perfectly precise. Therefore, we rely on models calibrated on large N-body simulations (e.g., Jenkins et al., 2001; Tinker et al., 2008; Bhattacharya

et al., 2011; Despali et al., 2016; Bocquet et al., 2016; Comparat et al., 2017; McClintock et al., 2019; Bocquet et al., 2020; Seppi et al., 2021). The abundance of dark matter halos is a powerful cosmological probe. It depends on the mass variance, which encodes the matter power spectrum, and the total amount of matter in the Universe (see Eq. 1.40). Because clusters trace massive halos, cosmological parameters are constrained by comparing the models to the number of galaxy clusters in observations.
In this context, robust modeling of the halo mass function is key. The majority of the models calibrated on simulations rely on the universality of the halo mass function: it depends on the statistic of peaks in a Gaussian field. Therefore, a set of simulations computed at a fixed cosmology is suitable to predict the halo mass function for different cosmological models (e.g, Jenkins et al., 2001; Bhattacharya et al., 2011; Angulo et al., 2012; Comparat et al., 2017; Seppi et al., 2021). However, the HMF deviates by 10–20% according to different numerical algorithms and mass definitions (Tinker et al., 2008; Despali et al., 2016; Diemer, 2020; Ondaro-Mallea et al., 2022). An innovative approach is the use of emulators trained on multiple sets of simulations with different cosmological parameters, where a model is built by interpolating the mass function measured in each box (McClintock et al., 2019; Nishimichi et al., 2019; Bocquet et al., 2020). We refer the reader to Angulo and Hahn (2022) for an extensive discussion. In addition, the evolution mapping approach from Sánchez et al. (2022) using σ_{12} has the potential to provide accurate redshift evolution of the mass function also at early times. Finally, most of these models are calibrated on N-body simulations that contain only dark matter. Significant progress has been achieved by recent works toward including baryonic effects on the halo mass function using hydrodynamical simulation (Bocquet et al., 2016; Debackere et al., 2021; Castro et al., 2021; Euclid Collaboration et al., 2022). More details about the exploitation of clusters in a cosmological context are provided in Sect. 1.2.2.

1.2 Clusters of galaxies

Clusters of galaxies are the most massive virialized structures in the Universe (see Allen et al., 2011; Kravtsov and Borgani, 2012, for reviews). They form hierarchically, from the merging of smaller halos, and reside in the nodes of the Large Scale Structure. Clusters host tens to hundreds of galaxies, their typical total mass spans from about 10^{13} to 10^{15} $M_\odot$, and their typical size from 1 to 5 Mpc. The largest fraction of their mass is in form of dark matter ($\sim$85%). Most of the baryonic mass is composed of the hot intracluster medium (ICM, $\sim$15%), while only a small percentage is in form of stars.

1.2.1 Observations of galaxy clusters

The first works related to clusters of galaxies date back to the 1700s, with the first observations of concentrations of galaxies. Charles Messier was mainly looking for new comets but ended up discovering other sources such as nebulae, star clusters, and galaxies. He published a final version of the *Messier catalog*, consisting of 110 objects (Messier, 1781). Slightly

later, F. Wilhelm Herschel extended such catalogs and published a collection of about 500 new objects (Herschel, 1802). He revealed that many nebulae observed by Messier were clusters of stars. Part of these sources corresponds to structures of the Local Group, such as Virgo and Coma (see Biviano, 2000, for a historical review). More recently, observations of the Coma cluster provided the first evidence of dark matter by the already mentioned work from Zwicky (1933).
The first galaxy cluster catalogs were published by Abell (1958); Abell et al. (1989), looking for concentrations of galaxies in the sky. Massive clusters can host hundreds of galaxy members. However, this method suffers from projection effects: many galaxies at different distances along the line of sight may appear as an overdensity of galaxies at the same position in the sky (Costanzi et al., 2019). The best way of dealing with projection effects is to use redshift measurements to disentangle member and field galaxies along the line of sight. Very large cluster catalogs have been compiled with this approach (e.g., Wen et al., 2012). Spectroscopic redshifts improve photometric measurements by about one order of magnitude (Clerc et al., 2020; Ider Chitham et al., 2020).

Galaxy clusters additionally exhibit observational features in other energy bands. Their hosting halos are filled with low particle density gas that is heated to extremely high temperatures ($\sim 10^{-3}$ cm^{-3}, 10^7 – 10^8 K, Sarazin (1986)). The ICM is a fully ionized plasma, that emits in X-rays due to thermal bremsstrahlung, as free electrons accelerate when approaching atomic heavier ions (free–free emission). X-ray observations therefore probe a different component of clusters, rather than the single galaxy members. A review of X-ray observations of clusters is provided by Rosati et al. (2002). Early large catalogs from X-rays were compiled using the ROSAT telescope (Truemper, 1982) by Voges et al. (1999); Böhringer et al. (2004). More recent works provided larger and more accurate cluster catalogs from XMM-Newton (Snowden et al., 2008; Adami et al., 2018) and, most recently eROSITA (Liu et al., 2022).
The emissivity due to thermal bremsstrahlung depends on the local density of free electrons and the cooling function Λ(T,Z):

$$\epsilon(r) \propto n_e(r) n_p(r) \Lambda(T, Z), \tag{1.44}$$

where T and Z are the temperature and metallicity of the ICM. For a fully ionized plasma $n_e \sim n_p$, so that the emissivity scales as the square of the local gas density. The shape of the bremsstrahlung spectrum shifts from a power law at low frequencies to an exponential decay at higher energies. The cutoff frequency $\nu_C = \frac{K_B T}{h_{Pl}}$ depends on the temperature of the ICM (h_{Pl} is the Planck constant). The measure of surface brightness provides a probe of the cluster density profile, which is usually modeled with a β profile (Cavaliere and Fusco-Femiano, 1976):

$$\begin{aligned} S(r) &= \frac{1}{4\pi(1+z)^4} \int \epsilon(r) dl \\ S(r) &= S_0 [1 + (r/r_c)^2]^{-3\beta+1/2}, \end{aligned} \tag{1.45}$$

where S(r) is the surface brightness (integrated along the line of sight), r_c is the core radius and β is related to the slope of the profile. A more generic description is given by the Vikhlinin et al. (2006) model:

$$n_e^2(r) = n_0^2 \frac{(r/r_c)^{-\alpha}}{(1+r^2/r_c^2)^{3\beta-\alpha/2}} \frac{1}{(1+r^\gamma/r_s^\gamma)^{\epsilon/\gamma}} + \frac{n_{02}^2}{(1+r^2/r_{c2}^2)^{2\beta_2}}, \tag{1.46}$$

which provides a better simultaneous description of the cluster core and outskirts, even with a fixed γ=3.
Because of the X-ray emission mechanism (Eqs. 1.44 and 1.45), the detection of galaxy clusters in the X-ray band depends on their dynamical state. Clusters hosted by relaxed dark matter halos had time to develop an efficient cooling in the central regions, which is manifested by a peaked emission in the core (cool core clusters). This effect may bias the X-ray detection toward relaxed objects, compared to more disturbed ones, where the gas did not cool down efficiently and they are therefore lacking the peak in the surface brightness (non-cool core clusters). This is the notion of cool core bias (Molendi and Pizzolato, 2001; Ettori and Brighenti, 2008; Eckert et al., 2011).

Another probe of the ICM is the inverse Compton scattering of CMB photons with the free electrons composing the cluster gas. This is known as the thermal Sunyaev–Zel'dovich effect (tSZ, Sunyaev and Zeldovich, 1972). The result is an average blue-shift of the CMB spectrum, observed in the millimeter band at a peak frequency of ∼217 GHz. The relative variation of the CMB temperature due to tSZ reads:

$$\begin{aligned} f(x) &= \left(x\frac{e^x+1}{e^x-1} - 4\right)(1+\delta_{\rm tSZ}(x,T)) \\ \frac{\Delta T_{\rm tSZ}}{T_{\rm CMB}} &= f(x)y = \int n_e \frac{K_B T_e}{m_e c^2} \sigma_T dl, \end{aligned} \tag{1.47}$$

where $\delta_{\rm tSZ}$ is a relativistic correction, x=$\dfrac{h_{Pl}\nu}{K_B T_{\rm CMB}}$, y is the Compton y-parameter, σ_T is the Thomson cross-section, and c is the speed of light. The tSZ signal scales linearly with the gas density. Because of its origin, the tSZ is independent of redshift and provides a great tool for detecting and studying clusters at early times (z$\gtrapprox$1).
An additional distortion of the CMB spectrum is caused by the bulk motion of clusters with respect to the CMB reference frame when the peculiar velocity has a component parallel to the line of sight. This is called kinetic Sunyaev–Zel'dovich effect (kSZ, Sunyaev et al., 2003) and is parameterized as:

$$\frac{\Delta T_{\rm kSZ}}{T_{\rm CMB}} = -\frac{v_{\rm pec}}{c}\tau_e, \tag{1.48}$$

where τ_e is the electron optical depth. The kSZ effect produces relative temperature changes of $\delta T_{\rm kSZ}/T \sim 10^{-5}$ and is negligible compared to tSZ relative differences of $\delta T_{\rm tSZ}/T \sim 10^{-4}$. For a review of the SZ effect, see Carlstrom et al. (2002).

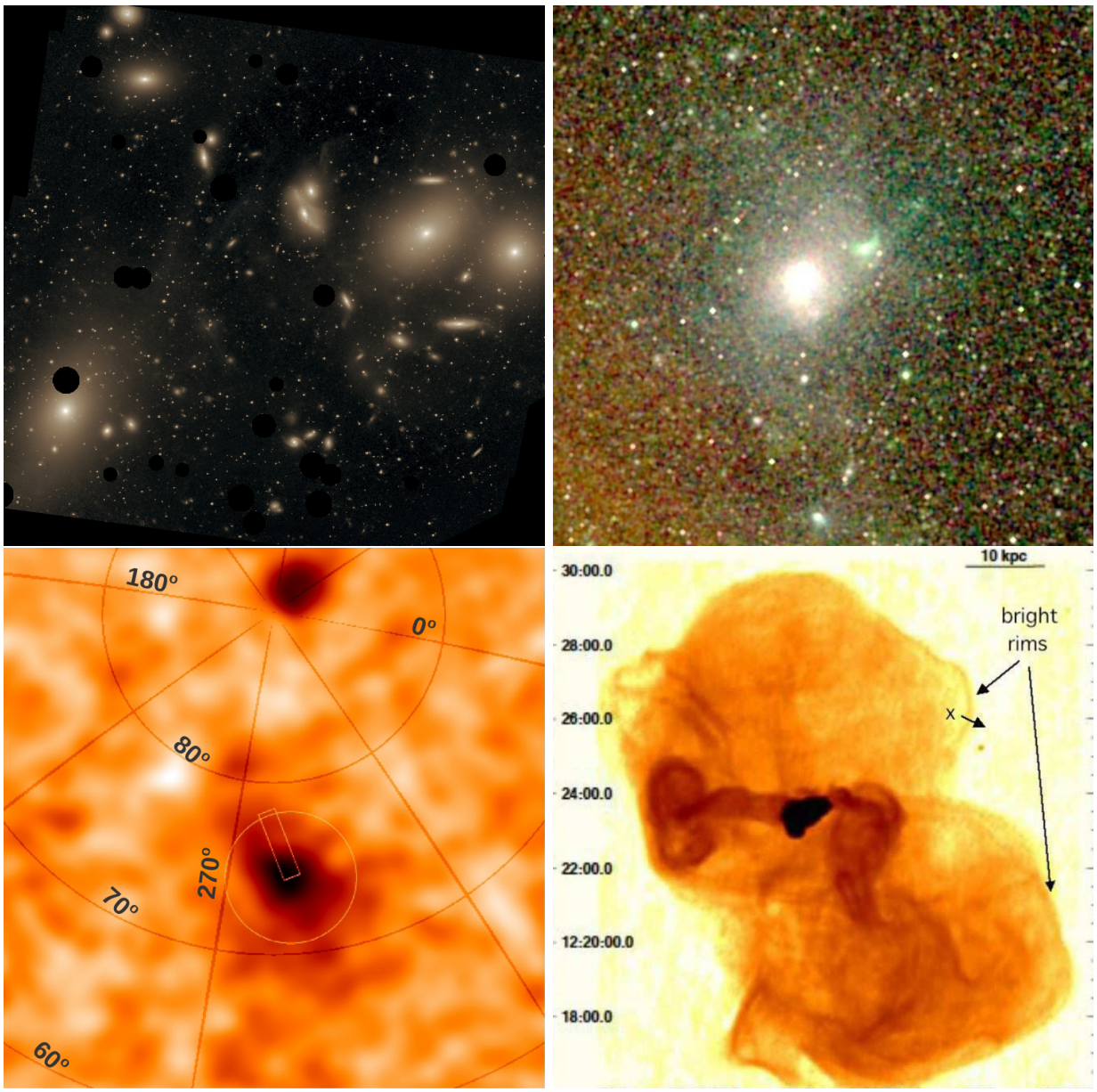

Abbildung 1.9: The Virgo galaxy cluster. The **Top-left panel**: Virgo in the optical band using data from the Burrell Schmidt telescope. Credit: Chris Mihos (Case Western Reserve University)/ESO. **Top-right panel**: X-ray data from eROSITA showing the emission from the hot ICM. The energy bands are 0.3–0.63 keV (red), 0.63–1.04 (green), 1.04–2.30 keV (blue). Credit: Jeremy Sanders. **Bottom-left panel**: Virgo as observed by Planck through the SZ effect (Planck Collaboration et al., 2016d). **Bottom-right panel**: the Very Large Array (VLA) view of the Virgo cluster in the radio band Mathews and Guo (2011).

Finally, clusters of galaxies are also observed in the radio band. The presence of energetic electrons and large scale magnetic fields created by mergers generate diffuse synchrotron radiation (Lazarian et al., 2012; Brunetti and Lazarian, 2016). Radio data is therefore probing non thermal cluster emission and highlight features related to dynamically unstable structures such as relics and shocks (van Weeren et al., 2010; Bonafede et al., 2010; Brüggen et al., 2012).
Figure 1.9 shows the Virgo cluster observed in different wavelengths. Virgo is the nearest cluster of galaxies, located about 16 Mpc away from us. Its estimated mass is $\sim 1.2\times10^{15}\mathrm{M}_{\odot}$ and it hosts ∼1500 galaxies. It is the nucleus of the Virgo supercluster, which hosts the Milky Way in its outskirts. The four panels show the Virgo cluster observed in the optical, X-ray, millimeter, and radio bands.

1.2.2 Clusters as cosmological probes

Various methods provide cosmological constraints using clusters of galaxies. We focus first on the cluster mass estimate and their number density as a function of mass and redshift. We describe alternative methods in a second stage.

Cluster counts

An observable quantity that is tightly correlated with mass is essential to use clusters of galaxies as cosmological probes by counting their abundance as a function of mass and redshift (see Eq. 1.40). Going back to the first catalog from Abell (1958) and the idea that light traces mass in the Universe, a quantity of interest is the total number of galaxy members. This concept is nowadays encoded in richness, which is often measured as the sum of the membership probability for each galaxy member (Rykoff et al., 2014):

$$\lambda = \sum_{i=0}^{N} P_{\mathrm{mem,i}}, \tag{1.49}$$

where the index i runs over the N members. Richness is used as a cluster mass proxy (Tinker et al., 2012; Ider Chitham et al., 2020). However, the measure of λ is often noisy, which reduced the cosmological constraining power (Saro et al., 2015; Capasso et al., 2019; Chiu et al., 2020).

There are many other ways of estimating the cluster mass. For example, assuming that the X-ray emitting ICM is in hydrodynamical equilibrium within a symmetrically spherical gravitation potential, one gets:

$$M(r) = -\frac{K_B T}{\mu m_p}\frac{r}{G}\left(\frac{d\ln\rho}{d\ln r} + \frac{d\ln T}{d\ln r}\right), \tag{1.50}$$

where K_B is the Boltzmann constant and m_p is the proton mass. However, massive clusters are rarely spherical objects in perfect equilibrium. Therefore, this method provides mass estimates that can be biased (Smith et al., 2016; Hurier and Angulo, 2018; Gianfagna et al., 2021). Cluster masses using X-ray data are also estimated using scaling relations between

mass and X-ray observables, such as the X-ray luminosity and Y_X, the product between the cluster gas mass and temperature (Maughan, 2007).

A similar approach is possible with the SZ effect: the total SZ signal integrated on the surface area covered by a cluster is proportional to the mass and temperature of the ICM:

$$Y_{\rm tSZ} = \int y d\Omega \propto M_{\rm ICM} T, \tag{1.51}$$

so that the $Y_{\rm tSZ}$ signal provides a direct probe of the gas mass and indirect information about the total cluster mass.

Another dynamical method to evaluate the cluster mass is to measure the member galaxies' velocity dispersion (Mamon and Łokas, 2005; Biviano et al., 2006; Oldham and Auger, 2016). The typical values in clusters are $\sigma_v \sim 10^3$ km/s, corresponding to a crossing time of $t_c \sim \dfrac{R_{\rm clu}}{\sigma_v} \sim 10^9$ yr: clusters are assumed to be virialized structures and it is possible to estimate their mass by applying the virial theorem. Assuming that the galaxies travel through a stable, spherical cluster potential, the mass–velocity dispersion relation reads:

$$M_{\rm vir} = \frac{3 R_{\rm clu}}{G} \sigma_v^2, \tag{1.52}$$

which provides a direct link to halo mass (Wojtak and Łokas, 2010). However, this method is also based on spherical symmetry and relaxation, which break down for disturbed clusters and recent mergers, and projection effects.

In principle, the most precise and least biased cluster mass estimate is provided by weak gravitational lensing (Bartelmann and Schneider, 2001; Hoekstra et al., 2013; Kilbinger, 2015; Umetsu, 2020; Giocoli et al., 2021). The gravitational field of a cluster shears and magnifies the images of distant background galaxies perpendicularly to the lens–image direction. Statistical treatment of the ellipticity of the background galaxy images is related to the tangential shear signal γ_T. The latter is proportional to the cluster gravitational potential projected along the line of sight and is, therefore, an excellent probe of the total mass of the cluster. For a review of gravitational lensing see Bartelmann (2010).
Since weak lensing data are usually partially available for large cluster surveys, a method is to calibrate a scaling relation between an observable such as X-ray luminosity or optical richness and weak lensing mass for a subsample of clusters (see e.g., Mantz et al., 2016; Pacaud et al., 2018; Bocquet et al., 2019; Ider Chitham et al., 2020).

Within a cluster cosmological experiment, a key ingredient is the survey volume, that is the volume within which each cluster is contained. This is fundamental for the comparison of the total number of clusters in a sample with the prediction from a model of the cluster volume number density. The comoving volume element (Hogg, 1999) is defined as:

$$dV(z, \Omega) = D_H \frac{D_M^2}{E(z)} d\Omega dz, \tag{1.53}$$

where D_M is the transverse comoving distance (Eq. 1.11) and Ω is the area covered by the survey.

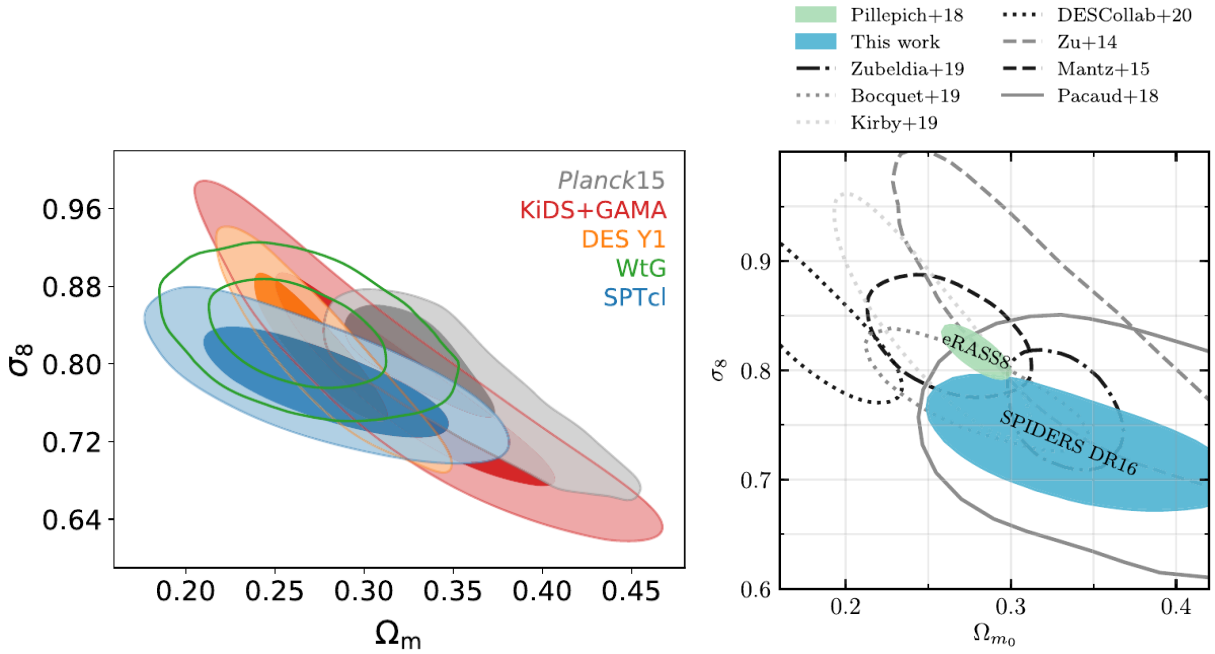

Abbildung 1.10: Cosmological constraints on Ω_M–σ_8 from different cluster count experiments compared to other probes. **Left-hand panel**: results from Bocquet et al. (2019), compared to CMB constraints (Planck Collaboration et al., 2014a), galaxy clustering and lensing (van Uitert et al., 2018; Abbott et al., 2020), and the Weighing the Giants experiment (WtG, Mantz et al., 2015b). **Right-hand panel**: results from Ider Chitham et al. (2020), compared to other cluster count experiments and a prediction using clusters detected by eROSITA from Pillepich et al. (2018b).

Finally, it is crucial to account for the fact that not all clusters in the Universe are detectable. Our telescopes and instruments can not pick up the emission from very faint and distant objects due to the limited sensitivity. The selection function accounts for this, by providing a probability of detection for the clusters involved in the experiment. It is possible to model the selection function empirically (Bocquet et al., 2019), or starting from reliable mock data (Clerc et al., 2018; Liu et al., 2021; Seppi et al., 2022).

All these ingredients are finally combined in the total predicted number of clusters as a function of observable and redshift:

$$\frac{dN(O,z)}{dOdz} = \int f(O,z)\frac{dP(O|M,z)}{dO}\frac{dn(M,z)}{dM}\frac{dV}{dz}dM, \tag{1.54}$$

where O is an observable (e.g., X-ray luminosity), $f(O,z)$ is the selection function, $\frac{dP(O|M,z)}{dO}$ is the mass–observable scaling relation, $\frac{dn(M,z)}{dM}$ is the model of the halo mass function, and $\frac{dV}{dz}$ is the comoving volume element. This method provides powerful constraints on cosmological parameters, mainly on the matter density parameter Ω_M and

the amplitude of the matter power spectrum σ_8, but also on the dark energy equation of state, which is parameterized as:

$$w(z) = w_0 + \frac{z}{1+z} w_a. \tag{1.55}$$

In the last two decades, cosmological experiments with clusters of galaxies reached a ∼10% and 5% precision on Ω_M and σ_8 (Vikhlinin et al., 2009b; Mantz et al., 2015b; Planck Collaboration et al., 2016c; Schellenberger and Reiprich, 2017b; Pacaud et al., 2018; Bocquet et al., 2019; Zubeldia and Challinor, 2019; Ider Chitham et al., 2020; Abbott et al., 2020; Costanzi et al., 2021; Garrel et al., 2022; Lesci et al., 2022a; Salvati et al., 2022). An example of the cosmological constraints of these two parameters is shown in Fig. 1.10. For detailed reviews see Borgani (2008); Allen et al. (2011); Kravtsov and Borgani (2012); Weinberg et al. (2013); Pratt et al. (2019); Clerc and Finoguenov (2022).

Alternative methods

Additional ways of constraining cosmological models with clusters have been explored. Their large scale distribution is a biased tracer of dark matter. The measure of the cluster two point correlation function or power spectrum is a sensitive probe of cosmology. Since clusters are highly non-linear structures, precise modeling of their bias relative to the dark matter distribution is essential (Kaiser, 1987; Sheth and Tormen, 2002; Tinker et al., 2010; Comparat et al., 2017). Such studies additionally provide constraints on the growth rate (Eq. 1.18) throughout cosmic time (Moscardini et al., 2001; Valageas and Clerc, 2012; Veropalumbo et al., 2014; Marulli et al., 2017, 2018, 2021; Lindholm et al., 2021; Moresco et al., 2021; Lesci et al., 2022b).

Combined SZ and X-ray observations of the same clusters allow using them as standard candles and measuring the Hubble constant H_0. In particular, the different dependence on gas density in Eqs. 1.44 and 1.47 allows measuring the angular diameter distance as:

$$D_A \propto \frac{\Delta T^2_{\rm CMB} \Lambda(T, Z)}{S(r) T^2}, \tag{1.56}$$

and constrain H_0 from Eqs. 1.11 and 1.12 (Carlstrom et al., 2002).

Clusters of galaxies produce deep gravitational potential wells that retain the gas ejected by Active Galactic Nuclei (AGN) and supernovae. Therefore, the chemical composition of the ICM and the ratio between the baryons and dark matter densities are expected to be similar to the cosmic ones. The baryon fraction is defined as:

$$f_b = \frac{M_{\rm gas}}{M_{\rm tot}} \approx \frac{M_b}{M_{\rm tot}} = \frac{\Omega_B}{\Omega_M}. \tag{1.57}$$

The baryon fraction is measured by combining X-ray data constraining the gas density profile and the gas mass, with weak lensing data or scaling relations to measure the total mass (Gonzalez et al., 2013; Laganá et al., 2013). It approaches the cosmic value on large scales in massive clusters (see Eckert et al., 2021, for a review). External information of Ω_B

from the CMB or the abundances of primordial elements provide a direct measure of Ω_M by inverting Eq. 1.57 (White et al., 1993; Allen et al., 2004; Ettori et al., 2009; Borgani and Kravtsov, 2011; Mantz et al., 2022).

An alternative method of using clusters as a cosmological probe is sparsity. Sparsity is the ratio of two different overdensity masses:

$$s_{1,2} = \frac{M_{\Delta_1}}{M_{\Delta_2}}, \tag{1.58}$$

where M_{Δ_1} and M_{Δ_2} are the masses corresponding to the overdensities Δ_1 and Δ_2, with $\Delta_1 > \Delta_2$. Sparsity encodes the mass distribution of dark matter halos, which is the result of hierarchical growth and depends on the amplitude of the matter power spectrum. Balmès et al. (2013) showed that sparsity is nearly independent of halo mass and it directly depends on the halo mass function at Δ_1 and Δ_2:

$$\int_{M_{\Delta_2}^{MIN}}^{M_{\Delta_2}^{MAX}} \frac{dn}{dM_{\Delta_2}} dM_{\Delta_2} = s_{1,2} \int_{M_{\Delta_1}^{MIN}}^{M_{\Delta_1}^{MAX}} \frac{dn}{dM_{\Delta_1}} dM_{\Delta_1}. \tag{1.59}$$

This technique does not significantly depend on selection and mass calibration effects (Corasaniti et al., 2018, 2021).

Finally, clusters of galaxies are a useful tool for testing the cosmological principle. Since clusters generate from initial peaks in the density fields, they collapse in a self-similar way (Kaiser, 1986). Under this assumption, clusters are scaled versions of the same object, depending on their virial mass and the evolution of the critical density. The self-similar model predicts power law relations between different cluster properties, such as X-ray luminosity and total mass. Although it provides a decent approximation to observed properties, the self-similar model fails due to baryonic effects, which make the scaling relations more complex than a simple power law. These scaling relations are a key property of galaxy clusters (Mantz et al., 2016; Bulbul et al., 2019; Bahar et al., 2022; Chiu et al., 2022). Their directional behavior and the dependence of their normalization on orientation are probes of cosmic isotropy, especially if the scaling relation is tight, for example, the one between X-ray luminosity and temperature (Migkas et al., 2020, 2021).

1.3 The extended ROentgen Survey with an Imaging Telescope Array (eROSITA)

The atmosphere of the Earth absorbs all photons in the X-ray band. Therefore, X-ray observatories need to be moved to outer space. The first X-ray source to be observed was the Sun in the late 1940s using V-2 rockets (Seward and Charles, 1995). Pioneering work from Giacconi et al. (1962) using an Aerobee rocket provided evidence for X-ray emission from the Moon and the discovery of the X-ray binary Scorpius X-1, the first extra-solar X-ray source ever detected. Since then, more than 20 X-ray satellites have been launched, starting with *Uhuru* (Giacconi et al., 1972), and including the *Einstein X-ray Observatory*

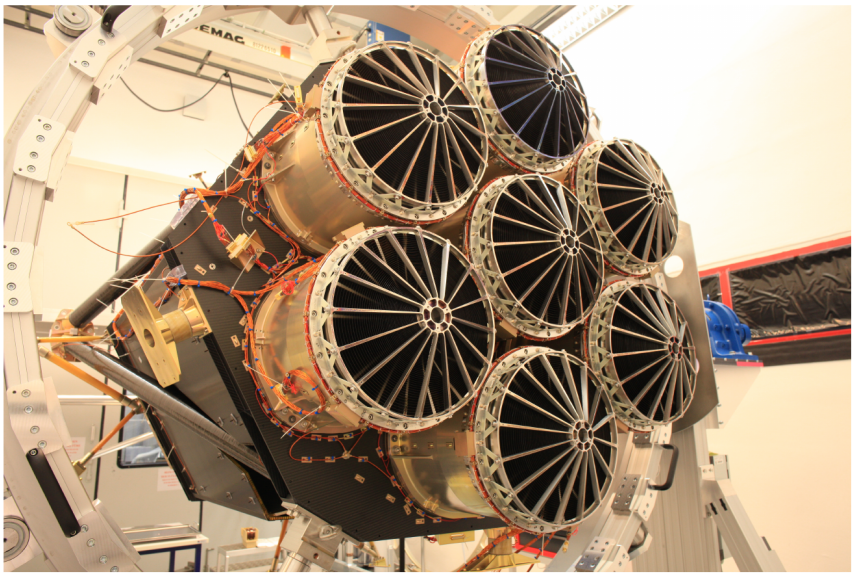

Abbildung 1.11: A picture of the seven eROSITA telescope modules. Figure taken from Predehl et al. (2021).

(Giacconi et al., 1979), *EXOSAT* (de Korte et al., 1981), *BeppoSAX* (Boella et al., 1997), *Chandra* (Weisskopf et al., 2000), *XMM-Newton* (Jansen et al., 2001), *Suzaku* (Mitsuda et al., 2007), and *Hitomi* (Takahashi et al., 2016). In particular, the *ROentgen SATellite* (ROSAT, Truemper, 1982) performed the first X-ray imaging all-sky survey, providing the first all-sky catalog of bright X-ray sources (RASS, Voges et al., 1999).

1.3.1 eROSITA onboard SRG

The *extended ROentgen Survey with an Imaging Telescope Array* (eROSITA, Merloni et al., 2012; Predehl et al., 2021) is the natural successor of ROSAT. It was designed and developed at the Max Planck Institute for Extraterrestrial Physics (MPE), in Garching, Germany. It is the soft X-ray instrument onboard Spectrum-Roentgen-Gamma (SRG). SRG-eROSITA was launched on the 13th of July 2019 from Baikonur, Kazakhstan, by means of a Proton-M rocket. It is now located at the Lagrangian point L2, where it follows an elliptical orbit with a semi-major axis of about 1 million km and a period of six months. eROSITA is designed as a sensitive, wide-field telescope, able to produce deep images of large areas of the sky. The goal of the mission is to complete eight all-sky surveys (eRASS1–eRASS8). One all-sky scan lasts for six months. The eROSITA mission is a collaboration between the German and the Russian consortia. The full sky is split in half between them.

eROSITA is composed by seven telescope modules with 54 nested Wolter-I mirror shells each (Wolter, 1952). They are shown in Fig 1.11. Each mirror assembly is coupled to a charge-coupled-device (CCD). The eROSITA point spread function (PSF) has a half energy width (HEW) of about 15″ on axis for each module. The spacecraft rotates with an angular velocity of 0.025 deg/s, so that one revolution around the ecliptic poles takes four hours.

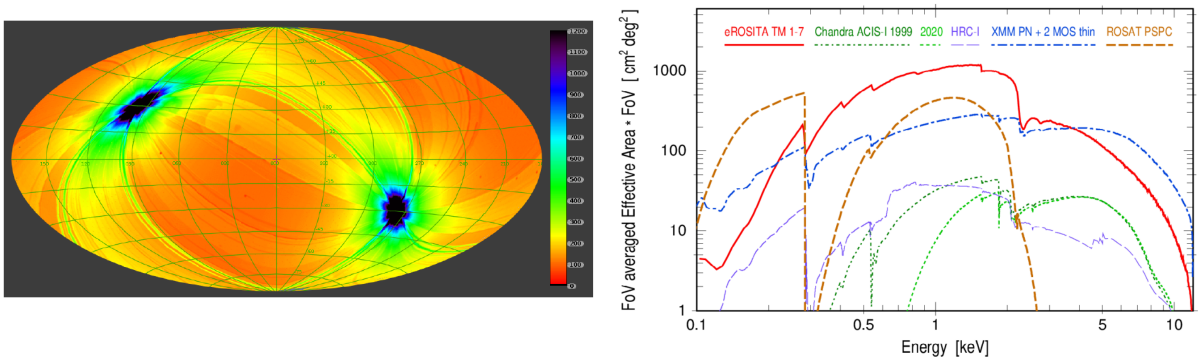

Abbildung 1.12: The scanning and collecting capabilities of eROSITA. **Left-hand panel**: vignetted exposure map of the first all-sky survey in galactic coordinates (eRASS1). The depth is larger at the ecliptic poles, covered by the spacecraft on each revolution. **Right-hand panel**: product of the eROSITA field of view and its effective area as a function of energy (grasp) in red, compared to Chandra ACIS-I (in green and purple), XMM-Newton (in blue), and ROSAT (in brown).

Consequently, the exposure at the ecliptic poles is larger compared to equatorial regions. The exposure map of eRASS1 is shown in the left-hand panel of Fig. 1.12. eROSITA has a larger field of view (FOV, 65′) compared to XMM-Newton (∼30′) and Chandra (∼17′). In addition, the quick scanning capabilities make the eROSITA grasp larger than its predecessors in the soft band (0.2–2.3 keV). The grasp is the product of the field of view and the average effective area and is a direct measure of the ability of a telescope to collect photons. It is displayed in the right-hand panel of Fig. 1.12. The final cumulative survey (eRASS:8) will be about 25 times more sensitive than RASS in the soft X-ray band (0.2–2.3 keV). eROSITA is the first instrument to provide an all-sky map in the hard band (2.3–10 keV).

The design of eROSITA is particularly suitable for performing a survey of galaxy clusters, thanks to a combination of its all-sky scanning capability, the small PSF compared to ROSAT (about 30 arcsec in survey mode average over the FOV), the larger FOV compared to Chandra and XMM-Newton (about 1 square degree), and its large effective area in the soft X-ray band (about 1300 cm^2 at 1 keV). The eROSITA telescope was designed to detect about 100 000 clusters, reaching an unprecedented sample size for the X-ray view of these objects and their astrophysical and cosmological purposes. eROSITA is also expected to detect about 3 million AGN and 20 000 normal galaxies, enabling detailed studies of the evolution of supermassive black holes throughout cosmic time and galaxy formation. The scanning strategy of the spacecraft additionally provides multiple observations of the same sky area at different times, a key ingredient for time domain studies, such as tidal disruption events (Malyali et al., 2021) and variability phenomena (Arcodia et al., 2021; König et al., 2022). A major result was the discovery of X-ray bubbles around the Milky Way by Predehl et al. (2020).

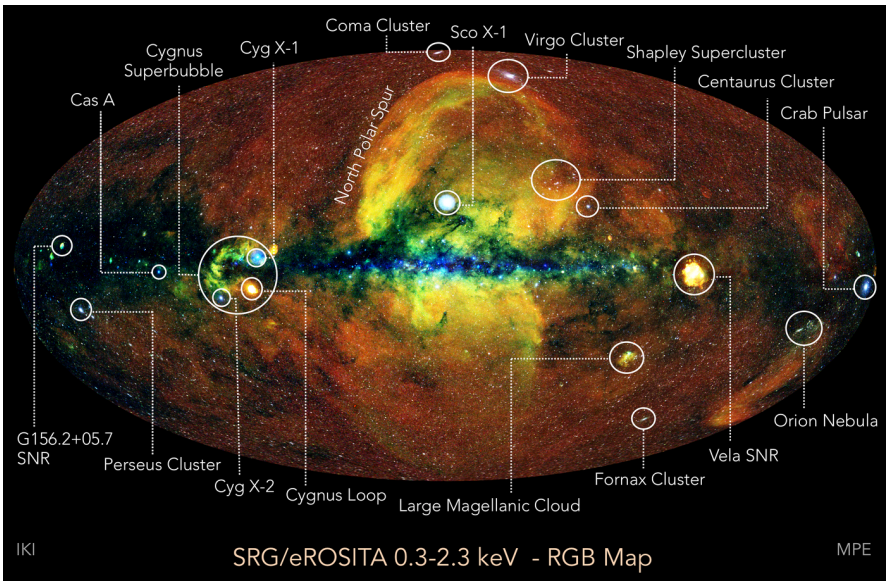

Abbildung 1.13: Map of the first eROSITA all-sky survey (eRASS1) in Aitoff projection. The figure is color coded according to photon energy (red: 0.3–0.6 keV, green: 0.6–1 keV, blue: 1–2.3 keV). Some of the brightest X-ray sources are annotated.
Credit: Jeremy Sanders, Hermann Brunner, Andrea Merloni and the eSASS team (MPE); Eugene Churazov, Marat Gilfanov (on behalf of IKI).

During the Calibration and Performance Verification Phase (CalPV), the eROSITA Final Equatorial Depth Survey (eFEDS, Brunner et al., 2022) has been carried out. eFEDS was designed to verify the survey capabilities of eROSITA. This mini-survey covers an area of $\sim$140 deg^2 in the equatorial region ($126° < \mathrm{RA} < 146°$, $-3° < \mathrm{DEC} < +6°$). It was covered with a vignetted (unvignetted) exposure time of $\sim$1.2 ks ($\sim$2.2 ks), a similar value compared to the final all-sky survey (eRASS:8) in the equatorial region. The first eROSITA all-sky survey (eRASS1) was completed in June 2020 [1]. It is shown in Fig. 1.13, where some galaxy clusters, such as Virgo, Centaurus, and Fornax, are clearly visible. Dust and gas along the galactic plane absorb the soft X-ray emission so that only the hard emission emerges. The most prominent sources are bright supernova remnants, such as Vela, or X-ray binaries, like Sco X-1.

1.3.2 Cosmological experiment

The main goal of eROSITA is to study the Large Scale Structure of the Universe, to constrain cosmological models and gain knowledge about dark matter and dark energy. The main probe is the abundance of galaxy clusters as a function of mass and redshift (see Eqs. 1.41 and 1.54). A forecast on the eROSITA cosmological capabilities was elaborated

[1] https://www.mpe.mpg.de/7461950/erass1-presskit

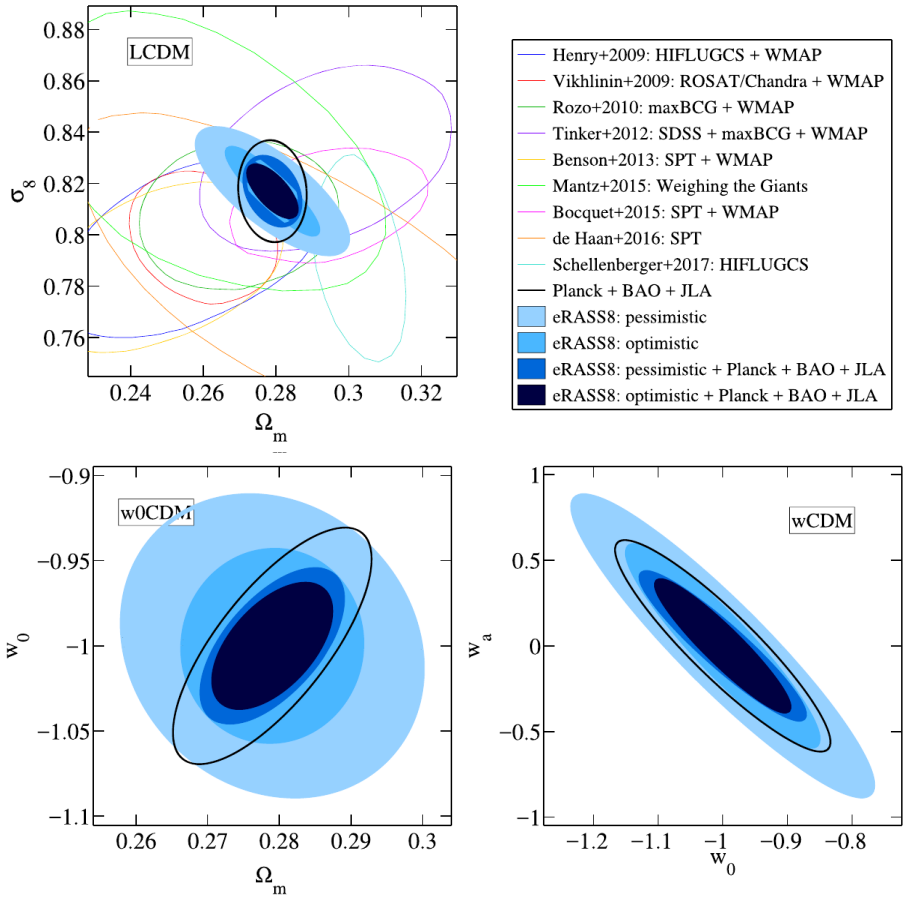

Abbildung 1.14: Forecast of the eRASS:8 cosmological experiment, combining cluster counts and angular clustering, for pessimistic and optimistic scenarios in light-blue shaded areas. The inclusion of Planck+BAO+JLA results (black line) is shown by the dark-blue shaded areas. The contours refer to 68% confidence level. **Top panels**: constraints of the ΛCDM model for Ω_M and σ_8, compared to previous cluster cosmological experiments. **Bottom panels**: constraints on model with varying dark energy equation of state. Figure taken from Pillepich et al. (2018b).

by Pillepich et al. (2018b), who showed that the Figure of Merit defined as

$$FoM_\Lambda = \frac{1}{\sqrt{\mathrm{detCov}(w_0, w_a)}}, \tag{1.60}$$

is expected to be larger than 100, which classifies eROSITA as a StageIV Dark Energy probe (Albrecht et al., 2006).

More predictions are shown in Fig. 1.14. The top panels show that eROSITA will be more precise than previous cluster count experiments when measuring Ω_M and σ_8. Contours with different shades of blue assume different levels of data quality. The pessimistic prediction in light blue considers only photometric redshifts, the priors on the mass–observable

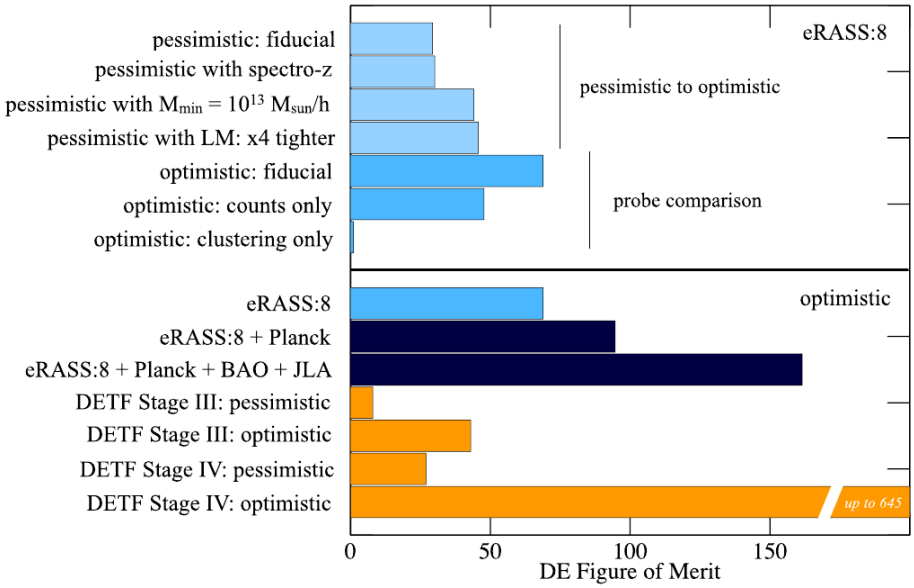

Abbildung 1.15: Prediction of the Dark Energy Figure of Merit (Eq. 1.60). The top panel shows different survey assumptions, the bottom panel shows a comparison to Dark Energy Task Force Requirements (Albrecht et al., 2006). The panel is adapted from (Pillepich et al., 2018b).

scaling relation match the errors from Vikhlinin et al. (2009a), and the inclusion of all clusters with more than 50 photon counts and more massive than $5{\times}10^{13}$ $M_{\odot}/h$. The optimistic one assumes instead spectroscopic redshifts, scaling relation priors that are four times better than Vikhlinin et al. (2009a), and the inclusion of groups down to $1{\times}10^{13}$ $M_{\odot}/h$. When combined with other probes, such as Planck, Joint Light Curve Analysis (JLA, Betoule et al., 2014), and BAO, eROSITA has the potential to provide constraints within about 2% and 1% for Ω_M and σ_8, respectively. The two bottom panels display constraints on the dark energy equation of state (see Eq. 1.55), with final uncertainties around 7% for w_0 and ± 0.25 for w_a.
A final prediction for the FoM obtained by assuming different survey assumptions are shown in the top panel if Fig. 1.15, adapted from Pillepich et al. (2018b). The lower panel shows a comparison to the recommendations from Albrecht et al. (2006), where an experiment is classified as Stage IV if it constrains the FoM between 27 and 645. The eROSITA prediction is included within this range.

1.4 Thesis overview

As highlighted in the previous sections, eROSITA has great potential for cosmological studies and will provide the best constraints using an X-ray-selected cluster sample ever. However, to reach the highest levels of precision, it is crucial to account for selection effects, uncertainties in the data, and properly characterize the samples of sources detected

by eROSITA. The eROSITA cosmological results will be biased if faint undetected clusters are not modeled properly, if the cluster sample is contaminated by random fluctuations of the X-ray background or bright AGN that appear as extended sources in the sky. These aspects impact the abundance of objects as a function of mass and redshift and will ultimately cause a wrong measure of the cosmological parameters if they are not adequately accounted for. Producing and analyzing mock data through simulations that reproduce observations is a great tool for testing all these effects.
The first goal of this thesis is to generate a digital twin of eRASS1, with realistic models of clusters and AGN, to study how different populations of sources are detected and characterized by eROSITA. This is fundamental to build a complete and pure galaxy cluster sample that is suitable for cosmology. This work is explained in Kap. 2, it was peer-reviewed and published in Astronomy and Astrophysics (Seppi et al., 2022).

Until the last decade, most of the uncertainty in cluster count experiments was due to a large scatter of the mass–observable scaling relation. In this era of precision cosmology, new surveys are providing very large samples of massive clusters, which statistically improves the constraints on the mass–observable relation. However, the better accuracy exposes uncertainties in the models of the halo mass function, which was a negligible source of error for previous surveys (Salvati et al., 2020). Accurate and precise models of the HMF are required to achieve percent level accuracy in the measure of cosmological parameters.
The second goal of this thesis is calibrating an accurate model of the halo mass function that additionally accounts for the dynamical state of dark matter halos. This work is explained in Kap. 3, it was peer-reviewed and published in Astronomy and Astrophysics (Seppi et al., 2021). The mass function model from Seppi et al. (2021) offers the possibility to marginalize on selection effects related to the dynamical state directly in the measure of the cluster abundance. It has the potential of reducing uncertainties in a cluster count experiment and producing unbiased cosmological constraints. However, a link between theoretical models and observations is lacking.

The third goal of this thesis is to study the dynamical state of clusters detected by eROSITA and build a connection between N-body simulations that are used to develop HMF models and observations. We use hydrodynamical simulations to compare variables related to the dynamical state of dark matter halos in N-body simulations to the offset between different definitions of the center for galaxy clusters observed by eROSITA. This work is explained in Kap. 4, it has been submitted to Astronomy and Astrophysics.

Finally, a summary of the main results and conclusions is described in Kap. 5.

Kapitel 2

Detecting clusters of galaxies and active galactic nuclei in an eROSITA all-sky survey digital twin

Our knowledge of the large-scale structure (LSS) of the Universe has dramatically improved in the past decades thanks to a variety of surveys at different wavelengths. A wealth of information about the matter distribution on cosmological scales is obtained by optical data from galaxy clustering, measured by the Two-degree-Field Galaxy Redshift Survey (2dFGRS, Colless et al., 2001), the Galaxy and Mass Assembly (GAMA) Survey (Driver et al., 2009), the VIMOS Public Extragalactic Redshift Survey (VIPERS, de la Torre et al., 2013), the Dark Energy Survey (DES, Abbott et al., 2020), the Kilo-Degree Survey (KiDS, Joudaki et al., 2018), the Hyper Suprime-Cam Subaru Strategic Program (HSC-SSP, Hikage et al., 2019), and the Sloan Digital Sky Survey (SDSS, Alam et al., 2021). Complementary data in the millimeter range trace the large-scale distribution of matter thanks to the lensing of the cosmic microwave background (CMB, Sherwin et al., 2012; Planck Collaboration et al., 2014c). In addition, large samples of extragalactic sources are provided by X-ray surveys, such as ROSAT (Boller et al., 2016) and eROSITA (Merloni et al., 2012; Predehl et al., 2021). It is important to consider both galaxy clusters and active galactic nuclei (AGN) in this context: they both trace the LSS. They are fundamental to shedding light on the hot and energetic large-scale structure of the Universe.

Clusters of galaxies populate the most massive bound dark matter halos in the Universe. They are the largest known virialized structures (Kravtsov and Borgani, 2012; Pratt et al., 2019) and are a great tool for cosmological studies (Tinker et al., 2008; Allen et al., 2011; Lesci et al., 2022a; Clerc and Finoguenov, 2022). Galaxy clusters are observed in optical data as an over-density of red galaxies (e.g., Rykoff et al., 2014; Abbott et al., 2020) or as peaks in weak-lensing convergence maps (e.g., Miyazaki et al., 2018a), by distortion of the CMB due to the Sunyaev–Zel'dovich (SZ) effect in the millimeter band (e.g., Staniszewski et al., 2009; Planck Collaboration et al., 2016a) and by extended emission in the X-ray band (e.g., Böhringer et al., 2004; Adami et al., 2018; Finoguenov et al., 2020; Liu et al., 2022). The combination of multiwavelength data is key for a complete description of galaxy

clusters. On the one hand, optical surveys have the highest source density, which provides the largest samples of clusters using photometric data (Oguri, 2014; Bleem et al., 2015a). On the other hand, pointed observations with interferometers in the radio and millimeter bands provide observations with extremely high angular resolution (Pasini et al., 2022). In addition, SZ surveys with telescopes such as *Planck* (Planck Collaboration et al., 2014a), the South Pole Telescope (SPT, Bleem et al., 2015b), or the Atacama Cosmology Telescope (ACT, Hilton et al., 2021) are effective in detecting high-redshift objects, thanks to the redshift-independent SZ signal. X-ray observations are particularly suitable to study clusters of galaxies. Clusters are the brightest extragalactic extended sources in the X-ray band (Rosati et al., 2002), they emit mainly due to thermal bremsstrahlung from the hot intra-cluster medium (Cavaliere and Fusco-Femiano, 1976) and their emissivity depends on the radial density profile.

Active galactic nuclei (AGN) are very luminous objects, powered by the accretion of rich gas reservoirs onto super-massive black holes, and constitute the majority of the extragalactic sources detected in X-ray surveys (see Padovani et al., 2017, for a review). A large sample of AGN enables studies of the general evolution of supermassive black holes (Kauffmann and Haehnelt, 2000), the properties of the host galaxy (Ferrarese and Merritt, 2000), the AGN clustering properties (Koutoulidis et al., 2013; Viitanen et al., 2019), and their link to the underlying dark matter large-scale structure (Fanidakis et al., 2011; Georgakakis et al., 2019), as well as different channels through which these objects are formed (Mayer and Bonoli, 2019), and the mechanisms triggering bursts of X-ray radiation (Arcodia et al., 2021).

With eROSITA, a new era in X-ray astronomy is now unfolding (Merloni et al., 2012; Predehl et al., 2021). Since December 2019, eROSITA is performing all-sky surveys. The sky is split in half between the German (eROSITA_DE) and Russian consortium (eROSITA_RU). The eROSITA_DE area is split into 2 447 tiles with a small overlap for data processing purposes. Of these, 2 248 are uniquely owned by the German consortium, and the additional 199 are shared. Each tile covers a unique area of ∼8.7 square degrees.
eROSITA is predicted to ultimately detect a total of about 10^5 clusters of galaxies after the final cumulative all-sky survey (eRASS:8), the largest sample of X-ray-selected galaxy clusters to date. This will allow a variety of studies involving the cluster X-ray luminosity function (Mullis et al., 2004; Koens et al., 2013; Finoguenov et al., 2015; Adami et al., 2018; Clerc et al., 2020; Liu et al., 2022), the clustering of galaxy clusters (Veropalumbo et al., 2014; Marulli et al., 2018, 2021; Lindholm et al., 2021), and provide powerful constraints on cosmological parameters such as the normalization of the power spectrum σ_8 and the matter content of the Universe Ω_M (Borgani, 2008; Vikhlinin et al., 2009b; Mantz et al., 2015b; Pierre et al., 2016; Schellenberger and Reiprich, 2017b; Pacaud et al., 2018; Ider Chitham et al., 2020; Garrel et al., 2022). A prediction of the eROSITA cluster count cosmology capabilities is studied by Pillepich et al. (2012, 2018b). A total number of about three million sources, most of which are AGN, are expected to be detected in eRASS:8, a factor of 20 better than ROSAT.

An efficient and accurate detection of extragalactic sources is key to properly sampling the cosmic web and making the most out of the large samples provided by eROSITA.

The identification of galaxy clusters in X-ray surveys like eROSITA is affected by Poisson count noise in the low photon count regime and by the redshift-dimming effect on the cluster surface brightness. Cluster samples selected from X-ray surveys are primarily flux-limited (e.g., REFLEX, Böhringer et al., 2004). The detection of clusters also depends on secondary effects, such as their extent on the sky, or the low surface brightness of very extended objects (Pacaud et al., 2006; Burenin et al., 2007; Finoguenov et al., 2020). In this context, the cool core bias and the dynamical state of galaxy clusters have also been studied in recent years (Hudson et al., 2010; Eckert et al., 2011; Rossetti et al., 2016; Andrade-Santos et al., 2017; Käfer et al., 2019; Ghirardini et al., 2021a). Relaxed clusters develop an efficient cooling toward their center, which enhances the X-ray emission in the inner region. Such peaked surface brightness profiles possibly bias the detection toward relaxed structures. This has an impact on cosmological studies using the halo mass function (Seppi et al., 2021).

The cross-correlation between clusters and AGN in the LSS creates an interplay between point and extended sources in the detection process. A detailed understanding of the point sources is fundamental to investigate not only the X-ray background and the completeness of the observed sample (Georgakakis et al., 2008), but also the fraction of clusters that are misclassified as a point source (Pacaud et al., 2006; Burenin et al., 2007). This happens because of the small size of high redshift clusters, the peaked emission from compact nearby groups, or the presence of a central AGN in the cluster, which can boost the detection of high redshift clusters (McDonald et al., 2012; Trudeau et al., 2020). This misclassification is mitigated by multiwavelength follow-up observations. For instance, Salvato et al. (2022) found 346 cluster candidates in the eFEDS point-source catalog by the identification of the red sequence using optical data. An extensive study of these objects is provided by Bulbul et al. (2022). An effective way of investigating the detection and selection effects in surveys is to simulate the observational process in its greatest detail. This approach has been explored using mocks in different wavelengths, from the optical band (Jimeno et al., 2017; Oguri et al., 2018), to the X-rays (Liu et al., 2013; Pierre et al., 2016; Clerc et al., 2018), and the microwave sky (Sehgal et al., 2010), or injecting simulated sources into real images (Suchyta et al., 2016; Everett et al., 2022). It allows accounting for instrumental effects and the observing strategy. Studying and quantifying effects that have an impact on the detection is then possible, comparing catalogs of simulated sources and the population that is detected in the simulation. Constant improvements in computational power and efficiency provide more detailed mocks. Recent progress in dark matter simulations allows to minimize the impact of cosmic variance thanks to the ability to simulate large volumes, but also resolve galaxy-like halos because of the small resolution (e.g., Klypin et al., 2016; Chuang et al., 2019; Ishiyama et al., 2021).

We study the eROSITA capabilities in the detection of extragalactic sources following this approach. Our goal is to understand the details of AGN and cluster detection and selection effects. These are two important subsequent steps. First, the detection should be optimized to maximize the ability to identify clusters and AGN, and make sure that the algorithm in question is detecting as many real sources as possible. After that, one can focus on selection criteria to clean the catalog of detected sources and obtain a certain sample according to

the scientific goal.
In this chapter, we use realistic end-to-end simulations to predict the population of objects observed by eROSITA, with a particular interest in extended sources, that are clusters of galaxies, and AGN. We focus on the eROSITA_DE sky area. We start from the simulations described by Comparat et al. (2019, 2020). We generate a half-sky simulation at the depth of the first eROSITA all-sky survey (eRASS1), the one reached after six months of operations. We follow the eROSITA scanning strategy. Photons are generated for 2438 eROSITA_DE tiles. The background is directly resampled from the eRASS1 observations. We extend the cluster model from Comparat et al. (2020) to galaxy groups down to $2{\times}10^{13}$ $M_{\odot}$ using the relation between X-ray luminosity and stellar mass (Anderson et al., 2015). Comparat et al. (2022) showed that such correction allows matching the relation between projected luminosity around eFEDS central galaxies and their stellar mass remarkably well. We run the eSASS (extended Science Analysis Software System) detection algorithm described by Brunner et al. (2022). We build a one-to-one association between simulated objects and the source catalog using the source ID of each simulated photon (Liu et al., 2021), properly linked to a cluster, AGN, star, or the background. We assess the performance of the detection in terms of completeness (fraction of simulated objects that are recovered in the source catalog) and purity (fraction of entries in the source catalogs that are assigned to the correct simulated object). Our study follows up on the work of Liu et al. (2021) on the eFEDS simulations. We take one step further, accounting for the larger variations of exposure and background level in eRASS1.
This chapter is organized as follows. We summarize the main features of the simulation and the X-ray model in Sect. 2.1. We describe the detection process, the handling of the catalogs, and the classification of the sources in Sect. 2.2. We provide our results in Sect. 2.3. We study the population in the source catalog, the cumulative number density of AGN and clusters as a function of flux, the completeness of these samples, their relation with purity and contamination, and measure the X-ray luminosity of clusters. We further discuss our results in Sect. 2.4, including the best strategy to build samples of clusters detected by eROSITA, accounting for the different exposure across the sky. Finally, we summarize our findings in Sect. 2.10. The results in this Chapter are presented in Seppi et al. (2022).

2.1 Simulated data

We follow the approach described by Comparat et al. (2020) and create all-sky simulations. A dark matter light cone is built with snapshots at different redshifts. Cluster and AGN models are used to predict X-ray emission (Comparat et al., 2019, 2020). We upgrade the cluster model to the galaxy groups regime. In this section, we review the main features of the simulations and models that are relevant for this analysis. The simulated data has been released with the article from Seppi et al. (2022).

2.1.1 Light cones from N-body dark matter simulations

A light cone is created with the UNIT1i N-body simulations (Chuang et al., 2019). These are computed in a Flat ΛCDM cosmology (Planck Collaboration et al., 2016c). The fiducial parameters are $H_0 = 67.74$ km s^{-1} Mpc^{-1}, $\Omega_{m0} = 0.308900$, $\Omega_{b0} = 0.048206$. The size of the simulation box is 1 Gpc/h[1] and the mass resolution is $1.2{\times}10^9$ $M_\odot$/h. It allows a detailed modeling of both clusters and AGN. It is suited for studying low mass structures down to 10^{11} $M_\odot$, AGN up to z∼6, and the eROSITA selection function (Liu et al., 2021).

2.1.2 X-ray model components

These simulations combine different source and X-ray background components. We describe each one of them in the following section.

Galaxy clusters

Comparat et al. (2020) introduce a new method to simulate the X-ray emission from galaxy clusters. The principle is to build mock observations using real data as a starting point (e.g., Kong et al., 2020; Everett et al., 2022). A total sample of 326 clusters is obtained by combining XMM-XXL (Pierre et al., 2016), HIFLUGCS (Reiprich and Böhringer, 2002), X-COP (Eckert et al., 2019b) and SPT-Chandra (Sanders et al., 2018). Their combination constitutes a relatively fair benchmark for eROSITA observations. Their X-ray properties are well measured inside R_{500c}, the radius encompassing an average density that is 500 times the critical density of the Universe at the redshift of the cluster $\rho_c = 3H^2/8\pi G$, where H is the Hubble parameter and G is the universal gravitational constant. From these clusters, a covariance matrix between redshift, temperature, hydrostatic masses, and emissivity is constructed. Simulated emissivity profiles are drawn from the covariance matrix by a Gaussian random process. These profiles are assigned to dark matter halos by a nearest neighbor process, considering mass and redshift. The brightness of the cluster core is linked to the dynamical state of the dark matter halo. The initial model is constructed using clusters with high counts and signal-to-noise ratio, making it reliable down to masses of $M_{500c} \sim 5{\times}10^{13}$ $M_\odot$.

In this chapter, we extend this model to galaxy groups for the eRASS1 simulation as follows. We use the relation between stellar mass and X-ray luminosity from Anderson et al. (2015) as a reference. The stellar mass is assigned to halos by an abundance matching scheme (see Comparat et al., 2019, and Section 2.1.2). We infer an average correction as a function of mass to align the scaling relation of the simulation to that of Anderson et al. (2015). The goal scaling relation between X-ray luminosity and the stellar mass of the central galaxy in each halo reads

$$\log_{10} L_{x,(0.5-2.0\mathrm{keV})} = 3 \log_{10} M_* + 7.8. \tag{2.1}$$

[1] h is the dimensionless Hubble constant, equal to the value of $H_0/100$.

This average correction bends the scaling relation predicted by Comparat et al. (2020) at low mass to predict lower luminosities for lower mass halos. Importantly, it preserves the scatter in the L_X–mass scaling relation. These values substitute the ones obtained by integrating the emissivity profiles from the original covariance matrix. For halos with a mass larger than $M_{500c} > 10^{14}$ $M_\odot$ the correction is negligible, but it becomes very important in the mass range $10^{13} - 5 \times 10^{13}$ $M_\odot$. In Sect. 2.5, both panels in Fig. 2.15 highlight the improvement of the model after applying the correction. The number density of sources as a function of X-ray flux (logN–logS) predicted for masses above $\log_{10} M/M_\odot > 13$ is in excellent agreement with observations (Finoguenov et al., 2007, 2015, 2020; Liu et al., 2022; Chiu et al., 2022; Bahar et al., 2022). With the eFEDS sample, the method is further validated. It offers a more complete picture of the cluster population. The relation between X-ray luminosity and M_{500c} in the second panel of Fig. 2.15 shows the impact of the correction, especially for groups. The predicted values of $\log_{10} L_x$ reach reasonable values of ~ 41 (and below) at $\log_{10} M/M_\odot \sim 13$. The improved model is in line with different sets of observations, considering that these are flux-limited samples, whereas the orange curve is built with the complete simulated clusters population (Lovisari et al., 2015; Schellenberger and Reiprich, 2017a; Bulbul et al., 2019; Lovisari et al., 2020; Chiu et al., 2022; Bahar et al., 2022). In general, our correction provides an excellent agreement between the new model and eFEDS clusters sample. We provide further details in Sect. 2.5. In total, we simulate 1 116 758 clusters.

Active galactic nuclei

Active galactic nuclei are simulated by an empirical model that reliably reproduces their number density as a function of X-ray luminosity, clustering, and redshift (Georgakakis et al., 2019; Comparat et al., 2019). It is based on stellar mass to halo mass relations (Moster et al., 2013) and abundance matching to reproduce the hard X-ray AGN luminosity function (Aird et al., 2015; Buchner et al., 2015) and their number density as a function of flux up to z = 6. It matches the observed AGN duty cycle (fraction of galaxies hosting an active nucleus) by construction (Georgakakis et al., 2019). The model extends to very low X-ray fluxes $\sim 1 \times 10^{-17}$ erg/s/cm^2, well under the eROSITA flux limit, which enable a prediction of the X-ray background due to faint AGN. For the construction of the AGN population in the eRASS1 simulation, the sky is first divided into 768 HEALPix[2] fields, which ensures faster processing, but also a smaller volume, sampling the luminosity function down to about 10^{-7} sources per Mpc^3. This prevents the simulation of extremely bright sources. The model of the AGN spectra is an absorbed power-law with Compton reflection and a soft scattered component by cold matter (in Xspec `tbabs*(plcabs+pexrav)+zpowerlw)*tbabs`). The spectral index of the power-law is equal to $\Gamma = 1.9$. Finally, a fine-grained K-correction is applied to the AGN population (Hogg et al., 2002). The simulation accounts for a cross-correlation between clusters and AGN since they are both generated from the same N-body simulation. We neglect secondary effects regarding the population of halos hosting AGN in

[2] https://healpix.sourceforge.io/

cluster environments. Further observational studies involving the fraction of active galaxies in clusters as a function of redshift and a comparison to field galaxies are required to develop such a model, (see Martini et al., 2013; Koulouridis et al., 2014; Noordeh et al., 2020). In total, we simulate 225 583 320 AGN, about 200 times more than the clusters. Among them, 93 311 810 produce at least one count within 60″ from the center.

Stars

Fluxes to be assigned to stars are drawn from the eFEDS logN–logS. We assign them to GAIA DR2 (Gaia Collaboration et al., 2018) true positions randomly. The spectrum is a 0.8 keV APEC model at redshift 0. This model is simple, but nonetheless sufficient to mimic the increase of stellar density toward the Milky Way for this simulation at the eRASS1 depth (Schneider et al., 2022; Salvato et al., 2022). In total, we simulate 373 316 stars.

Background

Our approach is similar to the one detailed by Liu et al. (2021), who decompose and re-simulate the eFEDS background, subtracting the contribution from the simulated faint AGN, that partially contribute to the cosmic X-ray background (CXB). However, this is not feasible in eRASS1, due to the nonuniform coverage of the sky and background emission. We update such a method for the eRASS1 simulation. Background photons are obtained by resampling the observed eROSITA background maps, masking identified point and extended sources. This allows the introduction of spatially varying background, that closely follows real data. We start from the eROSITA_DE eRASS1 event lists and source catalogs. Following the masking scheme devised by Comparat et al. (in prep.), the photons are split into two groups. First, we consider source photons: events located within 1.4 times the source radius of a detected source (see Sect. 2.2 for a definition of the source radius). Secondly, we select background photons: events located further than 1.4 times the source radius of any detected source. These thresholds guarantee conservative masking of the sources in the event list to obtain a background event list. The complementary set of events constitutes the source event list. The whole dataset is mirrored in the eROSITA_RU sky, to obtain an all-sky map. This is divided into 49 125 HEALPix regions, each of them covering ~ 0.84 deg^2. The X-ray spectrum and the images of the background events are extracted from these regions. All the spectra are merged into a single mean background spectrum. These inputs are combined to generate a specific SIMPUT[3] file for the mock background, that provides by construction a faithful reproduction of the observed eRASS1 background.

[3] https://www.sternwarte.uni-erlangen.de/sixte/sources, v-2.4.7

2.1.3 Mock observation

Photons are simulated with the SIXTE[4] software (Dauser et al., 2019), a dedicated end-to-end X-ray simulator. SIXTE is the official simulator for eROSITA. The result is a list of events with energy, position, and arrival time. This approach allows accounting for instrumental effects because the simulator relies on vignetting, energy-dependent PSF, ancillary response file (ARF), and redistribution matrix file (RMF) as input from calibration data. The setup follows the eROSITA all-sky scan strategy (Merloni et al., 2012; Predehl et al., 2021).
We use the same attitude file from the real observations for the eRASS1 simulation. The attitude file specifies the details of the scanning by the spacecraft. It follows the planned observing strategy, scanning one full great circle every four hours. In addition, we use the same good time intervals (gti) of the real survey. This allows us to account for details such as orbit corrections, when the cameras are switched off, or camera failures, making the simulation an ideal digital twin of the real eRASS1. The total number of events in the simulation, covering about 20 618 square degrees, with energy of 0.2–10 keV is 187 486 754. There are 118 905 555 photons in the soft band (0.2–2.3 keV). These numbers are indeed very similar to the real data, respectively equal to 194 350 024 and 118 815 616 counts. The ratios between these numbers are 0.965 and 1.001 respectively.

2.2 Data analysis method

In this section, we describe how the simulated event files are processed and analyzed. The final result is a catalog of sources identified by the detection algorithm. We refer to the latter as the source catalog in the rest of this work. Only event files in the eROSITA_DE sky are processed. We first generate the photons on the sky plane divided into 768 HEALPix regions and then create specific catalogs for each field. This way we do not simulate the same photons twice in the overlapping regions of different eROSITA tiles. Given our interest in cluster detection, we focus on a single band detection in the soft X-rays (0.2–2.3 keV), where the eROSITA effective area is the highest (Predehl et al., 2021).

2.2.1 eSASS detection

Each simulated tile is processed with the eROSITA Standard Analysis Software System (eSASS, version eSASSusers_201009) (Brunner et al., 2022). Starting from the calibrated event file, we produce $3.6° \times 3.6°$ images for the eRASS1 simulation and the corresponding exposure maps, using all 7 telescope modules, in the soft X-ray band 0.2–2.3 keV. The detection relies on a sliding box algorithm, that looks for overdensities of photons over the background map. It follows the subsequent steps.

1. `erbox`: the image is scanned by a sliding cell, which marks potential sources if the signal-to-noise ratio is higher than a given threshold. This initial list of potential sources

[4] https://www.sternwarte.uni-erlangen.de/research/sixte, v-2.6.0

Tabelle 2.1: Number of counts by sources detected with given values of detection and extension likelihood

	Clusters			**AGN**		
DET_LIKE	N events $<0.5\times R_{500c}$			N events $<30''$		
	ALL	CLU	BG	ALL	AGN	BG
5	21.3	8.7	10.1	5.1	3.7	1.3
8	25.8	11.4	11.4	6.6	5.0	1.5
10	30.4	13.8	13.0	7.7	6.0	1.6
15	42.6	21.3	16.9	10.1	8.4	1.7
20	48.8	25.4	17.9	12.5	10.6	1.7
25	74.4	36.6	23.6	14.6	12.7	1.8
50	100.8	62.3	29.8	25.2	22.6	2.4
75	152.4	96.3	42.5	35.6	32.3	3.0
100	209.0	127.2	61.5	46.5	42.1	3.7

	Clusters		
EXT_LIKE	N events $< 0.5\times R_{500c}$		
	ALL	CLU	BG
6	69.4	31.4	31.2
8	78.5	42.7	27.2
10	89.7	46.9	33.2
15	103.9	59.1	34.0
20	139.9	67.5	32.5
25	144.7	90.9	42.3
50	275.9	168.6	77.8
75	405.9	284.2	95.6
100	530.1	376.5	119.1

Notes. The first column in the upper (lower) table reports the value of detection (extent) likelihood measured on a source. For clusters, the other columns show the total number of counts generated by all sources (ALL, includes photons from clusters, AGN, stars, and the background) inside half R_{500c}, the ones only generated by clusters (CLU) and the ones produced by the background (BG). For AGN, we report the total number of events within $30''$, the ones generated by AGN and by the background.

contains a large number of false detection, but maximizes the completeness.
2. `erbackmap`: the potential sources are masked by constructing a detection mask and the image is interpolated to create an adaptively smoothed background image. This process is iterated three times, to converge toward a more robust background map (Brunner et al., 2022; Liu et al., 2021).
3. `ermldet`: each box marked as a potential source is analyzed by a maximum likelihood PSF-fitting algorithm, based on the position, count rate, and extent of the source. It compares the distribution of counts to a β model (Cavaliere and Fusco-Femiano, 1976) convolved with the eROSITA PSF. It allows a simultaneous fitting of multiple sources. Different choices of the minimum likelihood threshold control the purity of the sample, decreasing the false detection rate when increasing the threshold. This task produces a catalog of sources and a source map.
Sources are assigned a significance of the detection (detection likelihood), extension of the best fitting β model (extent) and significance of the extended model over the point-like one (extension likelihood). These parameters are computed by minimizing the C-statistic (Cash, 1979) in Eq.2.2:

$$C = 2\sum_{i=1}^{N}(e_i - n_i \ln e_i), \tag{2.2}$$

where n_i is the measured number of events in each pixel and e_i is the expected value from the model. The significance of each source is computed by comparing the best fitting model to the zero count case $\Delta C = C_{\rm null} - C_{\rm fit}$ (see Brunner et al., 2022, Sect. A.5). The probability that a source arises from a random background fluctuation is computed using the regularized incomplete Gamma function P_Γ.

$$P = 1 - P_\Gamma(\frac{\nu}{2}, \frac{\Delta C}{2}), \tag{2.3}$$

where ν is the number of degrees of freedom in the model. This is equal to three (four) for point (extended) sources, corresponding to positions on the pixels X and Y, count rate (and core radius of the β model) for our study, which only uses one detection band. The likelihood for each source is finally related to the natural logarithm of such probability:

$$\mathcal{L}_{\rm det} = -\ln P. \tag{2.4}$$

This gives a set of two fundamental parameters for each detection: DET_LIKE ($\mathcal{L}_{\rm DET}$), and EXT_LIKE ($\mathcal{L}_{\rm EXT}$). The first (second) one is related to the probability of identifying a spurious point (extended) source, exponentially proportional to −DET_LIKE (−EXT_LIKE). The core radius of the best-fitting extended beta model is also provided. It is set to zero for point sources, its minimum and maximum values are $8''$ and $60''$. A constant $\beta = 2/3$ is assumed for the model so that the slope of the profile is equal to -3 (see Eq. 1.45). We show that on average our model generates profiles that are compatible with this assumption in Sect. 2.5. The minimum thresholds of DET_LIKE and EXT_LIKE are extremely important in this step. They have a significant impact on the completeness and purity of the source catalog, see Sect. 2.3.4. We follow the same task processing as the

eFEDS data, choosing values of detlikemin = 5 and extlikemin = 6 (Brunner et al., 2022). The values of detection and extension likelihood are correlated to the number of events from a given source and from the local background by construction. AGN producing five counts on average are detected with DET_LIKE = 10. Clusters of galaxies require a larger amount of events to be detected. A value of DET_LIKE = 5 is measured for clusters with nine source counts and ten background counts inside half R_{500c}. Classifying the clusters as extended sources requires a larger number of events. A value of EXT_LIKE = 6 is measured for clusters with about 30 counts inside half R_{500c}. When the ratio between source and background photons increases, the detection and extension likelihood rise as well. A value of EXT_LIKE (DET_LIKE) of 25 is measured on average for clusters with 91 (37) counts against 42 (24) events generated by the background. We provide a summary in Table 2.1. It shows the average number of counts generated by all sources, including clusters, agn, stars, and background, and the ones only generated by clusters (AGN) and background in the top left (right) panels at fixed values of detection likelihood. The bottom panel displays the counts at given extension likelihood value.
4. `apetool`: we perform source aperture photometry and compute the sensitivity map for each simulated tile. This gives the minimum number of counts necessary to detect a point-like source as a function of position in the sky, and at a given Poisson false detection probability threshold.
5. `srctool`: we measure the radius that maximizes the signal-to-noise ratio for each source. We refer to this parameter as source radius (srcRAD).
6. `ersenmap`: we compute the sensitivity map for extended sources. This gives the minimum flux necessary for a source to be detected at a given DET_LIKE threshold.
7. `apetool`: we perform again source aperture photometry focusing on the extended sources and different apertures of 60, 90, 120, 150, 180, 240, 300, and 600″.

We perform the source detection in the soft (0.2–2.3 keV) X-ray band. In principle, one could choose specific detection and extension likelihood threshold according to different needs. We choose to characterize the extended sources without additional selections, using detlikemin = 5 and extlikemin = 6. This keeps our cluster catalog reasonably complete (down to some flux limit), without rejecting faint sources that are potentially interesting. Figure 2.1 shows an example of this whole process. It displays a wedge of the simulated light cone in the top panel, showing galaxies that trace the large-scale structure in grey and how this is populated by AGN in blue, and clusters and groups in red. The bottom panel shows the projection on the sky plane of the events emitted by the sources in the wedge. It displays simulated photons in the soft X-ray band (black dots), the simulated stars (green circles), AGN (blue circles), clusters (red circles), extended detections (magenta squares), and point-like detections (cyan squares). This tile gives a typical view of different possible cases. Red circles within a magenta square identify simulated clusters that are detected as extended, whereas red circles within a cyan square denote clusters detected as point sources. Similarly, input AGN and stars detected as point sources are shown by blue and green circles within cyan squares. Every circle (red, blue, or green) without a corresponding square denotes a simulated object that has not been detected. We show clusters and AGN respectively down to low flux limits of $3{\times}10^{-14}$ erg/s/cm^2 and $8{\times}10^{-15}$ erg/s/cm^2.

This explains the undetected objects in Fig. 2.1. Finally, background fluctuations that are detected as spurious sources are identified by squares without any circle.

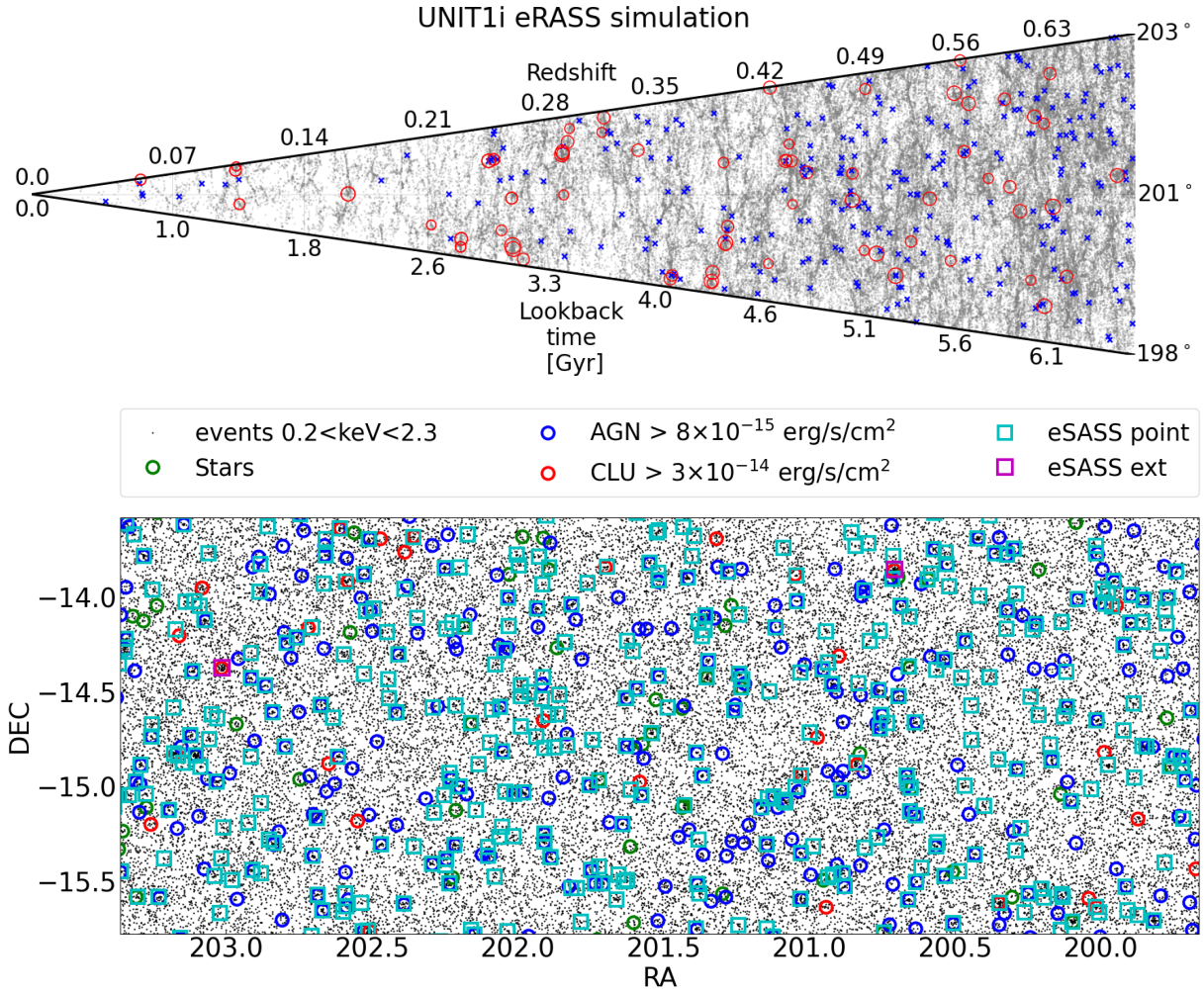

Abbildung 2.1: Large scale distribution of extragalactic sources and their X-ray view in the simulation. **Top panel**: light cone of the UNIT1i-eRASS1 simulation. The wedge shows the fraction of the sky enclosed by the same RA and DEC of the bottom panel as a function of redshift and lookback time. The galaxies tracing the large-scale structure are shown in grey. The AGN are denoted in blue. The red circles show clusters and groups. The size of the circle is proportional to the mass of the object. **Bottom panel**: central regions of tile 202105 of the eRASS1 simulation. This is the projection on the plane of the sky of light cone shown in the top panel. Photons with energies between 0.2 to 2.3 keV are shown by black dots, simulated stars by green circles, simulated AGN by blue circles, simulated clusters by red circles, eSASS extended detections by magenta squares, and eSASS point-like detections by cyan squares.

The X-ray background drives the detection process, especially for faint sources. We

compare the background maps computed on the simulation and on the real eRASS1 data. We find that the simulated background is overestimated by $\sim$ 10% compared to the observations. This is expected, because the cosmic X-ray background due to faint AGN is present both in the real eRASS1 background maps used to generate the background model, and as the simulated population of low-flux AGN.
We evaluate the impact of this 10% over-estimate of the background on the measured values of detection likelihood. We consider a wide range of counts per pixel values generated by a source (between 0.04 and 0.4) and by the background (between 0.001 and 0.009). These intervals are compatible with the source maps and background maps produced by eSASS. We expand these counts on a grid of 5×5 pixels, covering an area slightly larger than the eROSITA PSF. We compute the analytical value of detection likelihood by plugging these values into equations 2.2, 2.3, and 2.4. We repeat the process by increasing the background by 10%, computing the new value of $\mathcal{L}_{\rm det}$, and comparing it to the initial result with the unbiased background. We find that an overestimation of the background biases the detection likelihood to lower values. We measure a $\sim$ 4% negative impact on the calculation of detection likelihood for faint sources with DET_LIKE $\sim$ 5 and a 2.5% negative impact on more clear sources with DET_LIKE $\sim$ 20 due to a 10% overestimation of the background. We conclude that these effects have a minimal impact on the detection and characterization of faint sources around the detection limit, and do not significantly affect the study of more secure detections and the overall analysis of the population in the catalog. We provide further details and figures in Sect. 2.6.

2.2.2 Catalog description

We summarize the simulations and source catalog statistics in Table 2.2. The catalogs described above have been further cleaned because of the following reasons. The generation of event files was not completed correctly because of numerical issues in 6 HEALPix fields in the simulation, covering about 320 square degrees. These have not been considered in the analysis presented in the rest of this work. In addition, an area of about 260 square degrees around the southern ecliptic pole (RA$\sim$93°, DEC$\sim$-66°, where the exposure is maximal due to the survey scan mode) has been masked in the eRASS1 simulation. The generation of cluster events was not successful.
We focus on the extragalactic sky, masking the areas with galactic latitude $|g_{\rm lat}| < 10$ deg. The final area taken into consideration corresponds to 17 703.4 square degrees for the eRASS1 simulations.

Following the example of Liu et al. (2021), we merge simulated catalogs and source catalogs according to the integer identifier (ID) of each photon. Every simulated count has an ID that links it to the source that produced it. This method is more reliable than simply matching the catalogs (input and output) with coordinates, because it uses the origin of each simulated photon: a cluster, AGN, star, or the background. We summarize the algorithm in the following paragraph.
First of all, we assess whether a detected source has a simulated counterpart or not. For point (extended) sources detected by eSASS, we study the photons within aperture radii

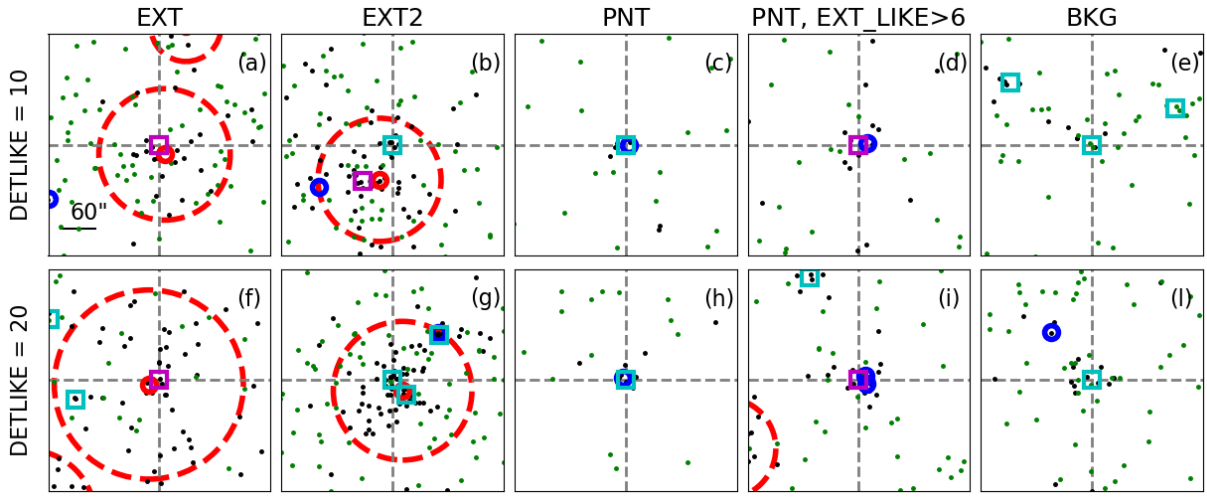

Abbildung 2.2: Examples of the eSASS catalog classification. Red (blue) solid circles show simulated clusters (AGN). Magenta (cyan) squares denote extended (point-like) eSASS entries, like in Fig. 2.1. The dashed red circles enclose $0.5\times R_{500c}$ of a simulated cluster. Soft X-ray photons from simulated sources are represented by black dots, the green ones come from the background. The first (second) row shows examples for sources with DET_LIKE = 10 (20). Columns show respectively: an extended detection uniquely assigned to a simulated cluster, a secondary detection assigned to an input cluster, a point detection uniquely assigned to an AGN, an extended detection uniquely assigned to an AGN, and a detection without any simulated input. All panels have the same physical size. A ruler of 60 arcseconds is shown in the top-left one.

of 20″ (60″). Their origin is stored in each photon ID. The entry in the source catalog is associated with the simulated source that issued the largest number of photons in the aperture radius. This assigns the ID of the simulated counterpart to the entry in the source catalog. We call this ID_Any. One caveat is that the simulation contains a large number of objects fainter than the eROSITA detection limit. Therefore, we only consider input sources that have at least two photons emitted during the mock observation. In addition, we set a lower counts threshold related to the local background counts, given by the counts corresponding to the 0.8 percentile point of the Poisson distribution, whose mean is equal to the number of background photons inside the given aperture radius.
Secondly, if an additional simulated counterpart is found, the one emitting the highest number of photons is assigned to ID_Any. The secondary counterpart is saved as ID_Any2. Finally, a simulated source can be split into multiple detected sources. This results in copies of the same ID_Any. We select the detection where the simulated object provides the highest photons count and consider a unique matching between the two (ID_Uniq). If ID_Any does not refer to a unique counterpart, in cases where there are multiple entries in the source catalog pointing to the same ID_Any, we use ID_Any2 if it is available. A one-to-one matching between the simulated objects and the source catalog can be obtained with ID_Uniq. We divide the source catalog into five classes using the IDs just assigned, following the example of Liu et al. (2021).

1. Primary counterpart of a simulated point source (PNT): detected source assigned to an ID_Uniq of an AGN or star. This is a secure point source detection.

2. Primary counterpart of a simulated extended source (EXT): detected source assigned to an ID_Uniq of a cluster. This is a secure cluster detection.

3. Secondary counterpart of a simulated point source (PNT2): detected source without an ID_Uniq, but assigned to an ID_Any of an AGN or star. This is a detection that corresponds to a fraction of a simulated point source but is not its primary counterpart. We refer to these as split sources corresponding to an AGN or star.

4. Secondary counterpart of a simulated extended source (EXT2): detected source without an ID_Uniq, but assigned to an ID_Any of a cluster. This is a detection that corresponds to a fraction of a simulated extended source but is not its primary counterpart. We refer to these are split sources corresponding to a cluster;

5. Background fluctuation (BKG): entry in the source catalog that is not associated with an ID_Any. This is a false detection, due to a random fluctuation of the background, and is classified as a spurious source.

The first two classes are additionally divided into three subclasses to study whether the source emission is contaminated by a secondary source. To quantify this, we analyze the photons within 60″ around every input source (denoted as ID_1). If we find at least three photons emitted by a source different than the target, and this number of counts is larger than the square root of the target number counts, we consider the source emitting such photons as contaminating. In this case, we save the ID of the contaminating source as ID_contam to the ID_1 source. This allows separating isolated (not contaminated) sources from clusters and AGN contaminated by another cluster and or AGN. These cases potentially lead to source blending.

Tabelle 2.2: Summary statistics of eSASS catalog for the eRASS1 simulation

	eRASS1 simulation		
AREA	**Class**	**Full** 20 617.8 deg^2	**Clean** 17 703.4 deg^2
CLUSTER	EXT	44 440	38 636
	EXT, $\mathcal{L}_{\rm EXT}$>6	5204	4220
	EXT2	8117	7300
	EXT2, $\mathcal{L}_{\rm EXT}$>6	177	148
AGN	PNT	708 735	574 733
	PNT, $\mathcal{L}_{\rm EXT}$>6	1653	1017
	PNT2	2296	1843
	PNT2, $\mathcal{L}_{\rm EXT}$>6	9	3
STAR	PNT	85 004	49 380
	PNT, $\mathcal{L}_{\rm EXT}$>6	313	178
	PNT2	561	361
	PNT2, $\mathcal{L}_{\rm EXT}$>6	1	1
BACKGROUND	BKG	284 654	229 559
	BKG, $\mathcal{L}_{\rm EXT}$>6	374	48
TOTAL	All	1 133 807	901 812
	$\mathcal{L}_{\rm EXT}$>6	7731	5615

Notes. The table reports the number of eSASS entries that are matched to a certain class of simulated objects (cluster, AGN, star, background). The catalog contains all sources with DET_LIKE >= 5. Each line shows different sources identified in the eSASS catalogs: the number of all matches (point-like and extended), their subsample with EXT_LIKE >= 6 (Extended), the ones relative to secondary matches (i.e., split sources, see EXT2 and PNT2), and the secondary matches that are classified as extended (EXT_LIKE >= 6). The values in the second column include all the simulated tiles, the values in the third one account for the additionally cleaning (see Sect. 2.2.2).

We summarize the simulations and source catalog statistics in Table 2.2.

We show different examples of classification of the sources in Fig. 2.2. The top left panel (a) shows an example of a simulated cluster that is detected as extended with DET_LIKE = 10. The position of the detection is well aligned with the position of the simulated object. The dashed red circle encloses $0.5\times R_{500c}$. The point detection in the center of the panel (b) is assigned to the bright simulated cluster just below, but it is not the primary detection, that is the extended one closer to the cluster center. This is the case of a split source. The third panel (c) highlights a simulated AGN (blue circle) properly detected as a point source (cyan square). The fourth panel (d) shows an example of contamination in the extent-selected catalog: an AGN detected as an extended source. Finally, the fifth panel (e) contains an extended detection without any simulated counterpart: a spurious source. In this case, most of the photons around the detection are coming from the background. This shows how background fluctuations end up decreasing the purity of the source catalog. The second row of the figure (panels f, g, h, i, and l) shows the same type of objects, but

with a higher value of detection likelihood equal to 20. We notice how the distribution of photons around faint detected clusters or AGN and spurious sources is very similar.
We compare the eSASS source catalog from the eRASS1 simulation to real eRASS1 data in Sect. 2.6.

2.2.3 Imaging and spectral analysis

We measure the temperature and luminosity of the simulated clusters detected as extended in the eRASS1 simulation, assuming the value of R_{500c} from the simulation. We compare them to the simulated quantities. We focus on secure clusters detected with EXT_LIKE > 20, spanning different ranges of exposure without additional selection on the sky area. Our approach is the same as the one described by Ghirardini et al. (2021a,b) and is summarized in this section.

1. **Source masking**: for each extended detection uniquely matched to a cluster, we mask every other source inside a circular region of $4 \times R_{500c}$. For extended sources, the masking radius is equal to the extent measured by eSASS. For point-like ones, it corresponds to the point where the count rate convolved with the eROSITA PSF is consistent with the background within 1σ. This value is fixed to 10 arcseconds when it is lower than such threshold.
2. **Background extraction and modeling**: we use the `srctool` command to extract the source spectrum in a circular region inside R_{500c} and the background spectrum in a circular annulus between $3 - 4 \times R_{500c}$. We model these two spectra simultaneously with the XSPEC software (v 12.10.1f, Arnaud, 1996), using C-statistic (Cash, 1979). The cluster emission is fitted by APEC model (Smith et al., 2001) and the Galactic absorption is modeled by TBabs (Wilms et al., 2000). The background model consists of a vignetted sky component and an unvignetted particle-induced one. The first describes photons focused by the telescope mirror and contains contributions from the Local Hot Bubble (apec), the Galactic Halo (tbabs×apec), and faint unresolved AGNs (tbabs×power-law). The second is due to instrumental effects and cosmic rays hitting the detector directly and is described by a combination of power-laws and Gaussian lines (Liu et al., 2021). We fix redshift and galactic column density to the simulated values and fit for temperature.
3. **Surface brightness fitting**: we proceed by measuring the cluster surface brightness inside R_{500c} and fitting the density profile following Vikhlinin et al. (2006) model, convolved with the PSF and projected onto the 2D image plane. The sky (particle) background model is folded with the vignetted (unvignetted) exposure map and added to the total model. The image is fit using the Monte Carlo Markov chain (MCMC) code emcee (Foreman-Mackey et al., 2013). We integrate the fitted 2D profile along the line of sight to obtain the surface brightness radial profile.
4. **Luminosity**: we finally convert the surface brightness radial profile to X-ray luminosity using an absorbed apec model in XSPEC. Given the measured temperature of a cluster, this provides the conversion factor from count rate to luminosity.

2.3 Results

In this section, we present our main findings about the detection process. We start from the point of view of the catalog of sources detected by eSASS. We refer to it as the source catalog. We focus on the cleaned catalog, see Table 2.2 and Sect. 2.2.2 for complete details. We give an overview of how the source catalog is populated by clusters, AGN, stars, and spurious sources (Sect. 2.3.1). We then move to the standpoint of the simulated sources and study which of them are detected. We demonstrate how the method is able to recover clusters and AGN as a function of their simulated flux (Sect. 2.3.2, 2.3.3). We detail how the detection of galaxy clusters depends on size and dynamical state.
We then combine these two points of view, quantifying the performance of the method (completeness, contamination, and spurious fractions), also accounting for the uneven depth of the survey (Sect. 2.3.4).
We study the sensitive area in the eRASS1 simulation as a function of limiting flux (Sect. 2.3.5) and finally verify that our measurement of the X-ray luminosity of clusters are compatible with simulated values (Sect. 2.3.6).

2.3.1 Population in the source catalog

We study the source population in the eSASS source catalog using fractions as a function of different cuts in detection and extension likelihood, using the classes defined in Sect. 2.2.2. We consider the full source catalog and the extent-selected sample (with positive values of EXT_LIKE). The result is shown in Fig. 2.3. We report the fraction corresponding to each class for different thresholds of detection and extension likelihood in Table 2.3. The histograms of the total number of sources and the fractional histograms in linear scales are collected in Sect. 2.7.

Full source catalog

The cleaned source catalog of the eRASS1 simulation contains 901 812 sources in total. Among them, 5615 are classified as extended.

Fraction of point sources

The majority of the catalog consists of point sources, mostly AGN and a few stars. They make up 93.8% of the catalog for detection likelihood larger than 10 in the eRASS1 simulation. For detection likelihood greater than 25, this fraction increases to 94.1%. This is driven by the predominant number density of the AGN population compared to other sources. In the whole cleaned catalog, 574 733 entries are associated with an AGN.

Fraction of clusters in source catalog

In the eRASS1 simulation, clusters of galaxies only consist of about 4.3% of the whole catalog for DET_LIKE > 5. Even when most of the false detections are removed, above

Tabelle 2.3: Population in the cleaned eSASS source catalog for different cuts of detection and extension likelihood.

	eRASS1						
CLASS, DET_LIKE	> 5	> 6	> 7	> 8	> 10	> 15	> 25
PNT	69.214	80.087	86.830	90.717	93.850	94.843	94.056
EXT	4.285	4.615	4.718	4.715	4.635	4.779	5.698
PNT2	0.244	0.195	0.143	0.109	0.065	0.037	0.028
EXT2	0.810	0.749	0.661	0.568	0.437	0.298	0.217
BKG	25.448	14.355	7.648	3.891	1.014	0.042	0.001
	EXT_LIKE > 6						
CLASS, DET_LIKE	> 5	> 6	> 7	> 8	> 10	> 15	> 25
PNT	21.282	21.292	21.253	21.201	21.212	20.513	18.807
EXT	75.156	75.299	75.452	75.634	75.951	77.175	79.763
PNT2	0.071	0.071	0.054	0.036	0.036	0.020	0.025
EXT2	2.636	2.624	2.614	2.589	2.456	2.096	1.405
BKG	0.855	0.714	0.627	0.539	0.346	0.196	0.000

CLASS, EXT_LIKE	> 6	> 7	> 8	> 10	> 15	> 25
PNT	21.268	15.724	12.243	7.464	3.758	1.648
EXT	75.169	80.803	84.383	89.137	92.863	95.655
PNT2	0.071	0.042	0.024	0.029	0.042	0.000
EXT2	2.636	2.787	2.831	3.108	3.252	2.697
BKG	0.855	0.645	0.519	0.261	0.084	0.000

Notes. The five classes (PNT, EXT, PNT2, EXT2, BKG) are defined in Sect.2.2. The fractions are reported in percentage units. The table is divided into three main quadrants. The first one describes the full catalog for different cuts of detection likelihood. The second one focuses on the extent-selected catalog (EXT_LIKE > 6) for different cuts of detection likelihood. Finally, the third one is about cuts of extension likelihood.

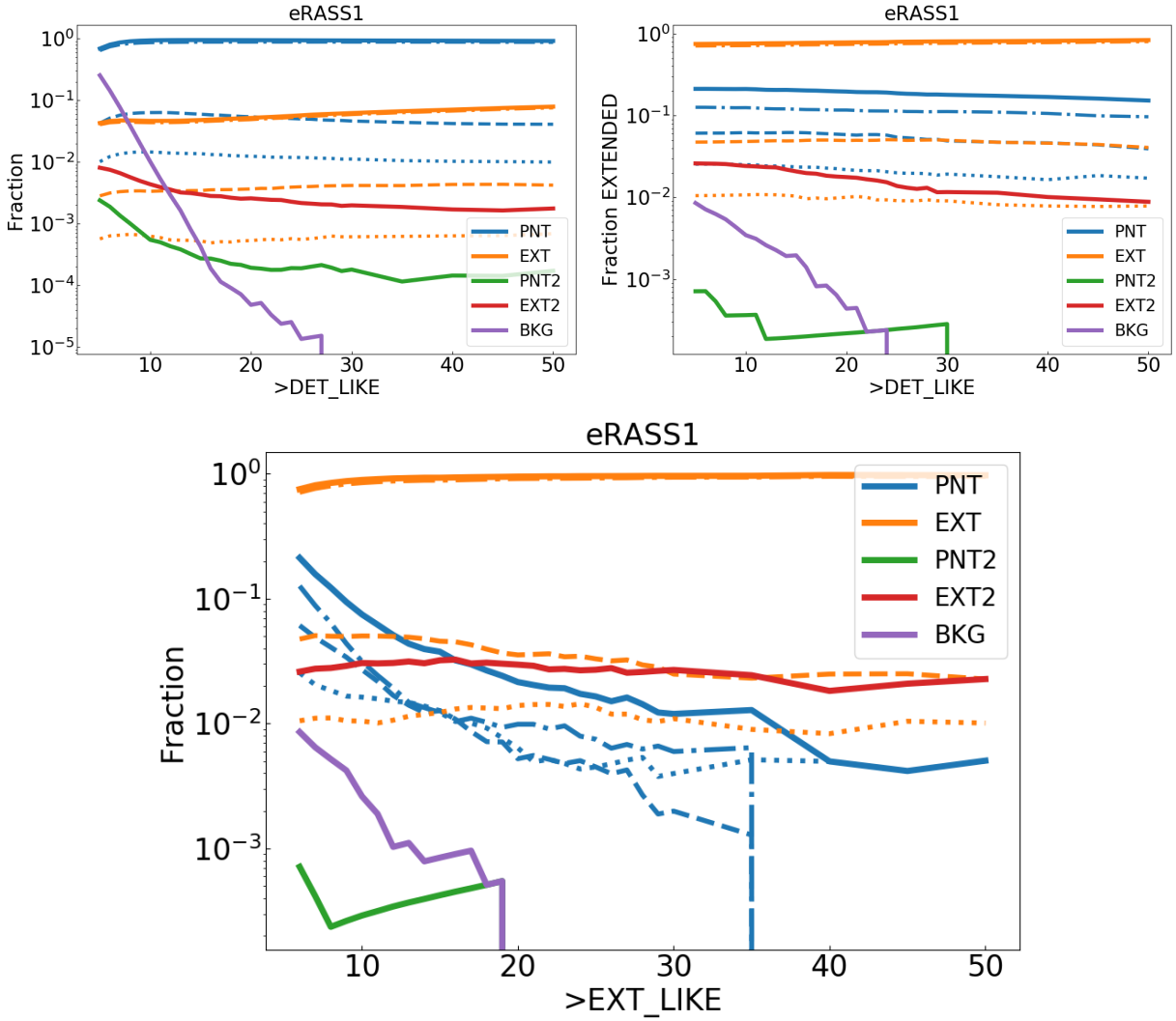

Abbildung 2.3: Population in the eSASS catalog. The total number of sources detected by eSASS in the eRASS1 simulation (cleaned, see 2.2.2) is 1 133 807 (901 812). The number of extended sources is 7731 (5615). **Left-hand panel**: Fraction of sources in the full catalog as a function of minimum detection likelihood. **Right-hand panel**: Fraction of sources in the extent-selected sample (EXT_LIKE >= 6) as a function of minimum detection likelihood. **Bottom panel**: population in the source catalog as a function of minimum extension likelihood. Lines of different colors show the classes defined in Sect. 2.2. The dash-dotted lines denote sources that are not contaminated by photons of a secondary source (no blending), the dashed ones identify sources contaminated by a point source, and the dotted ones show sources contaminated by a cluster.

DET_LIKE = 25, this fraction remains low, at about 6%. This difference between the fraction of AGN and clusters is driven by the intrinsic number density per square degree of these sources. For example, we simulate 18 clusters per square degree with flux larger than

10^{-14} erg/s/cm^2. At the same flux value, the input AGN are 100 per square degree (see Sect. 2.3.2). In addition, clusters need a larger amount of counts to be detected, especially as extended, compared to point sources: their extended emission requires a larger exposure to emerge over the background.

Fraction of spurious detections

A fraction of sources in the eSASS catalog is not matched to any input simulated object. These spurious detections are due to background fluctuations, that mimic the emission of a source. The detection likelihood encodes by definition the probability for each entry in the source catalog of being a false detection, as explained in Sect. 2.2. However, the analytical derivation does not account for additional effects during the measurement process. These include uncertainty in the background estimation, errors in the PSF-fitting, or issues related to hardware and calibration. Consequently, the false detection rate is larger than the one predicted by Equation 2.3.
The fraction of spurious sources drops significantly while increasing the detection likelihood threshold. We measure a spurious fraction of 25.4% for DET_LIKE $>$ 5 and 14% for DET_LIKE $>$ 6. The false detection rate is further reduced to 4% at DET_LIKE $>$ 8 and 0.001% for DET_LIKE $>$ 25. Progressive cuts in detection likelihood are therefore efficient in removing background fluctuations from the source catalog.

Fraction of split sources

Very bright and or extended input sources are possibly split into multiple detections. These are the one marked as secondary matches (PNT2, EXT2) in our classification scheme (see Section 2.2.2). The fraction of entries in the source catalog marked as a secondary match to a point source (PNT2) is always under 0.5%. Clusters are instead slightly more easily split into multiple sources, giving about 0.8% of entries cataloged as secondary matches to an extended source (EXT2). Together with decreasing the spurious fraction, increasing the DET_LIKE threshold gets rid of these low significance secondary detections, as this fraction decreases to $\sim$0.2 % at DET_LIKE $>$ 25. A total of 4627 clusters are split into more than one (point or extended) source in the eRASS1 simulation. About 70% of these are split into only two sources. Among the clusters that are split, the average number of split sources is 2.76. We find that the number of sources into which a cluster is split mainly depends on its flux, and secondary the size on the sky plane of the cluster itself. For example, more than 90% of the clusters with R_{500c} larger than 350 arcseconds are split into multiple sources. However, only the brightest objects are split into a large number of sources. A very bright and extended cluster with flux $\sim 10^{-11}$ erg/s/cm^2 and R_{500c} $\sim$500″ is split into 24 sources by eSASS on average. There is one particular case of an extremely bright and extended cluster ($F_X = 3.10\times10^{-11}$ erg/s/cm^2, $R_{500c} = 13.5'$) in the pole region that is split into 65 sources. These trends are highlighted in Fig. 2.4. The left-hand panel shows the fraction of clusters that are split into multiple sources as a function of flux and R_{500c}. The average number of sources that a cluster is split into is displayed on the

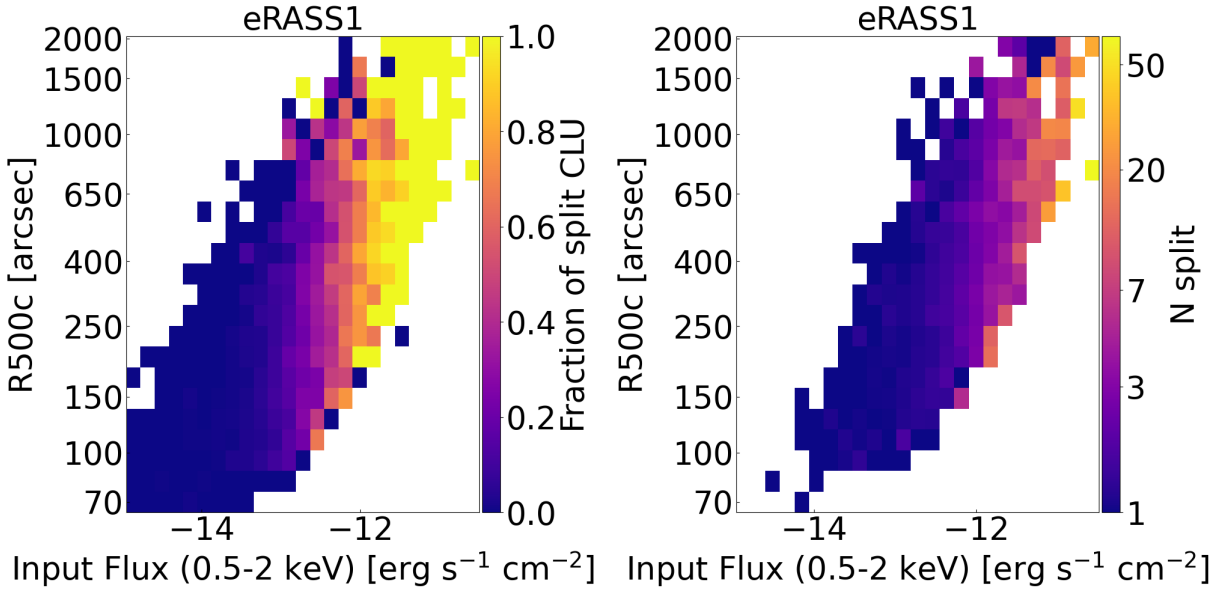

Abbildung 2.4: Number of split sources as a function of flux and R_{500c}. The left-hand panel shows the fraction of detected clusters that are split into multiple sources, the right-hand one displays the average number of sources which a cluster with given flux and size is split into. The blank spaces contain no input clusters.

right-hand panel.

Blends

We study the sources that are blended with a secondary one, according to the criteria defined in Sect. 2.2 to find objects whose emission is contaminated by another object. Most of the sources detected by eSASS are not contaminated by the emission of a secondary nearby object. 92% of the detected point sources are isolated. This number for clusters is 94%. In the full cleaned catalog, about 4% of the population consists of point sources contaminated by other point sources. This number increases to 6% for DET_LIKE > 10, because of the drastic drop of spurious sources. For point sources contaminated by clusters (i.e., detections whose primary match is an AGN or a star, but that contain photons emitted by a cluster) this fraction reduces to 1%. About 7% of the clusters in the full catalog are contaminated by other point sources. In such cases, the presence of the bright AGN enhances the emission from a physical source and helps the detection algorithm in the identification of a source at that position. The flux measured by eSASS for these sources will be biased (see Bulbul et al., 2022). More detailed modeling of the cross-correlation between AGN and clusters is required to reach conclusive statements about blending.

Extended source catalog

We now focus on the extent-selected sample, selected by EXT_LIKE >= 6. This is the minimum value of extension likelihood fixed by the choice of the parameter extlikemin = 6 (Sect. 2.2). Different values of this parameter impact the properties of the extent-selected catalog. We detect a total of 5615 sources as extended in the full cleaned eSASS catalog of the eRASS1 simulation.

Fraction of clusters

The eRASS1 extent-selected catalog is dominated in numbers by clusters of galaxies: 75.2% of the eSASS sources are uniquely matched to a cluster, with a 21.2% point source contamination. When increasing the detection likelihood threshold to 25, clusters make up 79.8% of the catalog. These numbers increase more significantly when cutting in extension likelihood rather than detection likelihood. For EXT_LIKE > 25, 95.6% of the eRASS1 sources are clusters. This is partially related to the significant decrement of background fluctuations, which is completely canceled at this value of extension likelihood. However, the main contribution is given by the drop of AGN that are mistakenly classified as extended sources, which reduces contamination significantly. This fraction changes from 21.2% for EXT_LIKE > 6 to 1.6% for EXT_LIKE > 25 in the eRASS1 simulation.

Fraction of AGN

The fraction of AGN in the extent-selected sample (EXT_LIKE >= 6) is constant at around 20% as a function of detection likelihood cuts. Even for DET_LIKE greater than 25, it still reaches 18.7%. It means that progressive thresholds of detection likelihood are not efficient in reducing the fraction of AGN detected as extended.
The contribution of detections that contain a fraction of point source signal (PNT2, split point sources) is minimal in the extended select sample. It is well below 1% for any cut in detection or extension likelihood. The fraction of entries classified as cluster signal (EXT2, split clusters) is around 2.6% for eRASS1. Increasing the extension likelihood does not have a significant impact on this number. This is due to the fact that the scaling of these secondary matches with EXT_LIKE is more similar to the one of primary matches, compared to the random background fluctuations. This is not true for cuts in detection likelihood, which keep a higher number of AGN in the extent-selected sample, reducing the relative contribution of both primary and secondary matches in the catalog. In fact, by increasing the DET_LIKE threshold in the eRASS1 catalog from 5 to 25, the fraction of secondary matches also drops from 2.6 to 1.4% for extended sources and from 0.071% to 0.025% for point-like ones respectively.

Fraction of spurious sources

Random background fluctuations in the extent-selected catalog are efficiently removed using different thresholds of DET_LIKE and EXT_LIKE. For the former, the spurious

fraction drops from 0.85% to 0.34% for detection likelihood larger than 5 and 10. The latter drops to 0.26% for EXT_LIKE > 10. There are no spurious sources above detection likelihood larger than 25 and extension likelihood larger than 20 in the extent-selected sample of eRASS1. The decrement of the false detection rate is steeper as a function of EXT_LIKE cuts. It means that, on top of reducing contamination, extension likelihood thresholds remove background fluctuations more efficiently than detection likelihood ones in the extent-selected sample.

Blends

We study sources blended with another source in the extent-selected catalog (EXT_LIKE >= 6). Focusing on the AGN that leak into the extent-selected sample, one can understand what caused the misclassification. For the whole extended eRASS1 sample, 6% of the catalog consists of AGN contaminated by another point source. This fraction is dominant over the ones contaminated by a cluster (2.5%). However, when increasing the EXT_LIKE threshold, the relation between these two classes changes significantly, to the point where for EXT_LIKE > 40, all the detections assigned to an AGN by our matching algorithm are actually blended with a cluster. Follow-up observations in the optical band have the potential to confirm these clusters, which lowers our estimate of contamination in the extended select sample due to bright point sources by ∼1%.

2.3.2 Simulated and detected sources

We now study which simulated sources are detected by eSASS. The detection process mainly depends on the net count rate of each source. Bright sources with large flux values provide a larger number of photons on the detector. Therefore, it is easier for the detection algorithm to identify them, compared to fainter objects dispersed in the local background. We investigate which simulated sources are detected by studying the number density as a function of the input flux threshold for AGN in Sect. 2.3.2, and for clusters in Sect. 2.3.2.

AGN logN–logS

We measure the cumulative number of detected AGN per square degree as a function of the input flux (0.5–2 keV band). We compare with the distribution of the simulated AGN (Comparat et al., 2019), with the observations from Gilli et al. (2007) and Georgakakis et al. (2008), and the collection from Merloni et al. (2012). The result is shown in the upper panel of Fig. 2.5. At the high flux end, the different shapes of the function denoting eRASS1 and other samples are expected due to the AGN simulation method in HEALPix fields as described in Sect. 2.1. It reduces the volume probed by the model and the total number of the brightest AGN consequently decreases, but this method guarantees a significant gain in computation time. Given our goal of studying the simulated objects that are detected, this has no impact on our purpose. In the lower panel, we show the ratio between the logN–logS built with the detected and simulated populations of AGN. Below the predicted

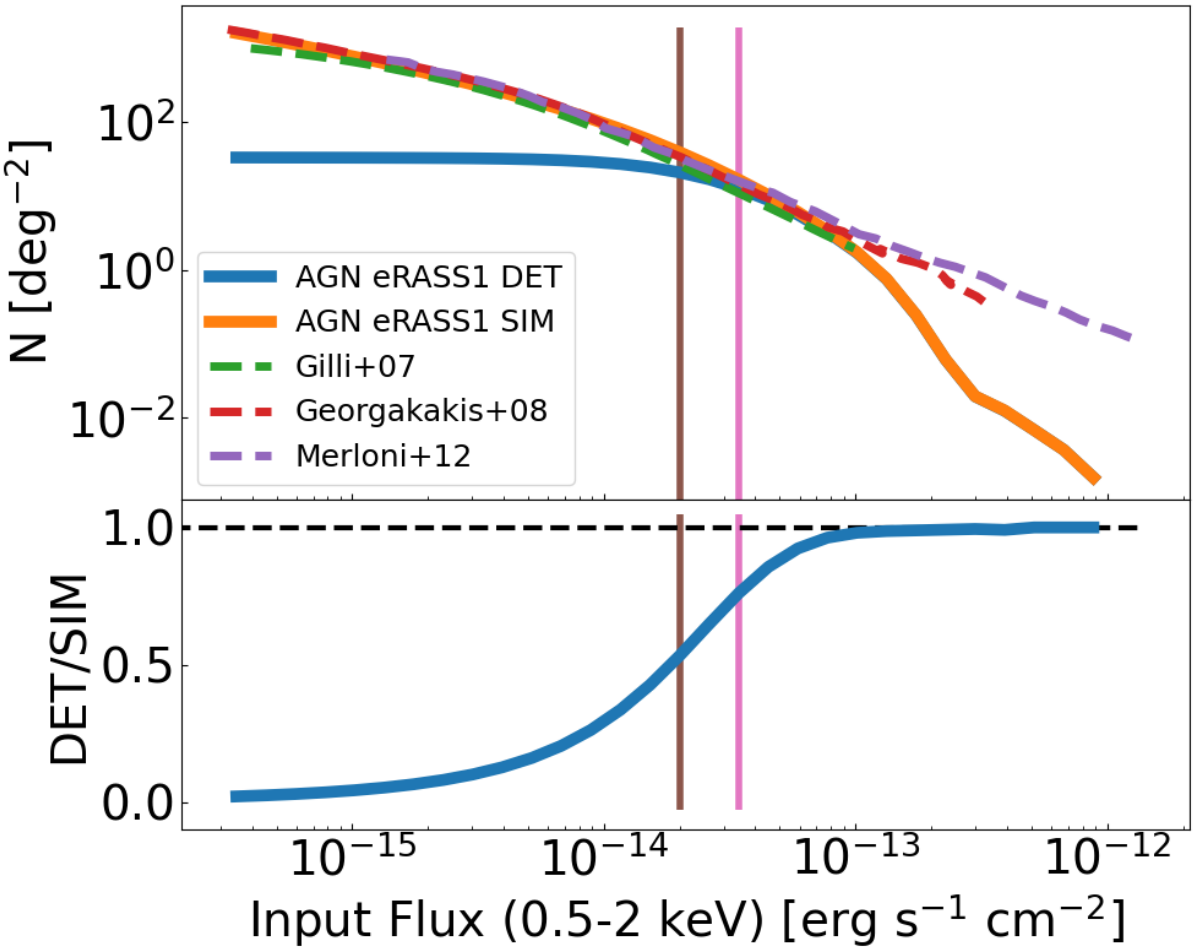

Abbildung 2.5: Cumulative number density of the AGN population. **Top panel**: The blue (orange) line shows the logN-logS built with the sample of detected (simulated) AGN. The green, red, violet dashed lines show the distributions from Gilli et al. (2007), Georgakakis et al. (2008) and Merloni et al. (2012). The brown and pink vertical lines locate the eROSITA flux value where the ratio between the detected and simulated populations is equal to 0.5 and 0.8, respectively. **Bottom panel**: Ratio between the logN-logS of detected and simulated AGN. A black dashed line denotes a ratio equal to 1.0.

eROSITA flux limits at $\sim 4{\times}10^{-14}$ erg/s/cm^2 for eRASS1 (see Merloni et al., 2012, Figure 4.3.1), the number density of detected AGN deviates from the simulated one (solid curves depart from the dashed ones). Toward high fluxes, the number density of detected AGN converges to the simulated one. The ratio between these two curves reaches a value of 0.5 at $\sim 2{\times}10^{-14}$ erg/s/cm^2 for eRASS1. These numbers rise to $\sim 3.5{\times}10^{-14}$ erg/s/cm^2 and a ratio of 0.8 between the logN–logS of detected and simulated AGN. This is in excellent agreement with the prediction of the eRASS1 sensitivity for point sources in the same soft band 0.5–2.0 keV from (Merloni et al., 2012). The completeness of the source catalog behaves smoothly as a function of flux and is in line with the expectations. We study the completeness fraction of AGN in more detail and provide analytical fits in Sect. 2.8.

Cluster logN–logS

We study the cumulative number density of clusters as a function of the input flux. We detect 0.1 clusters per square degree with flux larger than $4{\times}10^{-13}$ erg/s/cm^2 in the eRASS1 simulation. We detect all the clusters at the brightest flux end, as the ratio between the logN-logS built with detected and simulated clusters reaches a value of 1.0 for the eRASS1 simulation. It is equal to 0.5 for flux values of $\sim 3{\times}10^{-13}$ erg/s/cm^2. For the same flux limit, about 70% of the clusters with mass larger than $M_{500c}{>}3{\times}10^{14}$ M$_\odot$ are detected as extended sources. A ratio of 0.8 is reached for flux values of $\sim 1.5{\times}10^{-12}$ erg/s/cm^2 for eRASS1. These flux limit values are larger compared to the AGN ones. A different flux limit is thus expected between the two populations. The extension of the cluster model to galaxy groups allows a smooth transition between the faintest clusters that are not detected and the ones above the survey flux limit. The detection method is able to fully recover the bright end of the cluster sample. Around the flux limit, additional selection effects, such as the cool core bias or the size of the object on the sky plane, influence the detection process. In addition, at fixed simulated flux, due to their spatial extent, clusters will be detected with a lower likelihood compared to a point source with the same flux. We report the cumulative clusters number density as a function of flux in Fig. 2.6. In the upper panel, we show the cluster logN–logS for eRASS1. The blue line denotes the detected cluster population, while the orange line the simulated one. The green one adds a comparison to the eFEDS logN–logS (Liu et al., 2022). We additionally compare our result to The SPectroscopic IDentification of eROSITA Sources observational program (SPIDERS, Clerc et al., 2016; Finoguenov et al., 2020) denoted by the red dashed line, and the Extended Chandra Deep Field South (ECDF-S, Finoguenov et al., 2015), indicated by the purple dashed line. There is good agreement within these samples. The bottom panel shows the ratio between the detected and the simulated populations. All clusters with high flux are detected as extended. We present the challenges of the detection of extended sources in Sect. 2.3.3.

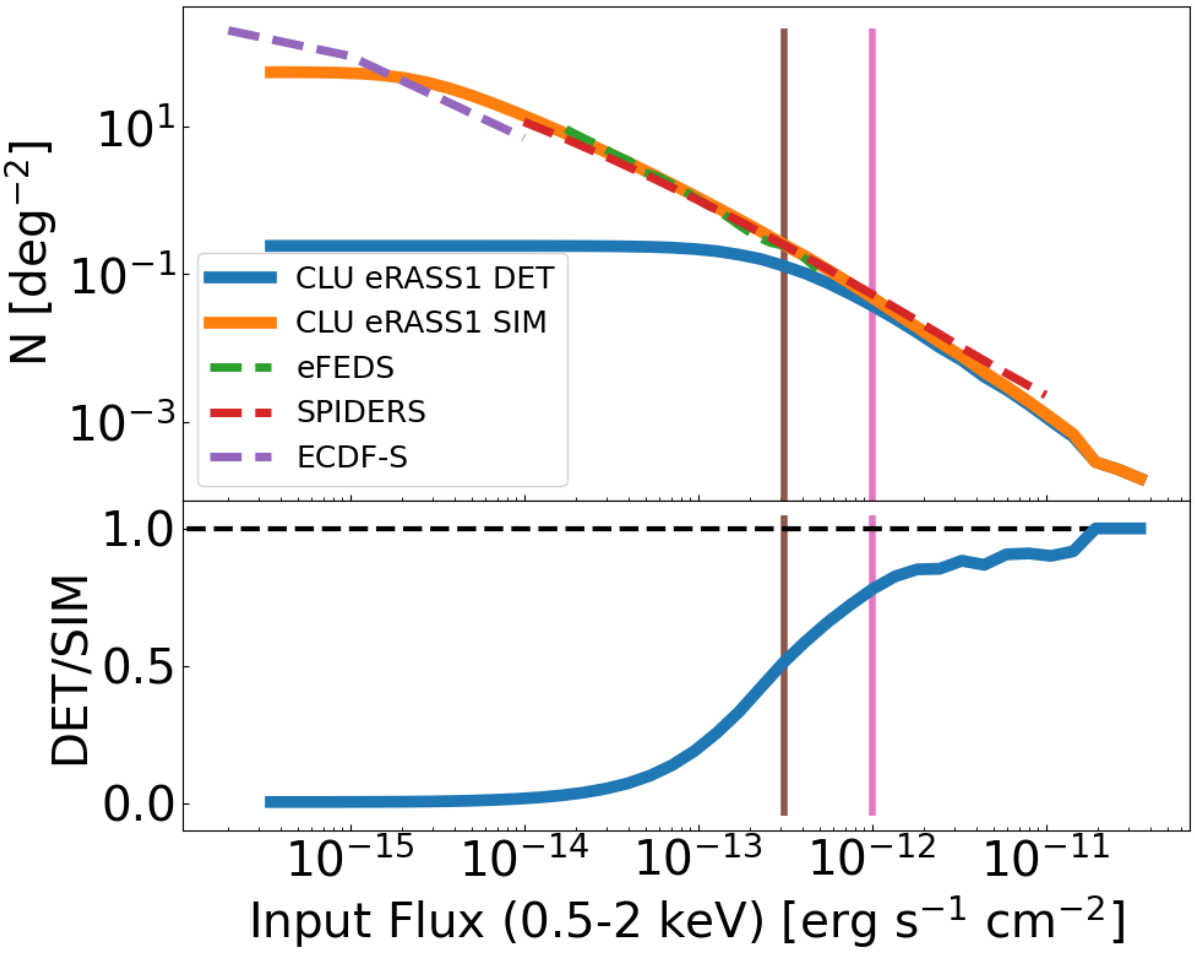

Abbildung 2.6: Cumulative number of clusters per square degree as a function of flux. **Top panel**: The solid blue (orange) line shows the logN-logS built with the sample of detected (simulated) clusters. The green dashed line shows the distributions of the eFEDS sample (Liu et al., 2022), the red one denotes the SPIDERS sample (Finoguenov et al., 2020), and the pink one the ECDF-S (Finoguenov et al., 2015). The brown and pink vertical lines locate the eROSITA flux value where the ratio between the detected and simulated populations is equal to 0.5 and 0.8, respectively. **Bottom panel**: Ratio between the logN-logS of detected and simulated clusters. A black dashed line denotes a ratio equal to 1.0.

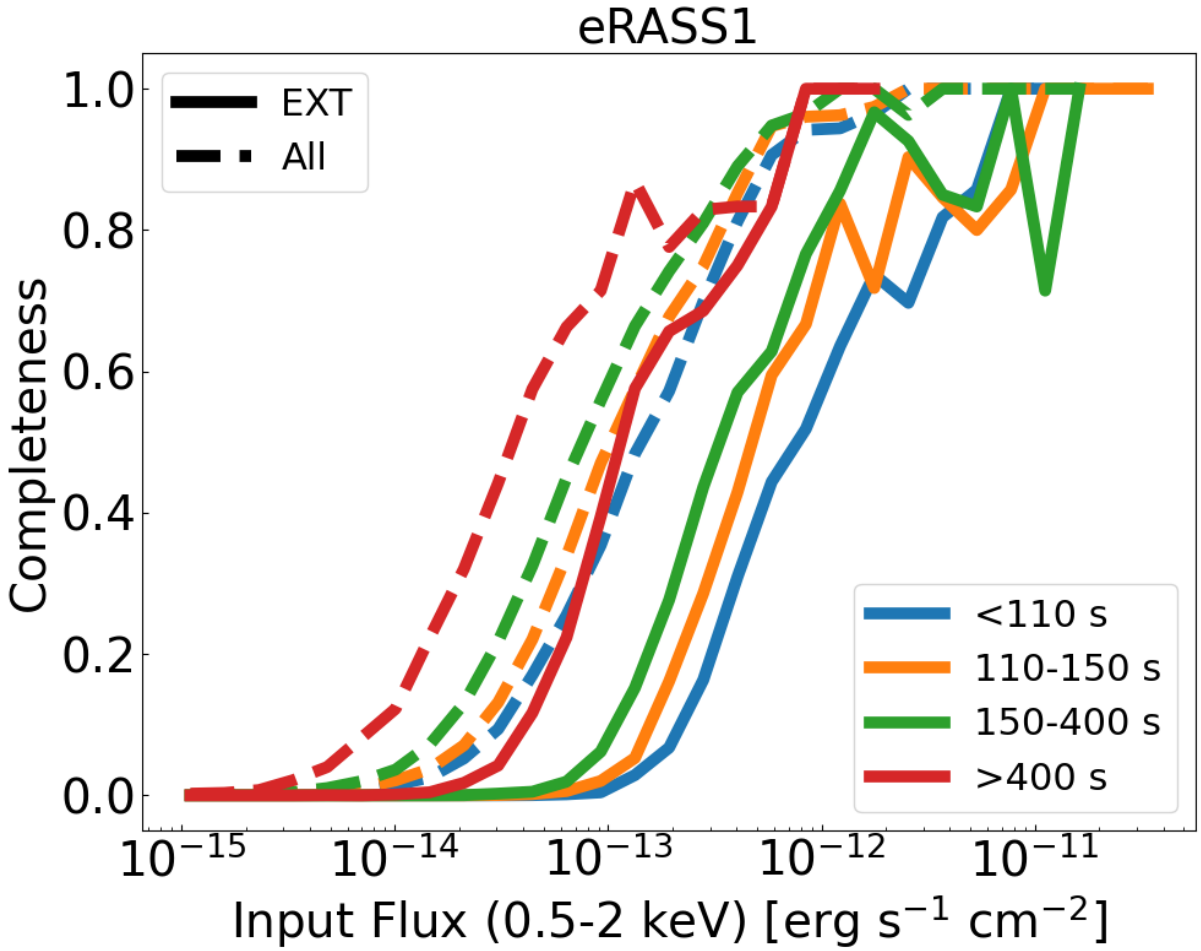

Abbildung 2.7: Fraction of simulated clusters with a counterpart in the eSASS catalog as a function of simulated soft X-ray flux. We do not apply any additional likelihood selection. Each color identifies an exposure time range. Solid lines denote clusters only detected as extended, while dashed ones include the ones detected as point sources.

Tabelle 2.4: Different exposure and properties of the eRASS1 simulations.

	eRASS1						
Exposure	Area [deg^2]	N_{CLU}/deg^2	Flux CLU 50%	Flux CLU 80%	N_{AGN}/deg^2	Flux AGN 50%	Flux AGN 80%
$<$ 110 s	6710	0.13	7.13×10^{-13}	3.39×10^{-12}	21.78	3.76×10^{-14}	7.02×10^{-14}
110 s – 150 s	4543	0.22	4.67×10^{-13}	1.2×10^{-12}	29.41	3.01×10^{-14}	5.31×10^{-14}
150 s – 400 s	6073	0.34	3.28×10^{-13}	9.72×10^{-13}	42.94	2.22×10^{-14}	3.98×10^{-14}
$>$ 400 s	377	1.05	1.12×10^{-13}	4.75×10^{-13}	93.71	1.10×10^{-14}	1.93×10^{-14}

Notes. Each column denotes: exposure interval, Area covered with the given exposure, number density of clusters detected as extended, flux limit where the completeness is equal to 0.5 for clusters detected as extended, flux limit where the completeness is equal to 0.8 for clusters detected as extended, number density of AGN detected as point sources, flux limit where the completeness is equal to 0.5 for AGN detected as points, and flux limit where the completeness is equal to 0.8 for AGN detected as points.

2.3.3 Cluster completeness

The completeness is defined as the ratio between the number of detected and simulated objects, see Eq. 2.5:

$$C = \frac{N_{\rm DET}}{N_{\rm SIM}}. \tag{2.5}$$

We measure the completeness of our source catalog as a function of the input flux in the 0.5–2 keV band. We study areas in the sky covered by different depths. We expect to measure higher completeness where the exposure is longer, which allows detecting a higher number of clusters. We consider four exposure time bins in this work, defining shallow, medium, deep, and pole regions. The respective intervals are $<$ 110 s, 110 s – 150 s, 150 s – 400 s, $>$ 400 s for the mock eRASS1. Such intervals are designed to identify three regions covering roughly a similar area on the sky, and a fourth, smaller one that encloses the pole with large exposure. Additional details are provided in Table 2.4. With this approach, we can quantify the gain of detected clusters thanks to deeper observations. We show the result for eRASS1 in Fig. 2.7. The lines are color-coded according to exposure time intervals. The solid lines show clusters of galaxies detected as extended, dashed ones additionally consider clusters detected as point sources with EXT_LIKE = 0. Adding the latter population increases completeness at a fixed value of flux. Focusing on the objects detected as extended, we measure a completeness fraction of 0.5 at 3.3×10^{-13} erg/s/cm^2 for regions around the average eRASS1 exposure of about 275 s, denoted by the green solid line. This result is comparable with previous predictions by Clerc et al. (2018), who measured a completeness value of 0.5 at $\sim 5\times10^{-14}$ erg/s/cm^2 in equatorial fields with eRASS:8 depth of about 2.0 ks. The decrement of completeness in the 150 s – 400 s range is due to a merging system, where only one eSASS detection with EXT_LIKE>6 is present. The latter is assigned to one of the two clusters, the one providing most of the counts around the detection. The second cluster is assigned to a nearby point-like detection instead. Adding the clusters detected as point sources increases completeness. For the depth interval 150 s – 400 s in eRASS1, the 50% completeness is reached at flux equal to 8×10^{-14} erg/s/cm^2. There is a flux difference of about 0.7 dex with the addition of this population.
The measure of completeness is positively correlated with exposure time. In the eRASS1

simulation, the fraction of clusters with flux $\sim 5{\times}10^{-13}$ erg/s/cm^2 that are detected as extended goes from 0.39 (exposure < 110 s) to 0.8 (exposure > 400 s).

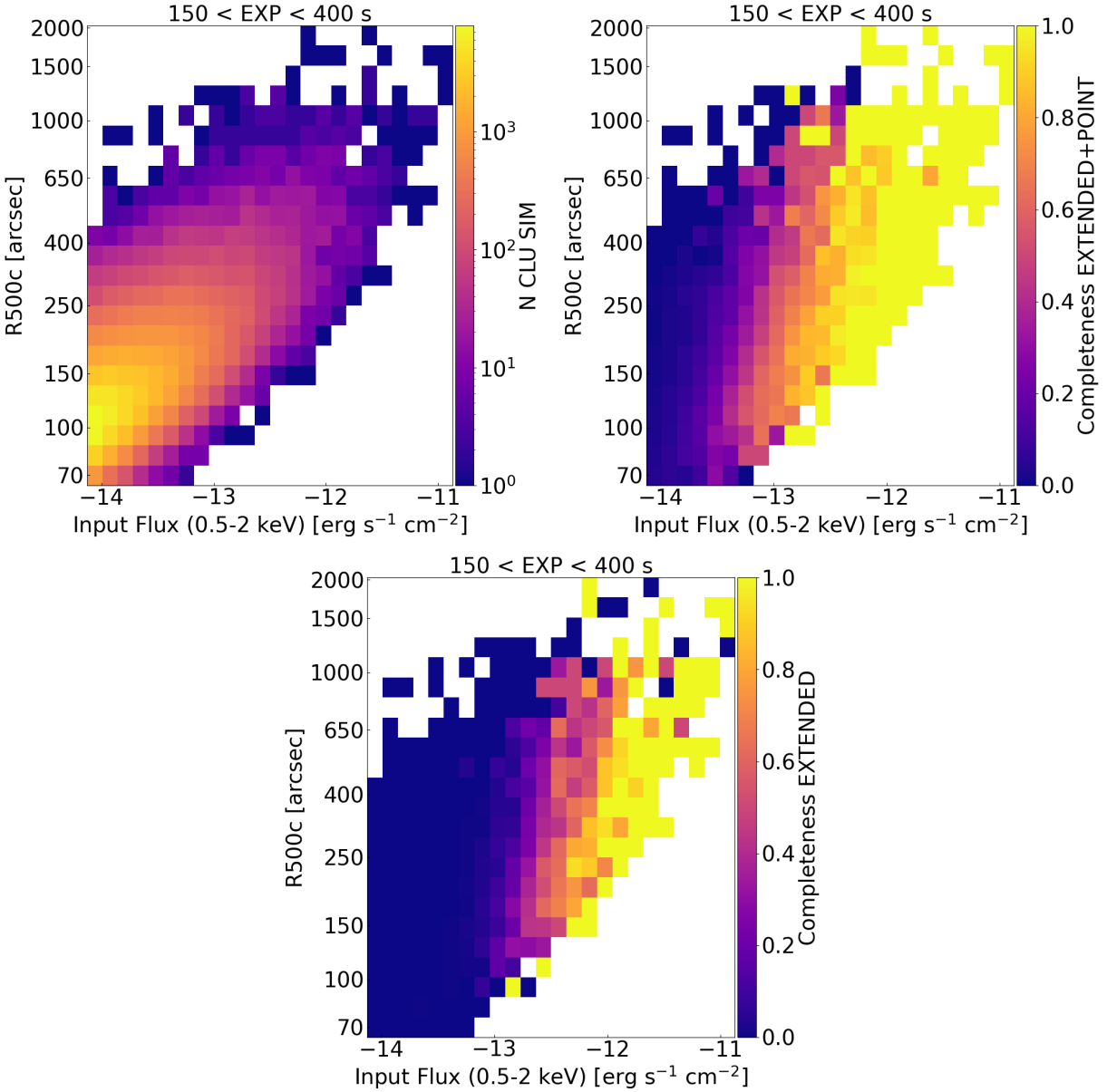

Abbildung 2.8: Simulated and detected clusters population as a function of the input flux and size on the sky. The figures refer to areas of the eRASS1 simulation covered by an exposure between 150 s and 400 s. The blank spaces contain no input clusters. **Left-hand panel**: number of simulated clusters in the flux–R_{500c} space. **Right-hand panel**: fraction of simulated clusters that is detected by eSASS, either as extended or point source. **Bottom panel**: fraction of simulated clusters that is only detected as extended.

The increase in the number of detected objects between the shallow and deep regions

is expected, but nevertheless remarkable. It translates into an increment of the number density of clusters detected as extended with exposure time. In the former, we detect and properly classify as extended 0.13 clusters per square degree (exposure < 110 s). In the latter, such number increases to 1.05. Our result is in agreement with previous works (Pacaud et al., 2006; Clerc et al., 2012, 2018). This only means that we recover a larger number of simulated clusters in deep areas, not that the detection is necessarily more efficient. A different fraction of spurious sources is also detected in areas with large exposure because the background has lower fluctuations. Its overall level might be larger, but its lower variability may also reduce the false detection rate. Such deep areas additionally suffer from a higher degeneracy between blended point sources and proper extended ones, as well as between AGN in clusters and cluster substructures, which has an impact on the measure of contamination. A detailed discussion is presented in the next section 2.3.4. We do a similar study for AGN and provide details and analytical fits in Sect. 2.8.

Completeness and apparent cluster size

We investigate the impact of the apparent physical size of the clusters on the sky on the detection. This information is encoded in the critical radius R_{500c}. We compare the number of detected objects to the simulated one on a 2D grid of flux and R_{500c}. Considering the angular size of the cluster on the sky (e.g., in arcseconds) instead of its physical size (in kpc) allows to additionally account for the impact of redshift, which makes distant massive large clusters appear smaller than nearby ones with similar mass.
We find that the detection of extended sources is not solely a simple function of flux and exposure time. At fixed flux and exposure time, the completeness varies as a function of the size of the clusters on the sky. In the eRASS1 simulation, bright clusters with flux$\sim 1\times10^{-12}$ erg/s/cm^2, located in an area covered by exposure 150 s – 400 s, and $R_{500c} = 180''$ are detected as extended with a completeness of 0.75. The rest of these sources are actually detected but misclassified as point sources. In fact, the completeness reaches a value of 1.0 when adding the population of clusters detected as point-like objects. At the same value of flux, for larger objects with $R_{500c} = 300''$, we measure a completeness of 0.84. The characterization of extremely large clusters is also challenging, because these can be split into multiple sources. In fact, the completeness decreases for large values of R_{500c}, above $400''$ and flux of 1×10^{-12} erg/s/cm^2. As R_{500c} increases, the surface brightness goes down rapidly. Therefore these cases represent the population of clusters which are very extended but with very low surface brightness, therefore they are harder to be detected. This is shown in Fig. 2.8. It displays the number of the simulated clusters population in the upper panels, the fraction of these objects that are detected as extended or point sources in the central ones, and finally only the ones classified as extended in the lower panels. It focuses on exposure intervals containing the average depth for our simulation, in the 150 s – 400 s range for eRASS1. This figure confirms the trends of increasing completeness with flux (see Fig 2.7). In addition, it demonstrates how the selection of extended sources is not a simple function of flux and exposure, but also of the size the object on the sky, encoded in our measure of R_{500c}.

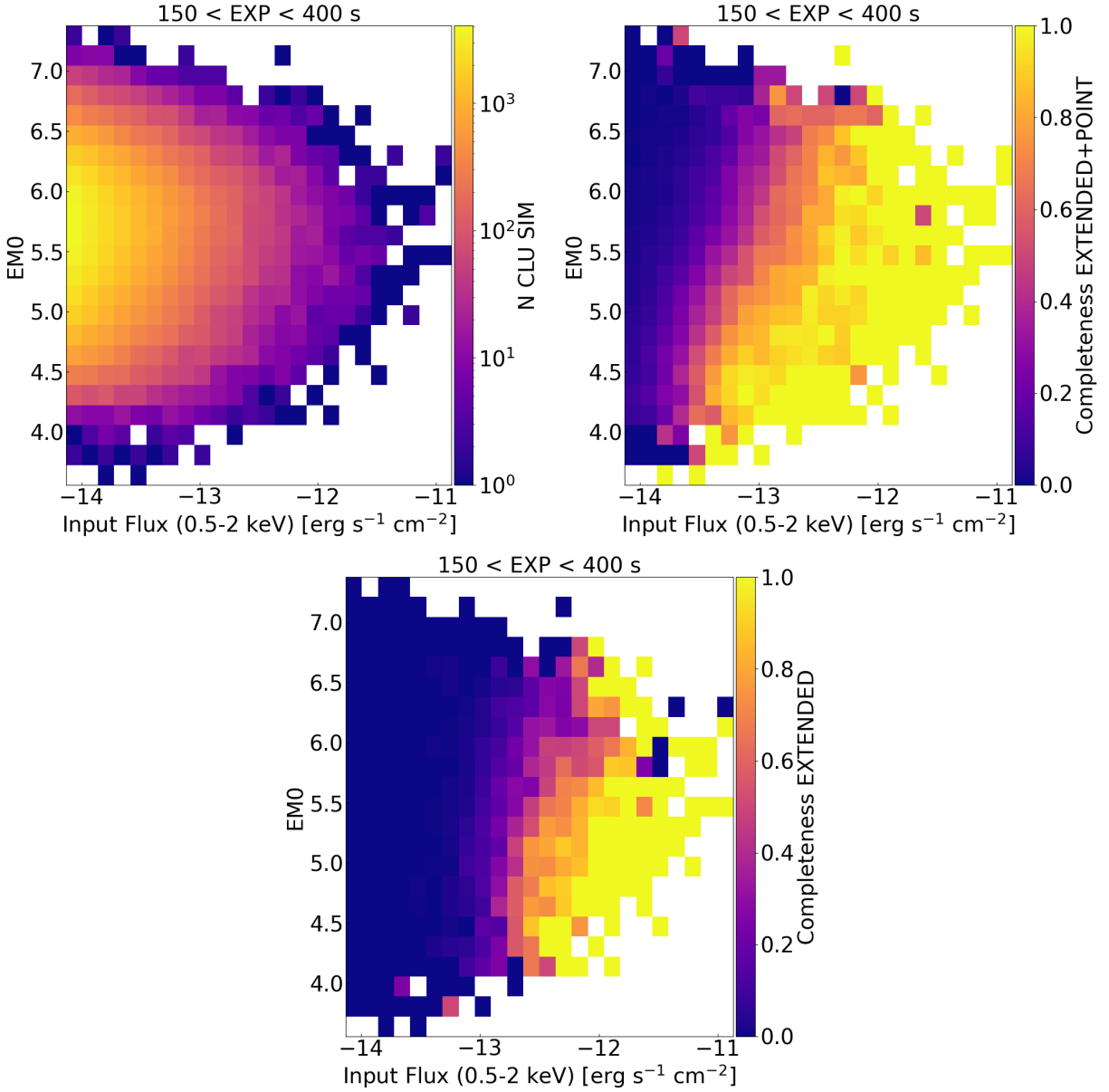

Abbildung 2.9: Population of simulated and detected clusters as a function of the input flux and dynamical state. The panels show areas of the eRASS1 simulation covered by an exposure between 150 s and 400 s. The blank spaces contain no input clusters. **Left-hand panel**: number of simulated clusters in the flux–EM0 space. **Right-hand panel**: fraction of simulated clusters that is detected by eSASS, either as extended or point source. **Bottom panel**: fraction of simulated clusters that is only detected as extended.

Completeness as a function of the central emissivity

We study the impact of the clusters dynamical state on the detection. Such property is related to the central cluster emission. In the simulations, we relate the emissivity in the central region of the cluster to a parameter of the dark matter halo ($X_{\rm off}$) which encodes its dynamical state. The offset parameter, $X_{\rm off}$, is the displacement between the halo center of mass and its peak of the density profile (Klypin et al., 2016; Seppi et al., 2021). The negative $\log_{10}$ of the central emissivity (EM0) is proportionally related to $\rm X_{off}$ (see Comparat et al., 2020, for more details). Dynamically relaxed dark matter halos (with low offset parameter) host clusters with peaked emissivity profiles (cool cores with high central emissivity, and low EM0 in this formulation). Conversely, disturbed halos (with large offset parameter) host noncool core clusters with flatter emissivity profiles. We measure the completeness fraction as a function of EM0 for clusters in different bins of flux (Fig. 2.9). This allows quantifying the impact of the cool core bias, which makes the detection more efficient toward clusters with a peaked emission in the core. We describe the results for the eRASS1 simulation in the following paragraph.
We find that clusters with low flux are hardly detected as extended. In this regime, where few objects are detected, they are mostly characterized as point sources. About 25% of the simulated objects with a flux of $\sim 3{\times}10^{-14}$ erg/s/cm^2, EM0 $\sim$ 5, and covered by an exposure between 150 s and 400s are detected, but none of them is classified as extended. At these low fluxes, we see evidence of the cool core bias. In fact, we detect only 7% of the extremely unrelaxed simulated clusters with EM0 = 6 at this flux value, as the completeness drops by a factor of $\sim$ 3.5 from relaxed to disturbed structures. The detection is generally more efficient for brighter objects. 82% of disturbed structures (EM0 = 5.5) and flux $\sim 3{\times}10^{-13}$ erg/s/cm^2 are identified by eSASS and 39% of them are characterized as extended. In addition, in this regime the cool core clusters are still detected as extended sources. For instance, at the value of EM0 = 5, every cluster brighter than $\sim 1{\times}10^{-12}$ erg/s/cm^2 is properly classified as extended. There is a smooth transition between these two regimes: 85% of the extreme cool cores (EM0 = 4.5) with flux $\sim 1{\times}10^{-13}$ erg/s/cm^2 are detected, but only 14% is identified as extended. Moving to the bright end of the flux distribution, the sample becomes less affected by the cool core bias. Among the clusters with flux $\sim 1{\times}10^{-12}$ erg/s/cm^2, relaxed (disturbed) ones with EM0 = 4.5 (EM0 = 5.5) are detected as extended in 100% (91%) of the cases. This transition is clear by comparing the central and bottom panels of Fig. 2.9. They remark the different behavior of the completeness for simulated bright cool cores between the clusters only identified as extended and the sample with the addition of the point-like detections in the eRASS1 simulation. When including the clusters detected as point sources (central panel), the population is skewed toward lower values of EM0, especially at low flux. This effect is mitigated in the extent-selected sample (bottom panel). We further discuss an explanation in Sect. 2.4.1. We conclude that the cool core bias strongly affects only the faint clusters detected as point sources. Its impact on brighter objects detected as extended is reduced. Our results suggest that a stricter selection focused on bright eROSITA clusters with larger values of extension likelihood provides a sample that is barely affected by the cool core bias. This is in agreement with Ghirardini et al.

(2021a), who do not find a clear preference for cool core clusters in the extent-selected eFEDS clusters, and Bulbul et al. (2022), who find steeper emissivity profiles and more concentrated objects only within the sample of eFEDS clusters detected as point sources.

2.3.4 Detection efficiency

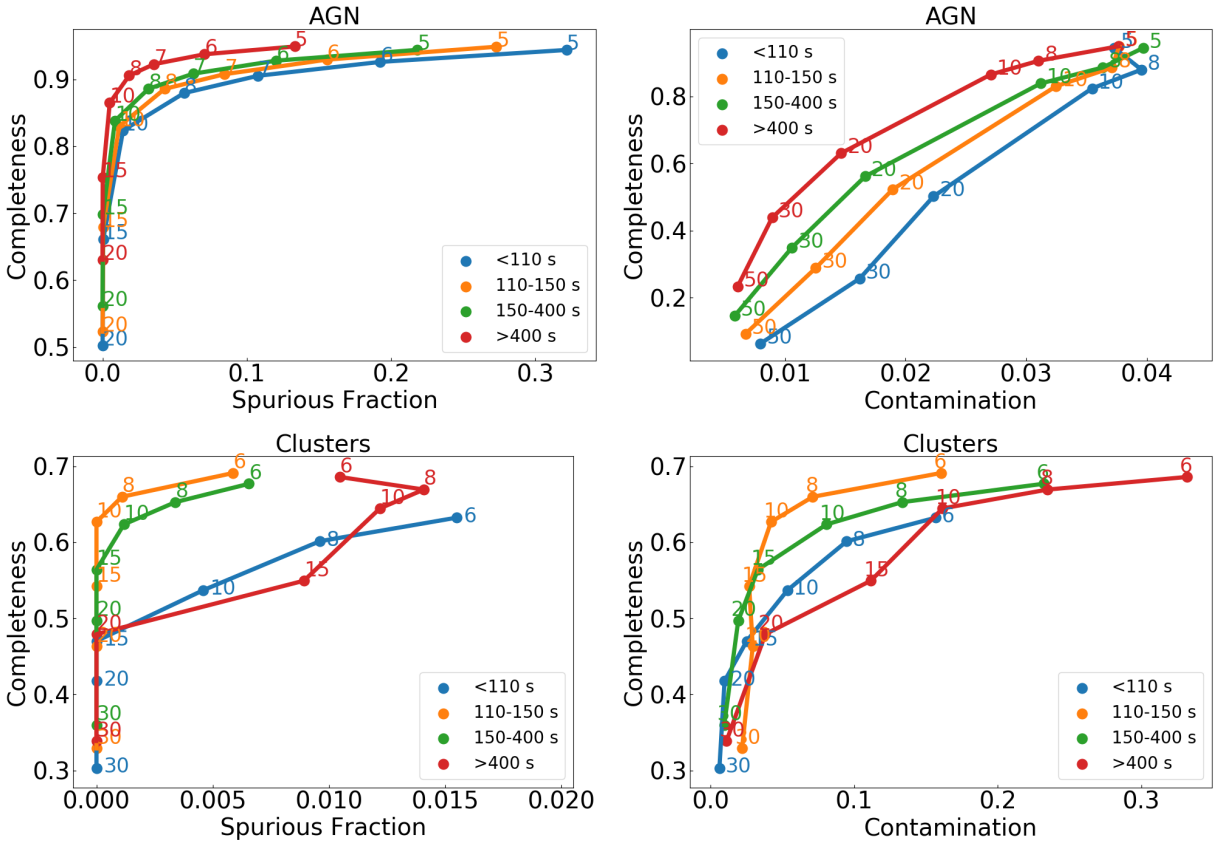

Abbildung 2.10: Efficiency of the eSASS detection for extragalactic sources in the eRASS1 simulation. The completeness is measured for simulated objects above the different flux limits for each exposure interval defined in Table 2.4. **Top panels**: Detection efficiency for AGN detected as point sources (EXT_LIKE = 0). The numbers denote DET_LIKE thresholds. **Bottom panels**: Detection efficiency for clusters. The numbers denote EXT_LIKE thresholds. No additional cuts of DET_LIKE are applied. **Left-hand panels**: Completeness as a function of spurious fraction. **Right-hand panels**: Completeness as a function of contamination. Different exposure intervals are shown in different colors.

Different aspects come into play when evaluating the performance of a detection algorithm. We quantify the ability to recover simulated sources, to properly classify them as point-like or extended, and to minimize the false identification of background fluctuations. The first one is completeness, that is the fraction of simulated objects with a given flux

that have been detected, see Eq. 2.5. When measuring this number, we choose different flux thresholds according to exposure time in the two simulations. We consider the same depth bins described in previous sections (see Table 2.4). We work with exposure intervals and flux values identifying the 50% completeness for clusters detected as extended from Fig. 2.7, and the 80% completeness for AGN. These are also reported in Table 2.4. We use these limits as thresholds and consider all the objects brighter than such values.
Secondly, one needs to account for contamination given by objects that should not be in the catalog of interest. For instance, contamination in a cluster catalog is given by bright AGN that are mistakenly classified as extended sources. This is measured by the fraction of entries in the extended source catalog that are assigned to a simulated AGN or star. For an AGN catalog instead, contamination is caused by faint and or cool core clusters that are erroneously detected as point sources.
Finally, it is important to consider the false detections, that are entries in the source catalog related to a random background fluctuation, not to a physical source. This causes a fraction of spurious sources in the eSASS catalog. Contamination and false detection rate are usually enclosed in the notion of purity. The purer a catalog, the fewer contaminants and spurious sources it contains.
We combine these aspects in a single concept: the detection efficiency, which encodes the completeness and purity of the source catalog. We measure completeness, contamination, and the fraction of spurious sources in the eRASS1 simulation for different intervals of exposure time, defined in Table 2.4. In addition, we account for different thresholds of the detection and extension likelihood to cut the catalogs and study how they impact the eSASS performance in terms of detection efficiency for AGN identified as point-like and clusters of galaxies characterized as extended. We report our results in the next paragraphs.

AGN

Increasing exposure time allows detecting a fixed fraction of sources down to lower fluxes. In the full catalog with DET_LIKE > 5, we measure similar values of the completeness fraction in distinct exposure bins, thanks to the choices of different flux limits. The values are larger than 90% for eRASS1. These numbers will depend on the given flux limit. Our choice of the value where the fraction of detected AGN is equal to 0.8 leads to measuring a higher completeness fraction when using such values as thresholds. In general, we measure a lower fraction of spurious sources in areas with larger exposure. This is because even though the total number of background photons is higher, their fluctuations are suppressed, which results in a lower false detection rate. In the shallow areas, about 32% of the full source catalog does not have a simulated counterpart and is classified as spurious. This number is reduced to 13% in regions around the southern ecliptic pole with the deepest exposure. We provide an analytical fit for the false detection rate as a function of DET_LIKE cuts and exposure time in Sect. 2.8 (see Equation 2.10 and Table 2.6). Progressive cuts in detection likelihood clean the source catalog from these false detections but reduce the fraction of simulated AGN that are detected. Given our choices of flux thresholds, the completeness drops from 94.4% (94.9%) to 82.4% (86.5%) from DET_LIKE > 5 to DET_LIKE > 10

in the shallow (polar) region of eRASS1.
The fraction of clusters that leak into the point source sample is around 4%. A higher detection likelihood cut of 20 reduces this contamination to 2.2% in shallow areas and 1.5% in the pole, but this results in a significant loss in terms of completeness, which respectively drops to 50% and 63%. At fixed completeness, we measure higher contamination in areas with lower depth. This means that a larger exposure time is key to properly distinguish AGN detected as point sources from clusters contaminating the point-like sample, that should be classified as extended ones instead. All these trends are clear in the top panels of Fig. 2.10. The left-hand one shows the fraction of detected AGN as a function of the false detection rate. The lines are color-coded by exposure time and the dots and numbers denote different cuts in detection likelihood. The right-hand panel displays the correlation between AGN completeness and the fraction of clusters wrongly detected as point-like objects.

Clusters

We perform a similar analysis for the population of clusters in the source catalog at different cuts of extension likelihood. The choice of flux limits corresponding to the 50% completeness in each exposure interval translates into completeness values of around 65% when using them as thresholds. The qualitative efficiency trends are slightly different from the AGN ones. For instance, we progressively measure a lower false detection rate from regions with exposure lower than 110 seconds (1.5%) to the ones covered by 150s – 400s (0.6%). However, the behavior of the spurious fraction in the pole region is different. In fact, it increases from extension likelihood larger than 6 to 8 and then drops as expected. This trend is related to the close interplay between the removal of false detections and bright AGN that leak into the extent-selected sample. Progressive cuts in extension likelihood are very effective for the latter case so that the number of spurious sources with respect to the total increases from EXT_LIKE $>$ 6 to EXT_LIKE $>$ 8, but it still decreases with respect to the number of real sources. Increasing to EXT_LIKE $>$ 10 brings the false detection rate to 0.5% in the shallow areas and 0.1% in deep regions. Together with rejecting spurious sources, increasing EXT_LIKE thresholds are very effective in reducing the fraction of contamination. The latter goes from about 30% for EXT_LIKE $>$ 6 in the pole region of eRASS1 to 4% at EXT_LIKE $>$ 20. In regions covered by the average depth of the survey, with exposure 150 s – 400 s for eRASS1, contamination goes from 23% to 1% for the same extension likelihood cuts of 6 and 20. Deep regions suffer from contamination more than shallower ones. This value goes from about 15% in shallow areas to 32% in the pole region for eRASS1. This is due to the larger amount of bright AGN photons that can be mistaken for extended objects, but also due to the higher chance of merging nearby point sources into a single extended detection.
In eRASS1, we measure similar contamination of about 15% on shallow areas ($<$ 110 s) cutting the catalog with EXT_LIKE $>$ 6 and the pole region ($>$ 400 s) with EXT_LIKE $>$ 10. The completeness is also close to 60% with these cuts. The average exposure of eRASS1 corresponding to roughly $\sim$ 275 s is included in the green curves. Cutting the catalogs at

EXT_LIKE > 20 provides about 50% of the simulated cluster above the chosen flux thresholds, with 2% contamination, and a null false detection rate in eRASS1.
The completeness-spurious fraction and contamination curves for clusters are shown in the bottom panels of Fig. 2.10. The figure is color-coded according to the exposure time. It highlights the results described above. Progressive EXT_LIKE cuts of 6, 8, 10, 15, 20, and 30 are written as text.
These are key steps toward selecting a sample of clusters of galaxies to measure cosmological parameters with eROSITA, which has to be as pure and complete as possible. We discuss an alternative way of characterizing clusters of galaxies using the maximum signal-to-noise radius (srcRAD) in Sect. 2.9.

2.3.5 Sensitivity

We compute the sensitivity maps for point sources in each simulated tile, using the `apetool` task, part of the eSASS chain described in Sect. 2.2. The sensitivity map is related to the probability of identifying a detection in a given energy band and at a given position on the detector. Our sensitivity maps are given in units of counts. These values depend on the Poisson false detection probability, which is defined as the probability of detecting photons generated by a random background fluctuation inside a radius of a given value as a source. We set this threshold to a standard value of $P = 4\times10^{-6}$, which corresponds to DET_LIKE$\sim$12 (see Equation 2.4). We consider apertures enclosing a local PSF encircled energy fraction equal to 60% (Brunner et al., 2022). Given the definition of detection likelihood (Sect. 2.2), these two quantities are related by DET_LIKE = -ln(P). The final sensitivity map depends on the estimated background map, the detection mask, and the exposure map. Additional details are provided by Georgakakis et al. (2008). For each simulated tile in the simulation, we obtain the lower count rate detection threshold by dividing sensitivity and exposure maps. We convert to flux by dividing the count rate by the energy conversion factor (ECF) in the soft X-ray band between 0.2 and 2.3 keV. The ECF is computed following Brunner et al. (2022), with an absorbed power-law model of slope equal to 2.0 and varying galactic absorbing column density (nH) equal to the average value in each tile. It is equal to 1.074×10^{12} cm^2/erg for an nH value of 3×10^{20} cm^{-2}. The result is the survey flux limit in areas of the sky covered with different exposure. We compute the cumulative distribution function of the flux limit for each tile and normalize it by the unique area covered by the sensitivity map. We sum up such quantity for all the simulated tiles. The result is the Area covered by the simulated first eROSITA all-sky survey as a function of limiting flux.

We show the normalized survey area in figure 2.11. It displays the area curve for the eRASS1 simulation and a comparison to the eRASS:8 sensitivity prediction from Merloni et al. (2012). The dashed orange line is an extrapolation of the eRASS:8 prediction to the depth of eRASS1, obtained by re-scaling the curve to the predicted eRASS1 limiting flux (see Table 4.4.1 in Merloni et al., 2012), and multiplying by an additional factor of 1.403, converting the flux of an ideal absorbed power-law AGN model with $N_H = 3\times10^{20}$ cm^{-2} and photon index $\Gamma = 1.8$ from the 0.5–2.0 keV band to the 0.2–2.3 keV one. The

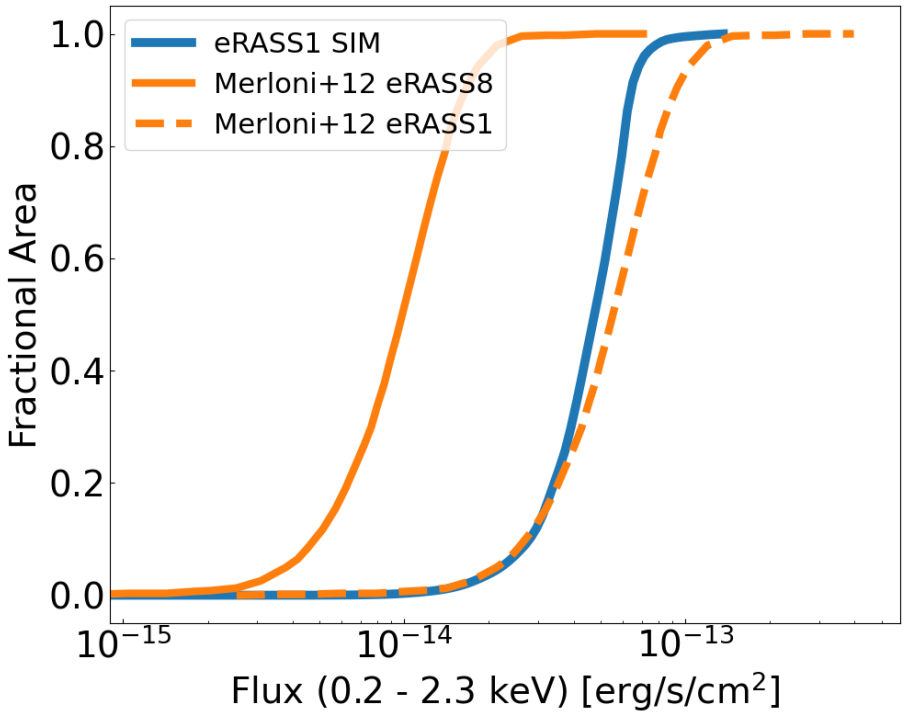

Abbildung 2.11: Simulated eROSITA fractional survey area as a function of flux limit. The eRASS1 simulation is denoted by the blue line, the prediction by Merloni et al. (2012) for eRASS:8 is shown in orange. The dashed line denotes an extrapolation of the eRASS:8 prediction to the depth of eRASS1.

agreement between the prediction and our measurement at the faint end is excellent. An offset at larger fluxes is expected because the former is based on an analytical derivation of the Poisson probability for false detections (see Sect. 4.3.1 of Merloni et al., 2012). A number of assumptions are taken into account to compute the final sensitivity, related to the background, the foreground absorption, and the exposure. Moreover, a further contribution is given by different Poisson probability thresholds, equal to 3×10^{-7} against 4×10^{-6} for this work. Our measure using the sensitivity maps computed by eSASS additionally accounts for a diverse and more realistic treatment of the X-ray background, as well as the true exposure derived from the real eRASS1 scanning process. This allows accounting for the nonuniform depth of the survey, compared to the prediction that is carried out at a fixed exposure equal to the average value for eRASS1 and does not include the higher sensitivity in deeper regions. This causes the difference between our measure and the prediction at the bright flux end. We find that 50% of the area is covered with a flux limit of 4.7×10^{-14} erg/s/cm^2 in the 0.2–2.3 keV band in the eRASS1 simulation.

2.3.6 Imaging and spectral analysis

We measure temperature and X-ray luminosity for a subsample of randomly selected clusters that have been detected as extended by eSASS. Our approach follows Ghirardini et al. (2021b) and is described in Sect. 2.2.3. This sample spans a wide range of exposure times, from equatorial shallow regions to deeper ones close to the southern ecliptic pole. It consists of 873 objects. In order to test our measurements, we compute a weighted mean of the measured luminosity in input luminosity bins with 0.1 dex width. We use weights that are equal to the inverse of the uncertainty on the value of measured X-ray luminosity. The result is shown in Fig. 2.12. The blue shaded area shows the average value of the recovered X-ray luminosity, enclosing the L_X standard deviation in each bin. This is always compatible with a perfect one-to-one relation, shown by the orange dashed line, and with a linear fit[5] in the form of $\log_{10}L_{X,M} = m\ \log_{10}L_{X,SIM} + q$, denoted by the green dashed line. The lower panel displays the residual plot.
We notice that the residuals of the linear fit slightly shifts from negative to positive values for increasing luminosity. The slope of the linear relation is $m = 1.026\pm0.001$. First of all, the fit of the density profile for faint clusters is more challenging, because they provide a lower amount of counts than bright objects. In addition, the temperature spectral fitting also requires a larger amount of photons to be precise. For fainter objects, this makes the conversion factor between count rate and luminosity more uncertain. This effect is partially mitigated by the fact that system with lower luminosity show also on average a lower gas temperature. These systems have more emission lines and a bremsstrahlung cutoff at a photon energy with high effective area, which partially reduces the number of counts needed to ultimately measure the temperature. The combination of these two factors biases the recovered X-ray luminosity toward lower values. The scatter slightly shifts toward positive values for luminous clusters. Bright structures are more probable to be extended on

[5] `https://scipy.org/` (Virtanen et al., 2020)

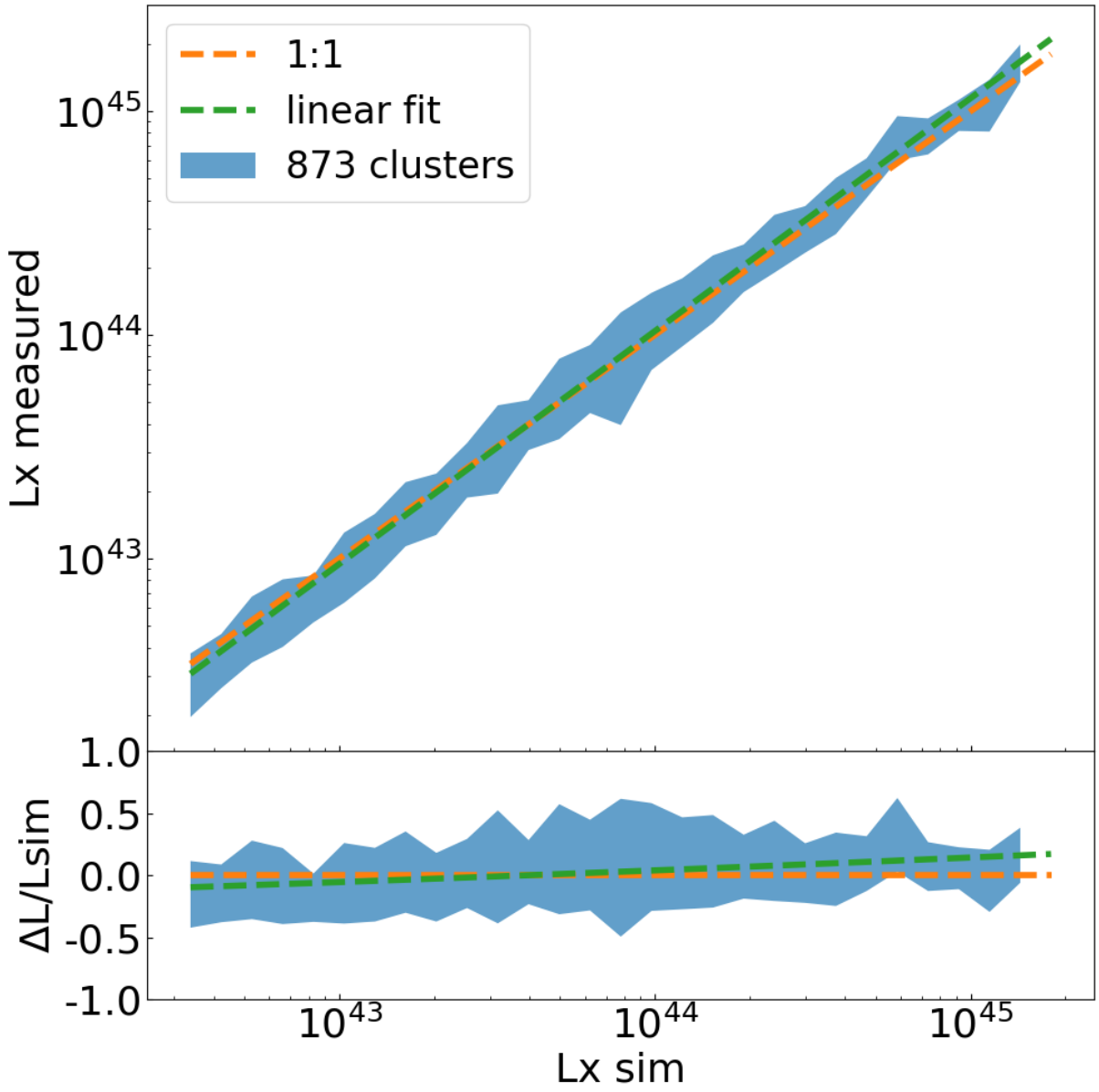

Abbildung 2.12: Measure of X-ray luminosity. **Top panel**: Comparison between average values of measured X-ray luminosity as a function of input ones. The blue shaded area encloses the average measured luminosity within 1σ uncertainties. The dashed orange line shows a perfect one-to-one relation. **Lower panel**: Residual plot normalized by the input luminosity.

the sky, which increases the total net count of events that are not generated by the cluster itself but are potentially considered in the surface brightness fitting. This happens if point sources or other extended sources are not properly masked or if the background is not perfectly modeled. An additional component of the scatter in the relation between simulated and measured X-ray luminosity is given by bright nearby clusters. For these objects, the background extraction region is very large and can span areas with nH fluctuations. This may bias the temperature measure and ultimately the X-ray luminosity. These are all minor effects that do not affect our results on average. We find an excellent agreement between the input luminosity values and the measured ones.

2.4 Discussion

In the following section, we further discuss the importance of a proper characterization of source samples in the eROSITA surveys, and different strategies to build cluster samples for eRASS1.

2.4.1 Biases of the survey sample properties

Understanding the properties of large samples of sources from surveys such as eROSITA is crucial to exploit their scientific potential to the fullest. Accurate and precise detection and classification of sources in astronomical surveys is, therefore, an essential task. A multitude of factors make the process complex: the nature of the sources themselves, the characteristic of the telescope, and the detection pipeline. In general, it is important to understand and quantify the causes of errors and misclassification.
For instance, fluctuations of the X-ray background are potentially detected and classified as a source by eSASS. In this context, a biased measure of the X-ray background impacts not only the number of false detections, but also the detection likelihood of identified sources, because photons emitted from a source might be mistaken for background photons, or vice versa. An accurate estimate of the false detection rate is crucial to assess the fraction of spurious sources in a given sample. We showed that this can be achieved with realistic end-to-end simulations, identifying entries in the source catalog that are not matched to a simulated counterpart.

Another key factor is the contamination in the extent-selected sample (see Fig. 2.3). It is important to figure out why it occurs and how it can be reduced. Contamination is caused by different aspects. The main contribution is given by bright point sources, that are classified as extended. In the cleaned eSASS catalog of our simulation, 1017 extended detections are assigned to a simulated AGN, about 18% of the total extent-selected catalog. Secondary effects include close pairs of bright AGN, that can mimic the emission of an extended object when the detection algorithm is not able to resolve and disentangle the point sources. 446 among the 1017 AGN detected as extended are contaminated by another point source in our simulation (see Sect. 2.2.2). In addition, bright nearby stars can appear extended on the sky, further contaminating the cluster sample. 178 extended

detections are assigned to a star in the cleaned catalog of our simulation. Areas around bright known stars from the optical band can easily be masked in the real survey, which minimize the contamination due to stars. This effect is even magnified in areas with deep exposure, where random background fluctuations have a higher chance of being identified as an extended source. We find a total of 48 extended false detections.

Such cases end up in secure extended detections with large values of DET_LIKE, which explains why choosing extremely high thresholds of detection likelihood do not lower contamination. Instead, a cut in EXT_LIKE is needed to reduce the contamination due to point sources in the extent-selected catalog. This is clear in the bottom panel of Fig. 2.3. It is possible to argue that a direct comparison of the population in the catalog after cutting at the same value of extension or detection likelihood is misleading, due to the intrinsic difference between them. In particular, EXT_LIKE has typically smaller values than DET_LIKE for extended detections. In the extent-selected sample, the 0.25, 0.5, 0.75, and 0.90 quantiles for EXT_LIKE (DET_LIKE) are 7.95 (22.69), 12.05 (37.19), 22.85 (65.22), 41.35 (114.99). For instance, if we focus on the 0.5 quantile, the AGN contamination is equal to 4.6% for EXT_LIKE = 12.05 and to 18.2% at DET_LIKE = 37.19. Therefore, we still conclude that applying extension likelihood cuts is a more efficient way of decreasing contamination. In observations, this can also be solved by a multiband approach, for example doing an optical follow-up of extended X-ray sources (see Salvato et al., 2022, for an example). This allows keeping all the cluster candidates in the catalog, that has the highest possible completeness. In a second step, one can look for overdensities of red galaxies around each X-ray detected cluster. If there is evidence of a red sequence, the cluster will be confirmed (see Finoguenov et al., 2020, for an example). Otherwise, the catalog will be cleaned from a spurious or contaminating source, increasing the purity of the sample, while keeping the completeness level unchanged. This is a key ingredient toward precision cosmology with X-ray-selected clusters (Ider Chitham et al., 2020).

With optical follow-up observations, one can not only find contaminating point sources classified as extended but also identify real clusters of galaxies that are misclassified as point sources (Green et al., 2016; Bulbul et al., 2022). Understanding why extended sources are classified as point-like ones is key to correct this bias and properly characterize as many clusters as possible. A cluster ends up classified as a point source because of different reasons. The first one is brightness. These are usually faint objects, whose extended emission at the outskirts struggles to emerge over the local background. The second one is related to their cores. Clusters with a peaked emission in the center are possibly mistaken for point sources. In fact, we find that clusters with low flux and cool core are detected as point sources (see Fig. 2.9). Furthermore, high redshift clusters, even if intrinsically bright and extended, cover a tiny area on the sky, possibly smaller than the PSF of the telescope. Finally, clusters of galaxies hosting an AGN are potentially dominated by the emission of the latter. All these cases give rise to contamination and or misclassification for clusters of galaxies that leak into the point source sample. A purer cluster sample affected by less systematics may be obtained by a detection algorithm that excises the core region. This is because the cluster's outskirts have been shown to evolve in a self-similar way, with low scatter (McDonald et al., 2017; Käfer et al., 2019; Ghirardini et al., 2019). A more direct

definition of the sample in terms of cluster mass is therefore achievable this way. This idea was implemented in clusters studies by Vikhlinin et al. (1998). A recent implementation is described by Käfer et al. (2020), where the X-ray images are filtered by a series of spatial wavelet filters with different scales, which allows isolating the extended emission from galaxy clusters. However, such a method requires a larger amount of counts to detect a cluster, which lowers the completeness of the sample.
The misclassification and contamination of clusters are additionally relevant for AGN. Simply selecting AGN from the point-like catalog means missing the bright objects that are mistakenly classified as extended, and accounting for faint clusters contaminating the point source catalog. However, we showed that this can be addressed by estimating completeness and contamination from realistic simulations, which provides the fraction of false detections, the contaminants, and the sources missed by the detection scheme according to desired selection criteria.

Completeness purity trade-off

Perfectly complete and pure samples of sources are ideally desired for astrophysical and cosmological studies. This means that, above the flux limit for a given experiment, a perfectly efficient detection and selection scheme should provide all the physical sources, making the source catalog as complete as technically feasible, identifying also very faint objects. Depending on the scientific application, it should also produce a catalog containing only the sources of interest, making the sample as pure and clean as possible. This means minimizing the rate of false detection: background fluctuations that are classified as physical sources. The concept of purity also includes contamination. For instance, in the extent-selected sample contamination is caused by bright AGN or stars, which should be classified as point sources instead. The number of such objects should also be minimized. The concepts of completeness and purity are closely related: maximizing the first means pushing the limits of the algorithm, and trying to identify the faintest physical objects. These are easily mistaken for random background fluctuations, which ultimately ends up costing a higher fraction of spurious sources in the final catalog.
In the context of the eROSITA surveys, the trade-off between completeness and purity is affected by various parameter choices made to select clusters. Different extension likelihood cuts are an example. Choosing a very low threshold will keep the catalog complete on the one hand, but on the other, the risk of introducing AGN in the catalog is higher, which increases contamination. Instead, higher likelihood thresholds will give a cleaner sample, at the cost of reducing the fraction of detected objects. This is evident in Fig. 2.10, where progressive EXT_LIKE cuts degrade completeness, but improve purity, reducing the false detections and contamination. We stress that our choice of various flux limits for different exposure times guarantees a comparable benchmark between areas covered by varying depth.
Different choices regarding the parameters characterizing the source catalog should be taken according to the specific scientific goal. For example, if the goal is to work with a secure catalog from the start, higher thresholds should be chosen. This will minimize the

spurious sources and the contamination, making such cluster sample pure. However, the completeness will also be reduced. Instead, if the goal is to select the highest possible number of clusters at first, a very low EXT_LIKE limit is best. Such studies may involve the evolution of the luminosity function. A secondary step might then be required to clean the catalog, for example with multiwavelength observations such as an optical follow-up, allowing the confirmation of cluster candidates if there is evidence for a galaxy red sequence. This approach allows reducing contamination thanks to the multiwavelength information, while keeping the completeness level high, because no additional X-ray selection is applied. It also makes the cluster sample more secure, because it probes two distinct properties: the intra-cluster medium through X-rays, and the galaxy members in the optical and infrared bands. Samples defined in this way are particularly suitable for cosmological experiments. In this context, it is important to model the contamination and completeness levels together. For example, Aguena and Lima (2018) quantified the bias on the measure of cosmological parameters due to the imperfect modeling of completeness and purity in a cluster count experiment. They assumed a DES-like survey and found that a proper description of completeness and purity is key to measure unbiased cosmological parameters without degrading the constraining power especially when including low-mass clusters. A detailed description of the cluster selection (see Fig. 2.10) is therefore essential, since eROSITA will discover many new low-mass clusters and groups. Finally, other studies such as clustering may require a sample of objects contained in a well-defined volume. These can be constructed by rejecting faint and distant sources (see Sect. 2.4.2).

Impact of source size and cool core bias

Given the morphological complexity of clusters of galaxies, their detection is not a simple function of flux and exposure time.
For example, it has been shown that the size of the cluster on the sky does have an impact on the identification (Pacaud et al., 2006; Burenin et al., 2007; Clerc et al., 2018). On the one hand, the detection algorithm can easily detect bright nearby clusters and characterize them as extended. On the other hand, high redshift clusters, even if bright and large, may cover an area on the sky that is close to or smaller than the telescope PSF. The same holds for nearby groups with very low mass. Such objects are easily mistaken for point sources in the detection process. This makes the detection of clusters more complex. In Fig. 2.8, we show that the fraction of clusters detected as extended is not only a function of flux and exposure, but it additionally depends on the size of R_{500c} on the sky, even fixing the former two variables. This effect is more visible for clusters with a smaller radius, whose extended emission struggles to emerge over the background, compared to larger clusters with a similar flux. These objects are actually detected by eSASS, but classified as point sources, as expected.
Furthermore, the dynamical state of the clusters plays a role in the detection and classification. Dynamically relaxed structures had time to develop an efficient cooling toward the central regions, which enhances their central X-ray emission, resulting in a peaked surface brightness profile. This makes it easier for these types of objects to emerge over the back-

Tabelle 2.5: Number of clusters in the volume-limited and flux-limited samples for areas covered with different depth.

	Number of clusters			
Exposure [s]	Volume-limited		Flux-limited	
	DET	SIM	DET	SIM
> 0	262	282	734	893
> 110	349	392	829	1044
> 150	361	414	752	992
> 400	80	100	146	200

ground and biases the detection toward them. This is the notion of cool core bias (Eckert et al., 2011). However, clusters with a peaked profile can resemble the emission from a point source. In such cases, the peaked emission toward the central regions dominates over the tail at larger radii. This is not easily identified by eSASS, which ends up classifying the cluster as a point source. The net effect is that the detection is biased toward cool core clusters, but they might be easily misclassified as point sources.
The link between this effect, the exposure time, and the background has a significant impact on the detection process. On the one hand, a large exposure for a cool core cluster makes the large ratio between photons from the core and photons from the outskirts more clear over the background, making them look more similar to point sources than analogous objects covered by a shallow exposure. This will increase the probability to misclassify such clusters as point-like objects. On the other hand, with increasing depth the signal-to-noise ratio of the cluster outskirts will increase relative to the local background. In principle, a more accurate estimate of the background is also possible in this regime, thanks to the lower variability. These aspects should instead help the identification of clusters as extended. The characterization of the cool-core fraction in a cluster population also depends on the selection of the sample. Ghirardini et al. (2021a) measured the dynamical state of eFEDS clusters combining a set of quantities (such as concentration, central density, photons asymmetry, ellipticity) and did not find a prominent cool-core bias on the extent-selected sample. More detailed simulations at deeper exposures (e.g., eRASS:8) are needed to investigate this topic. Nonetheless, most of the brightest clusters are properly identified as extended (Fig. 2.9). We conclude that the eSASS algorithm minimizes the impact of the cool core bias on the vast majority of the sample of clusters detected as extended sources. It mostly affects the low-flux clusters, where only the cool cores are detected, but classified as point sources.

2.4.2 Construction of volume-limited samples

In the context of cosmological experiments, a well-defined sample of galaxy clusters is crucial. The eROSITA all-sky surveys naturally produce samples that are mainly flux-limited. Such samples are made up of clusters that reach the survey flux limit, which mainly

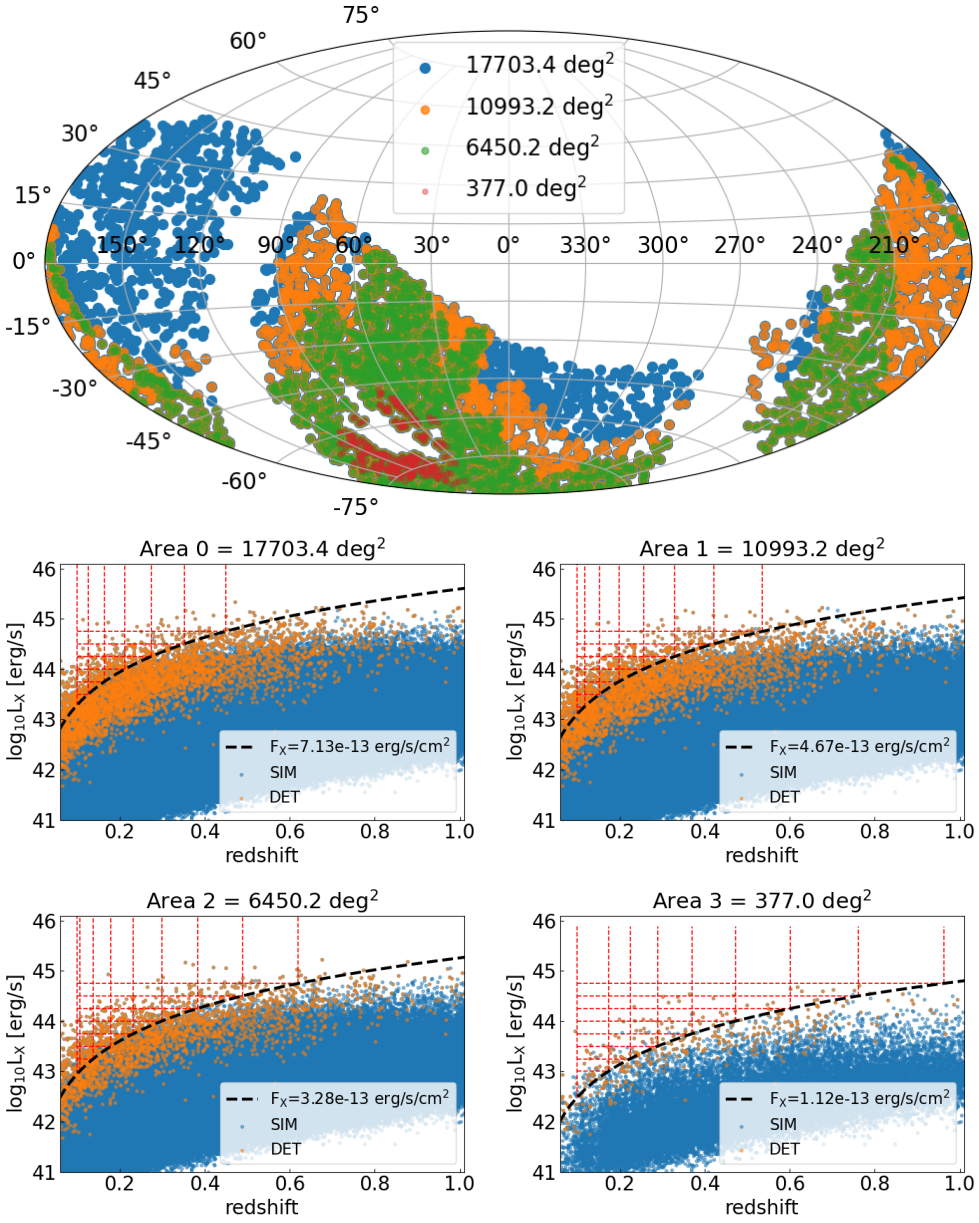

Abbildung 2.13: Selection of a volume-limited cluster sample in the eRASS1 simulation. **Top panel**: sky map with the cluster population in areas covered by different depth. Areas 0, 1, 2, and 3 respectively cover regions with exposure larger than 0 s, 110 s, 150 s, and 400 s. They are cumulative areas with respect to the ones defined in Table 2.4. **Bottom panels**: population of simulated and detected clusters in the luminosity–redshift plane. The black dashed lines denote the chosen flux threshold at each depth (see Table 2.4). The red dashed lines locate different areas above the given flux limits. The volume-limited sample is constructed with the objects within the regions delimited by these lines.

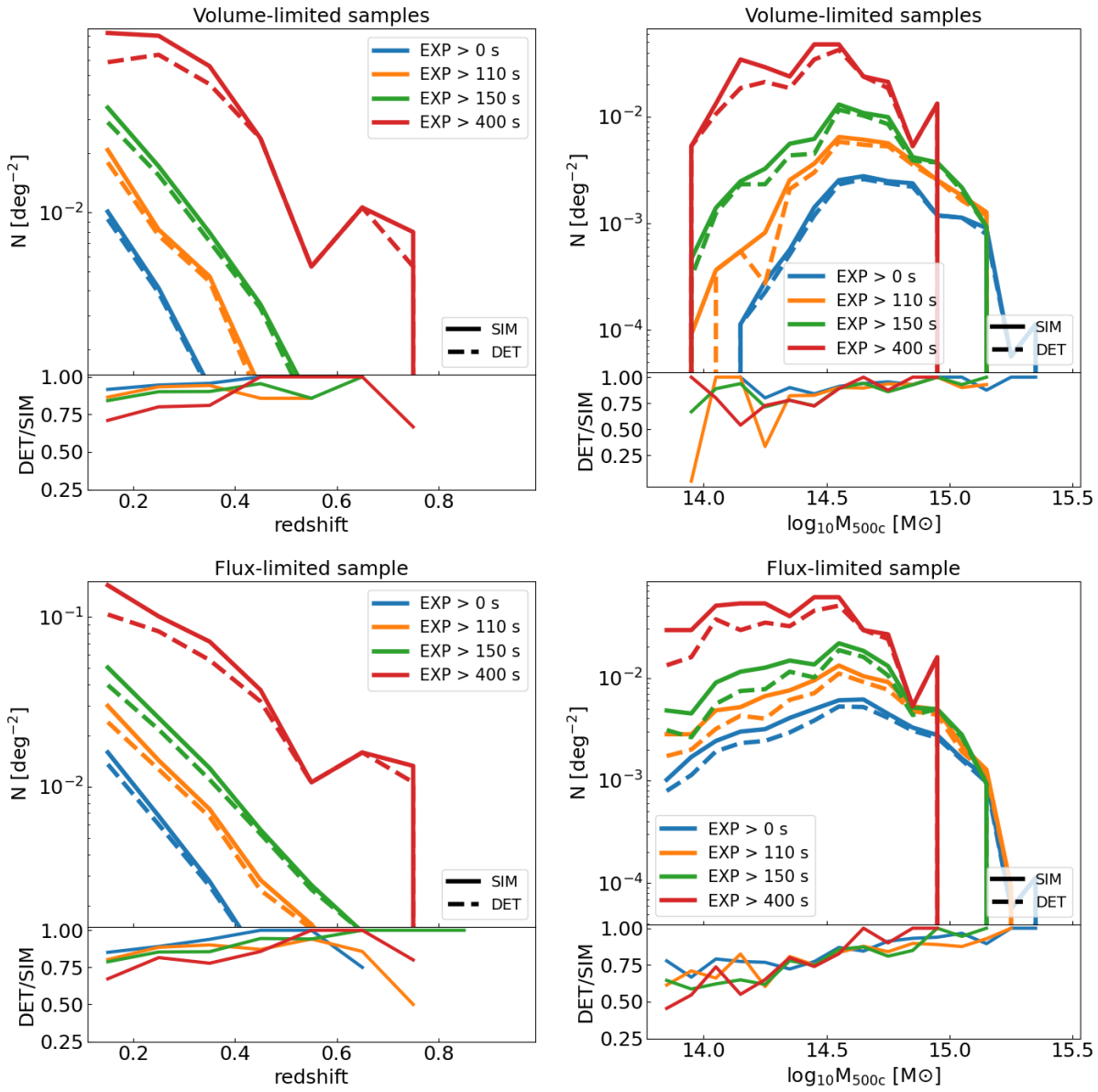

Abbildung 2.14: Comparison between the volume-limited and flux-limited samples built with clusters detected as extended and simulated ones. The top panels display the volume-limited samples, the bottom panels show the flux-limited ones. **Left-hand panels**: relative contribution to the total cluster number density as a function of redshift for the four different populations. The lower plot shows the ratio between the N(z) built with the samples of detected and simulated clusters. **Right-hand panels**: relative contribution to the total cluster number density as a function of mass for the four different populations. The lower plot shows the ratio between the N(z) built with the samples of detected and simulated clusters.

depends on the telescope sensitivity and the scanning strategy. Therefore, a higher number density of objects is detected at low redshift and luminosity, compared to the high-z regime where only the brightest sources are detected. In addition, given that the sky coverage by eROSITA is not uniform in terms of depth, fainter objects can be detected in deep areas with a higher probability. Therefore, the source catalog will have different properties in regions with different exposure. These differences can be mitigated by building volume-limited samples, made up of clusters with a given set of properties inside a well-identified volume, that is encoded in the choice of the maximum distance (or redshift) up to which one is interested to build the sample. Then, in order to obtain an unbiased sample of objects, one can only consider sources whose luminosity is larger than the value corresponding to the flux limit at the chosen redshift. This provides a sample with a constant distribution of the number density as a function of redshift. Practically, a volume-limited sample is built from a flux-limited one by getting rid of the sources that are less luminous than a given threshold and further away than a certain distance (or redshift). The relation between these values of luminosity and redshift is set by the flux limit of the survey (see Eq.2.6), where $d_{\rm L}$ denotes the luminosity distance, that depends on redshift:

$$F = \frac{L}{4\pi d_{\rm L}^2(z)}. \tag{2.6}$$

Within such ranges of luminosity and redshift, a selection function built from simulations is less uncertain and allows unbiased results in cosmology experiments. A volume-limited sample provides an even sampling of the large-scale structure, accounting for the observational limits of the survey.

Because eROSITA does not cover the sky with a uniform depth, we build different volume-limited samples applying higher flux limits in areas with lower exposure. We use z = 0.1 as the lower redshift boundary and consider the exposure intervals and flux limits corresponding to 50% completeness defined in Table 2.4. We account for the K-correction in the relation between flux and luminosity (Equation 2.6). It guarantees that the flux is always measured in the same energy band for clusters at different redshift. We start by considering all the clean extragalactic eROSITA_DE sky (see Sect. 2.2.2) and applying the largest flux limit, identifying the 50% completeness in the shallowest areas of the survey. We proceed by reducing the area, gradually excluding shallower regions, and applying deeper flux cuts. The result is shown in Fig. 2.13. The first panel shows the cluster population in areas covered by a different exposure, that have been used to construct the volume-limited samples. The other four panels show how simulated (in blue) and detected clusters (in orange) populate the luminosity–redshift plane. The black dashed lines denote deeper flux limits as the area shrinks. The red dashed lines locate regions of this plane where the clusters are luminous enough to be above the flux limit at a given redshift. The volume-limited samples are finally built by considering the clusters inside the areas identified by these lines. Table 2.5 reports the number of clusters in the volume-limited samples and the corresponding flux-limited ones.

This method provides a collection of clusters detected as extended sources that are an even subsample of all the simulated clusters within the same ranges of luminosity and redshift.

This is shown in Fig. 2.14. The two top panels show how the volume-limited samples built with clusters detected as extended (dashed lines) compare to the one made up of simulated clusters (solid lines). The plot is color-coded according to the exposure time. The relation between the two samples in terms of redshift and M_{500c} is shown in the right-hand and left-hand panels, respectively. The bottom panels show the corresponding flux-limited samples. The ratio between the number density distribution of detected and simulated clusters as a function of redshift is roughly constant for the volume-limited samples, and the majority of clusters with masses down to $M_{500c} \sim 10^{14} M_{\odot}$ within our selection are detected. This means that our method has the potential to identify a cluster sample that provides an even sampling of the large-scale structure at different redshift and exposure time. The same trends of the cluster number density as a function of mass and redshift for the flux-limited samples are qualitatively similar. For the second one, the completeness is lower compared to the volume-limited case, because the cuts in luminosity and redshift exclude clusters that are close to the flux limit and have a lower probability of being detected. Using the full flux-limited sample to measure cosmological parameters maximizes the number of clusters, but a robust selection function around the upper redshift limit of the survey is required. The advantage for the volume-limited samples is that they are contained in a well-defined cosmological volume. This potentially makes the definition of the survey volume less uncertain in a cluster counts cosmological experiment.

2.5 Extension of the model to galaxy groups

In this section, we provide further details about the extension of our improved cluster model to lower masses (see Sect. 2.1), comparing it to the eFEDS cluster sample. Along with the (Anderson et al., 2015, AN15) correction using stellar mass, we also tested an improvement exploiting the X-ray luminosity - halo mass relation, following Le Brun et al. (2014) (LB14). Such correction reads:

$$\log_{10} L_{\mathrm{x},(0.5-2.0\mathrm{keV})} = 2 \log_{10} M_{500c} + 14.5. \tag{2.7}$$

This correction gives a shallower slope in the cluster logN–logS (see Fig. 2.6) than AN15. AN15 provides a better agreement to observations than LB14, especially at low luminosities $< 1{\times}10^{43}$ erg/s. LB14 underestimates observed values by a factor of ~ 2 at $1{\times}10^{42}$ erg/s. AN15 provides a great correction for the X-ray luminosity to stellar mass relation by construction, while LB14 does not align well with observations. The same holds for the X-ray luminosity to temperature relation. The AN15 version gives excellent agreement to eFEDS data for low luminosity clusters. The L_x–T_x relation obtained from LB14 is too steep. We ultimately choose the AN15 correction over LB14, as it produces a logN–logS and scaling relations that align better with observations (see Fig. 2.15). The prediction of the L_X–M_{500c} relation is slightly underestimated at the high mass end compared to data, see Fig. 2.15. This makes our approach conservative, since the most massive and luminous objects are detected more easily, see also Fig. 2.7. On the other hand, the fact that observations suffer from the Malmquist bias at the low mass end possibly affects our

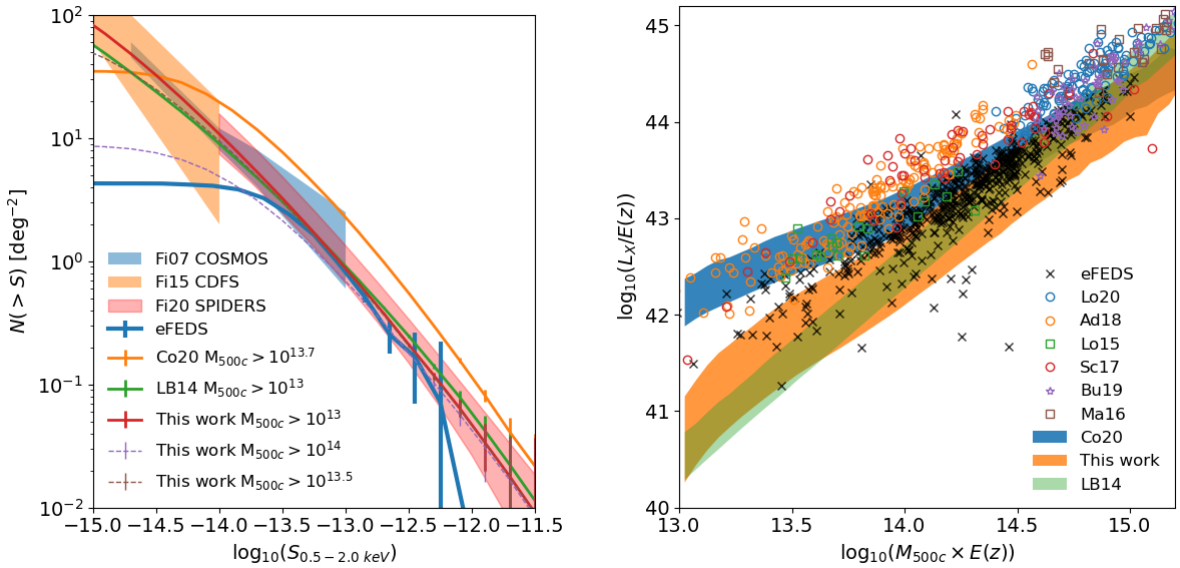

Abbildung 2.15: Improved cluster model. **Left-hand panel**: Number density of sources as function of flux. The solid orange (red) line shows the prediction of the model before (after) applying the correction. The shaded areas in blue, orange, and red denote the logN–logS from Finoguenov et al. (2007, 2015, 2020). The green and blue lines show a comparison to Le Brun et al. (2014) and Liu et al. (2022). The dashed pink and brown lines denote the model corrected for higher mass thresholds. **Right-hand panel**: Relation between X-ray luminosity and mass. The blue (orange) shaded area shows the prediction of the model before (after) applying the correction. The green shaded area denotes the relation from Le Brun et al. (2014). Additional samples are shown by blue circles (Lovisari et al., 2020), orange circles (Adami et al., 2018), green squares (Lovisari et al., 2015), red circles (Schellenberger and Reiprich, 2017a), pink stars (Bulbul et al., 2019), and brown squares (Mantz et al., 2016).

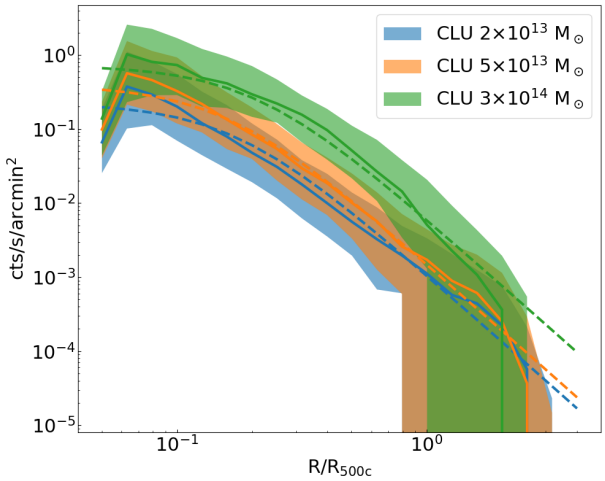

Abbildung 2.16: Surface brightness profiles of the simulated clusters. The radius is normalized to R_{500c}. The solid lines show the average profile, the shaded areas denote the 1σ scatter around the mean. The dashed lines show the best-fitting beta model for each average profile.

correction using AN15. Nonetheless, the addition of the eFEDS cluster sample shows the ability of the model to reproduce observations also in the regime of galaxy groups.
We verify that the shape of the cluster profiles generated with the new model is on average compatible with a beta model. We measure the radial profile of events generated by three samples of 100 simulated clusters with masses of 2×10^{13}, 5×10^{13}, and 3×10^{14} $M_\odot$ as a function of R_{500c}. We fit each one of them with a beta model (see Sect. 2.2). The result is shown in Fig. 2.16. The solid lines show the average surface brightness profile for each one of the three samples, the shaded areas denote the 1σ scatter around the mean value. The dashed lines denote the best-fitting beta model to each average profile. We fix β=2/3, leaving the core radius as a free parameter. This is the same assumption taken by the `ermldet` task (see Sect. 2.2). The agreement between the average profile and the beta models is good. Even if the profile of a single object can significantly deviate from a beta model, our model generates profiles that are on average compatible with the assumptions taken by eSASS in the source detection chain.

2.6 Comparison to data

We compare the source catalog of the eRASS1 simulation to the one obtained by processing the real data with the same eSASS set-up, described in Sect. 2.2. There is good agreement between the mock and the real data. This is shown in Fig. 2.17. The mock is denoted by the blue solid line and the real data by the orange one. The distributions of the photon energy shown in the top panel are in excellent agreement, especially for the soft energy range in our interest. The central and bottom panels show the cumulative distributions of

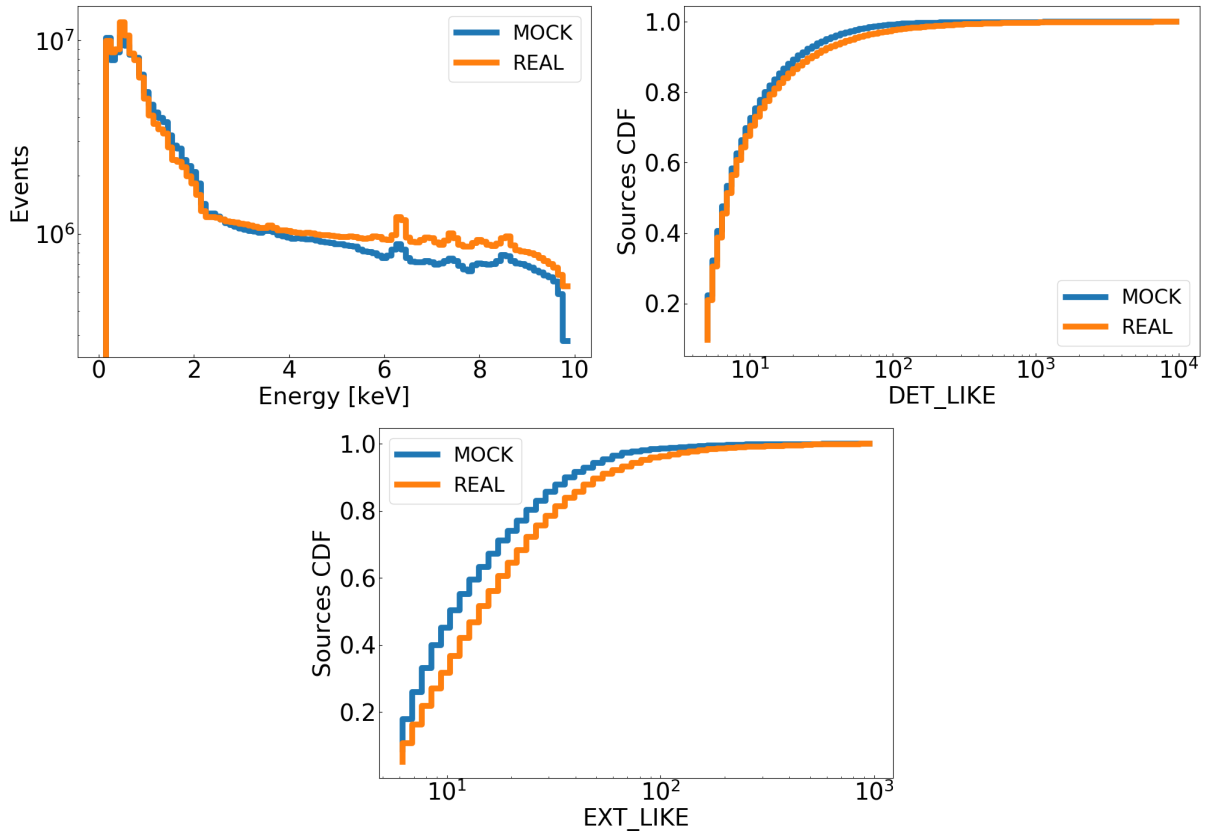

Abbildung 2.17: Comparison between the eRASS1 simulation and the real data. These are respectively denoted by the blue and the orange solid lines. **Left-hand panel**: distribution of the photon energy. **Right-hand panel**: cumulative distribution of the sources as a function of detection likelihood. **Bottom panel**: cumulative distribution of the sources as a function of extension likelihood.

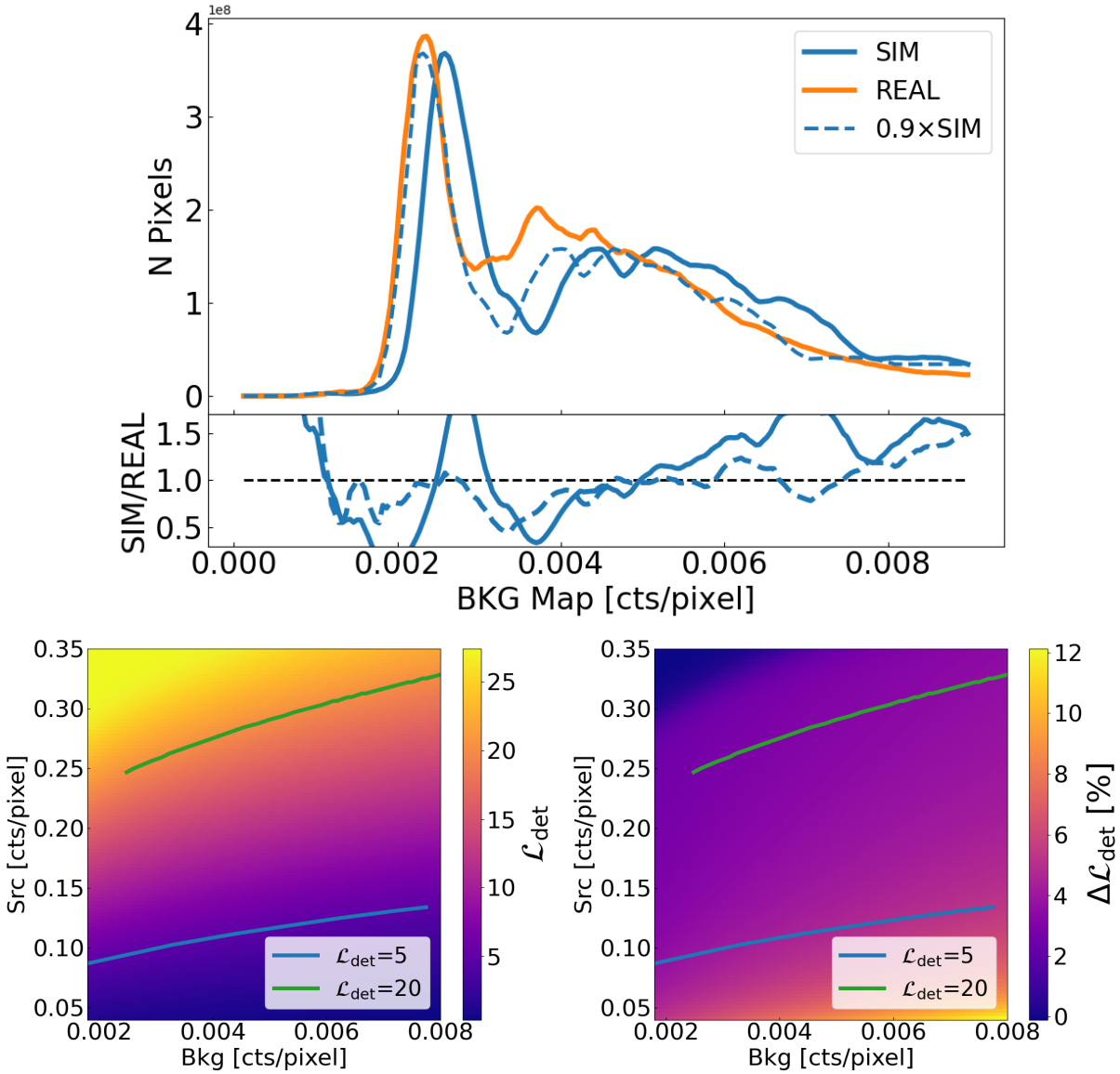

Abbildung 2.18: Background evaluation in the eRASS1 simulation. **Top panel**: Comparison between the mock and real background maps. The lines identify the number of pixels showing a given value of the background map. The mock data is denoted in blue, the real eRASS1 in orange. The dashed blue line shows the simulated background re-scaled by 0.9. The lower panel shows the ratio between the mock and real data. **Bottom panels**: Impact of a 10% overestimation of the background on the analytically computed value of detection likelihood. The left-hand panel shows DET_LIKE as a function of counts in each pixel given by a source and by the background. The panel on the right shows the corresponding percentage error on detection likelihood caused by a 10% larger background. The blue and the green solid lines respectively denote DET_LIKE = 5 and 20.

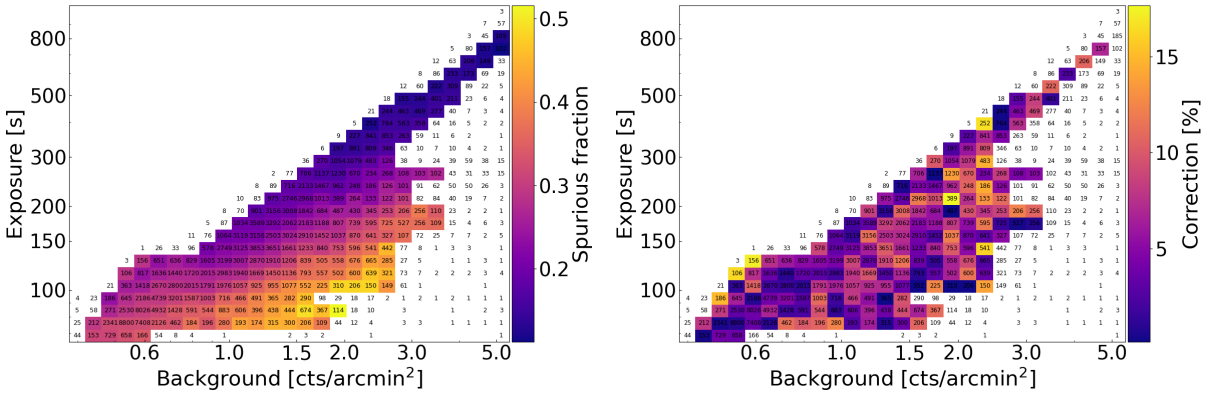

Abbildung 2.19: Evolution of the false detection rate in the eRASS1 simulation. **Left-hand panel**: Spurious fraction as a function exposure time and background level in the eRASS1 simulation. Each bin containing more than 100 sources is color-coded by the false detection rate. **Right-hand panel**: Correction of the prediction of the spurious fraction for the eRASS1 data using the simulation due to the 10% overestimate of the background. The x-axis is binned with a progressive 10% increment. The total number of spurious sources in each bin is written as text.

detection and extension likelihood. A small difference between the two is expected, caused by the fewer number of bright simulated AGN due to the steep logN–logS at high flux (see Fig. 2.5). This also contributes to the difference between the photon energy distributions at the hard end, above 5 keV. Nonetheless, the cumulative distributions show that the mock catalog and the real one have similar properties.

In addition, we compare the background maps measured on the eRASS1 simulation to the ones obtained from real data (see Sect. 2.2). Figure 2.18 shows the total number of pixels with a given value of the background map, expressed in counts per pixel. The real data is identified by the orange line and the simulation by the solid blue one. On the one hand, the peaks of these two curves differ by about 10%. In fact, a re-scaling of the simulated background by a factor of 0.9 (denoted by the dashed blue line) aligns well with the real eRASS1 maps. This is expected because the cosmic X-ray background component is slightly over-estimated in the simulation. The mock data contains the population of faint simulated AGN. However, this contribution is partially present also in the real eRASS1 maps that are used to create the background model. On the other hand, in some areas, the real background is higher than the mock data. This is because the model has been generated using a mean spectrum but in the eRASS1 data some local instabilities cause such higher background level.
In Sect. 2.2 we verified that such overestimation of the background has a negligible impact on the measured values of detection likelihood. This is reported in the bottom panels of Fig. 2.18. We show the value of detection likelihood as a function of source and background counts per pixel (on the left), and the corresponding relative error due to the background

over-estimate (on the right). The relative error is computed as

$$\Delta\mathcal{L}_{\rm det} = \frac{\mathcal{L}_{\rm det,UN} - \mathcal{L}_{\rm det,B}}{\mathcal{L}_{\rm det,UN}}, \tag{2.8}$$

where $\mathcal{L}_{\rm det,UN}$ is the unbiased value of detection likelihood, and $\mathcal{L}_{\rm det,B}$ is the value of detection likelihood biased by a 10% overestimation of the background. The solid lines in blue and green denote values of $\mathcal{L}_{\rm det} = 5$ and 20, respectively. There is a ∼4% impact on the value of detection likelihood for faint sources with DET_LIKE∼5.
Finally, we quantify whether the 10% overestimate of the background significantly impacts our prediction of the false detection rate for the eRASS1 data using the digital twin. For this goal, we measure the spurious fraction on a two-dimensional grid of exposure time and background level. At fixed exposure, we build a binning scheme for the background level such that successive bins are 10% greater than the previous one, according to $X_{i+1} = 1.1 \times X_i$, where X represents the background level bins. The upper panel of Fig. 2.19 shows the spurious fraction in the exposure-background level plane. The grid contains 95% of the real eRASS1 catalog. At fixed background level, the spurious fraction decreases as a function of exposure time. Indeed the deeper data allows suppressing fluctuations of the background. At fixed exposure time, the false detection rate increases as a function of the background level, because the probability of picking up a random fluctuation is larger. This makes our prediction of the false detection rate conservative, because at fixed exposure, the real eRASS1 has a lower background compared to the simulation.
Given our choice of the binning scheme, we can compare successive background level bins at fixed exposure time to estimate a correction for the prediction of the spurious fraction in the eRASS1 data using the simulation. We compute the relative difference between the bins $(f_{\rm spur,i+1}\text{-}f_{\rm spur,i})/f_{spur,i+1}$, where the index i runs on the background level bins for each exposure. The correction for each bin is shown the lower panel of Fig. 2.19. We average over the bins containing more than 100 spurious sources, in order not to be affected by noise. We find a mean correction of 5.7%. We conclude that our measure of the spurious fraction in the digital twin is a conservative prediction of the false detection rate in the real data, and it is not significantly affected by the 10% overestimate of the background.

2.7 Population histograms

In this section, we collect panels showing the histograms and linear fractions relative to the population in the source catalog, described in Sect. 2.3.1 and Fig. 2.3. These are shown in Fig. 2.20. The panels on the left show the total number of sources for different cuts in detection or extension likelihood. The panels on the right show the relative fraction for each source class.

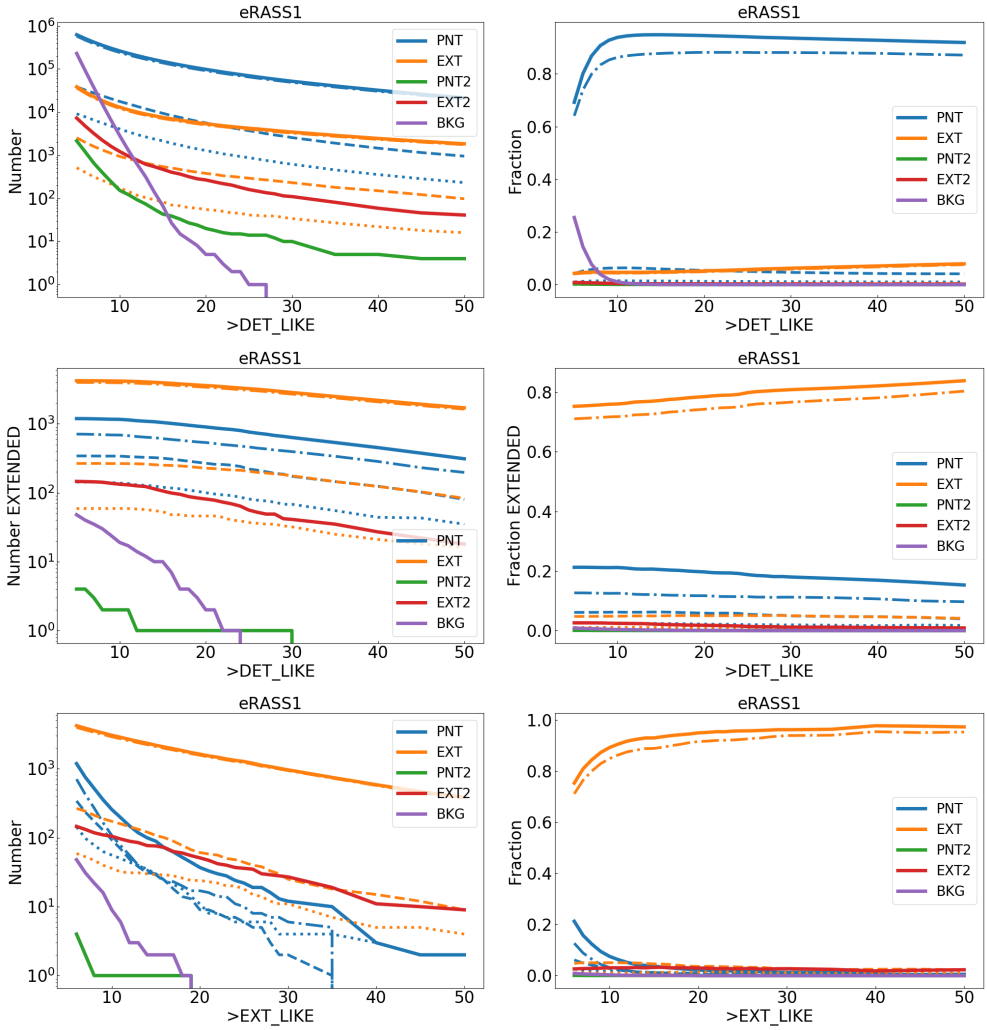

Abbildung 2.20: Population in the detected source catalog. The total number of sources in the cleaned catalog of the eRASS1 simulation is 901 812. The number of extended sources is 5615. **Top panels**: Number of sources in the full catalog and fractions of the population classes in linear scale as a function of minimum detection likelihood. **Central panels**: Number of sources and fractions of the population classes in linear scale in the extent-selected sample (EXT_LIKE >= 6) as a function of minimum detection likelihood. **Bottom panels**: population in the source catalog and fractions of the population classes in linear scale as a function of minimum extension likelihood. Lines of different colors show the classes defined in Sect. 2.2. The dashed-dotted lines denote sources that are not contaminated by photons of a secondary source (no blending), the dashed ones identify sources contaminated by a point source, and the dotted ones show sources blended with a cluster.

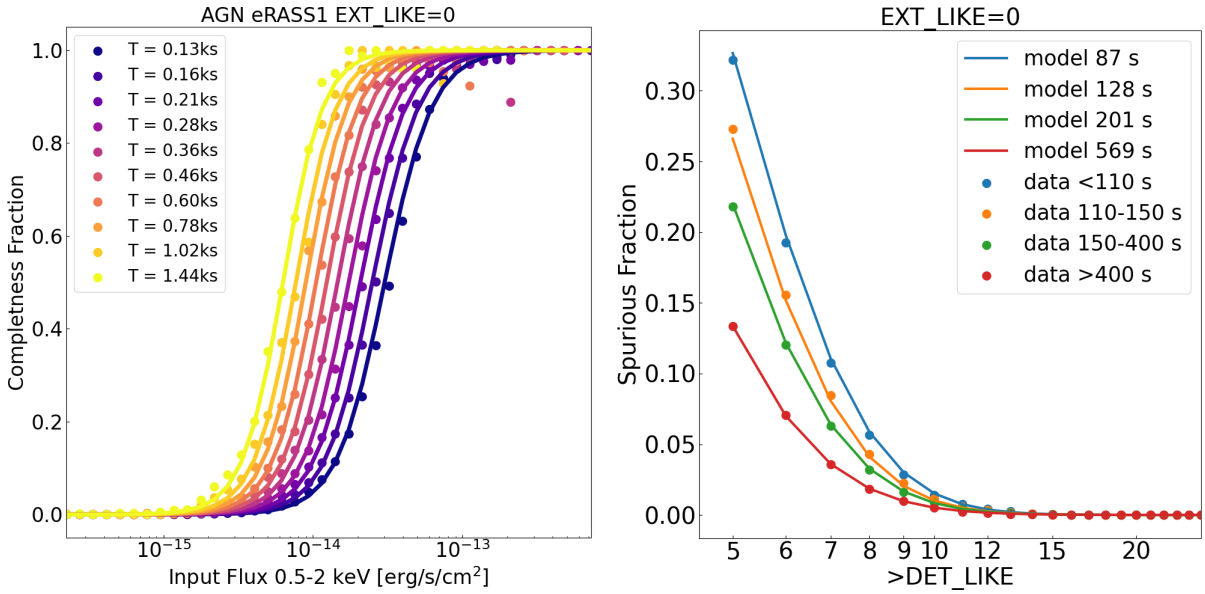

Abbildung 2.21: The point source sample. **Left-hand panel**: fraction of AGN detected as point-like objects as a function of the input flux in the soft X-ray band for different exposure times. The circles show the values measured comparing input and source catalogs, the solid lines our best fit model in Eq.2.9. **Right-hand panel**: fraction of spurious sources in the point source sample as a function of detection likelihood cuts for different exposure times. The full circles denote the false detection rate measured in the simulation, the solid lines identify the model described by Equation 2.10 computed at the average exposure time corresponding to each bin.

2.8 AGN

We provide analytical fits to the completeness fraction of the sample of simulated AGN that are detected as point sources. Similarly to Sect. 3.6, we measure the detected fraction in terms of input flux and exposure time. We model these trends according to a modified sigmoid function. Our model reads:

$$\begin{aligned} b &= q(\log_{10} T)^{\mathrm{w}}, \\ c &= q1(\log_{10} T)^{\mathrm{w1}}, \\ C(F,T) &= \frac{1}{1+10^{-3}\mathrm{e}^{-b\log_{10} F+c}}, \end{aligned} \tag{2.9}$$

where q = 4.59, w = 0.41, q1 = 11.01, w1 = 0.16 for eRASS1. We measure the exposure time T in seconds and the flux F in erg/s/cm^2. We show the result in Fig. 2.9. The values extracted by matching the source catalog and the simulated AGN are identified by circles, the best-fit model is shown by the solid lines, color coded by exposure time. This model is not intended to provide a complete description of the AGN selection function, but it gives a useful benchmark. We notice that it works particularly well for exposure times between 160 s and 600 s for eRASS1, containing most of the eROSITA coverage in terms of observing time.
In addition, we provide a functional form to describe the fraction of false detections for different cuts of detection likelihood and exposure time in the point-source sample. This is described by the following equation:

$$\begin{aligned} A &= a_1 \times T^2 + b_1 \times T + c_1, \\ B &= a_2 \times T^2 + b_2 \times T + c_2, \\ \mathrm{BKG(A,B)} &= \frac{0.85}{(A \times \mathrm{DET\_LIKE}^B + 1)^{4.2}}, \end{aligned} \tag{2.10}$$

where T is the exposure time in seconds, DET_LIKE is a cut in detection likelihood, and the values of the parameters are reported in Table 2.6. Such a model grasps the details of this trend. It is shown in the right-hand panel of Fig. 2.21. The dots denote the false detection rate measured in the simulation in each exposure time interval as a function DET_LIKE threshold, while the solid lines denote the model computed at the average exposure time corresponding to each interval.

Finally, we study the accuracy of the position of AGN detected as point sources (EXT_LIKE=0). We study the offset between the simulated and detected positions and how it relates to the positional error computed by eSASS. Such error is the sum in quadrature of the error on the pixel position multiplied by the pixel scale and is named RADEC_ERR. We find that 99.48% (99.75%) of these point sources are contained by a ratio between the offset and RADEC_ERR lower than 5 (6). This is especially true for secure detections with DET_LIKE $>$ 10, whereas sources with smaller values of detection likelihood show larger positional errors and populate the bottom right corner of

Tabelle 2.6: Parameters describing the spurious fraction in the point source sample as a function of detection likelihood thresholds and exposure time.

a_1	b_1	c_1	a_2	b_2	c_2
-1.9059×10^{-8}	4.4167×10^{-5}	-1.2476×10^{-4}	2.9317×10^{-6}	-3.070×10^{-3}	2.8982

Notes. The model is described by Equation 2.10.

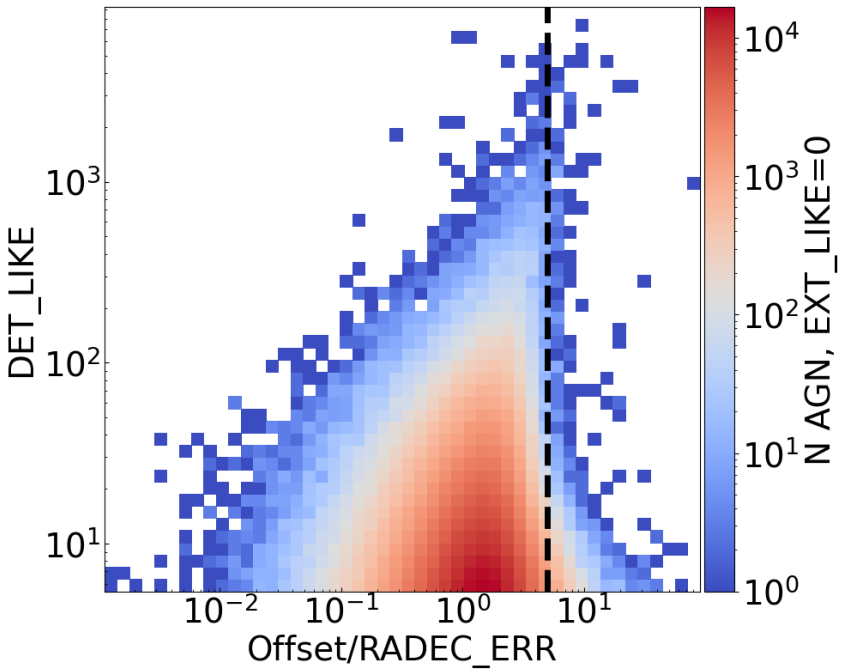

Abbildung 2.22: Positional accuracy of the AGN detected as point-like (EXT_LIKE = 0). This figure shows a 2D histogram in the Offset/RADEC_ERR – DET_LIKE parameter space, and the black dashed line denotes a cut at Offset/RADEC_ERR = 5. The bins are color-coded according to the number of detected AGN in each bin.

Fig. 2.22, which displays how AGN detected as point-like occupy the DET_LIKE – Offset/RADEC_ERR parameter space. The figure is color-coded according to the number of sources in each bin.

2.9 Cluster characterization

Our goal is to characterize a cluster sample that is as pure and complete as possible. On the one hand, we want to maximize the clusters detection rate, making sure that most of the simulated ones are recovered by eSASS. On the other hand, we want to keep the contamination low. This means not only rejecting spurious sources, that are entries in the source catalog that do not correspond to any physical object, but also reducing the contamination due to bright AGN and stars detected as extended objects. Simply applying a high threshold of detection likelihood is not enough to do this, as explained in Sect.2.2 and Fig. 2.3. Therefore, we now focus on the catalog of extended sources, with detection likelihood larger than 6. It contains 7731 entries, 75.2% are clusters, 21.2% are AGN, 3.6% are either spurious sources or secondary matches to simulated objects ($\sim$ 0.9% and 2.7% respectively, see Fig. 2.3), and 5% are stars. Our goal is to single out a complete and pure cluster sample in terms of observables, such as properties measured by the eSASS detection algorithm. We focus on two parameters: the source radius and the extension likelihood. We show the entire source population in this parameter space in the left-hand panel of Fig. 2.23. Clusters are identified by blue circles, AGN by yellow triangles, stars by green squares, and spurious sources by red diamonds. Although most of the sources seem to span the entire srcRAD interval, only clusters reach very high values larger than 200 arcseconds. In addition, galaxy clusters populate the high EXT_LIKE end of this panel. We conclude that a double selection in terms of source radius and extension likelihood is relevant for future cosmological experiments using eROSITA galaxy clusters. We further study the population of detected clusters in terms of extension likelihood, srcRAD and counts in the right-hand panel of Fig. 2.23. Clusters with a larger amount of counts are detected at higher values of EXT_LIKE and show a larger srcRAD. This suggests once again how focusing on the top-right corner of this parameter space, selecting sources with large extension likelihood and source radius, allows one to identify secure clusters emitting a large number of photons. Such correlation also shows the impact on clusters selection of srcRAD. In particular, high count clusters are all located at the high srcRAD end: there are 255 detections with srcRAD $>$ 200 arcseconds and 250 are uniquely matched to a cluster. However, objects with less than 100 counts are detected at different values of srcRAD, indicating that this parameter is less relevant in the selection of low count clusters.

2.10 Summary

Thanks to the eROSITA X-ray telescope, we are detecting clusters of galaxies and active galactic nuclei in the X-ray band at an extraordinary rate. This has a multitude of science

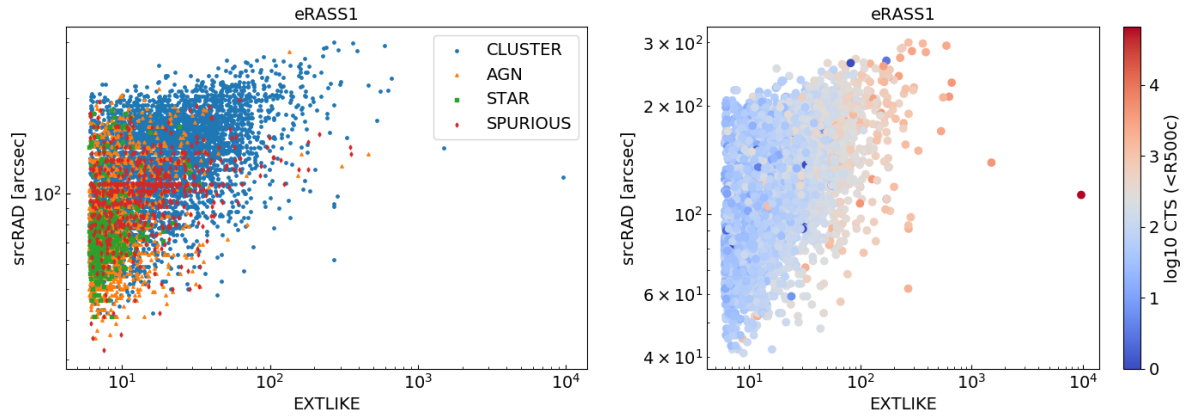

Abbildung 2.23: Distribution of the eSASS sources as a function of srcRAD and EXT_LIKE. **Left-hand panel**: entire source catalog for the eRASS1 simulation in the srcRAD-EXT_LIKE parameter space. Clusters are identified by blue circles, AGN by yellow triangles, stars by green squares, and spurious sources plus secondary matches to simulated objects by red diamonds. **Right-hand panel**: detected clusters color-coded by simulated counts in the 0.2–2.3 keV band inside R_{500c}.

applications, from the evolution of accreting supermassive black holes (Fanidakis et al., 2011), to major steps forward in cosmological studies with X-ray-selected clusters samples (Pillepich et al., 2018b). In this context, it is key to deeply understand the detection and selection of these objects, alongside the properties of the sources recovered by a given detection scheme. Using the models described by Comparat et al. (2019, 2020) we produce a half-sky simulation following the observational strategy of eROSITA, to the depth of the first all-sky survey. The simulated objects include clusters of galaxies, AGN, and stars. The population of simulated clusters is a truthful representation of real clusters of galaxies because the model is built from real observations. The background is simulated following an approach similar to the one detailed Liu et al. (2021). For the eRASS1 simulation, we resample directly the real background maps. This provides an accurate digital twin of real eROSITA data. The result is shown in Fig. 2.1. We run the eSASS detection algorithm on the simulation. We compare the background maps measured on the simulation and the real eRASS1 data. The simulated background is overestimated by ∼10%, but this has a minor impact on the computation of the detection likelihood of each source (see Fig. 2.18). We build a one-to-one correspondence between the source catalog and simulated objects thanks to a photon-based matching algorithm. We classify sources with five different labels: (i) uniquely identified with an AGN or star (PNT), (ii) uniquely identified with a cluster (EXT), (iii) fraction of AGN or star (PNT2), (iv) fraction of a cluster (EXT2), (v) background fluctuation (BKG). Various examples at values of detection likelihood equal to 10 and 20 are shown in Fig. 2.2.
We study the population in the source catalog as a function of different cuts in detection and extension likelihood. We find that the former is efficient in removing spurious

sources from the catalog, which reduces the false detection rate. However, it does not reduce the contamination in the extended select sample due to bright AGN. Instead, progressive EXT_LIKE thresholds are better suited for this task. In addition, at large values of EXT_LIKE > 35 in eRASS1, all the contaminating AGN contain cluster emission. This reduces our estimate of contamination by ∼1%. These results are shown in Fig. 2.3.
Our detection algorithm perfectly recovers the bright end of the number density of objects as a function of flux for both clusters and AGN (see Fig. 2.5 and 2.6). The eSASS detection scheme is suitable for detecting clusters of galaxies. We compare the number of simulated and detected clusters in four different intervals of exposure time. Three of them cover a similar sky area, the fourth one is smaller and centered around the ecliptic poles with extremely high depth. In areas covered by the average simulation depth, we detect half of the simulated structures at flux values of 3.3×10^{-13} erg/s/cm^2 for eRASS1 (see Fig. 2.7). We show how the selection of clusters is not a simple function of flux and exposure time, but the objects with different angular sizes on the sky plane are also detected differently, for instance clusters with smaller extent can be detected as point sources (see Fig. 2.8). This is in agreement with previous work (Pacaud et al., 2006; Burenin et al., 2007; Clerc et al., 2018; Finoguenov et al., 2020). We study how the relaxation state impacts the detection, by exploiting the central emission parameter EM0. The detection is biased toward relaxed clusters with a low EM0. However, the effect is mostly relevant for clusters detected as point sources, as eSASS tends to classify some of these relaxed clusters with EXT_LIKE = 0. This is particularly evident in the low flux regime around 10^{-13} erg/s/cm^2, where a high fraction of simulated objects has a counterpart in the source catalog, but such counterpart is extended (EXT_LIKE >= 6) for only a few of them. The overall detection and characterization of clusters of galaxies are more efficient at the bright flux end, where they are still detected as extended sources (see Fig. 2.9). In the extent-selected sample, the impact of the cool core bias is minimal. These results are in agreement with the eFEDS sample (Ghirardini et al., 2021a; Bulbul et al., 2022).
We combine completeness and purity into the single concept of detection efficiency. We see how choosing specific flux thresholds for varying exposure times (see Table 2.4) allows detecting AGN and clusters with similar levels of completeness in areas covered with different depths. Figure 2.10 shows that the false detection rate in shallower areas is larger. This is due to the lower signal-to-noise ratio, which causes higher relative fluctuations of the background. This is clear for the point-like sample. Progressive cuts in detection likelihood remove the majority of the spurious sources in the point-like sample. The false detection rate drops from 21.5% in areas covered by 150–400 s exposure at DET_LIKE > 5 to < 1% for DET_LIKE > 10. The fraction of clusters mistakenly assigned to the point sample is low, below 4% for every DET_LIKE cut. Similar considerations can be done using thresholds of extension likelihood for clusters. In this case, the pole region shows different behavior of the false detection rate due to the cut in extension likelihood that is very efficient in removing spurious sources and also contaminating AGN in the extent-selected sample. Higher thresholds of extension likelihood are required to lower this value. Progressive EXT_LIKE cuts are very effective in reducing contamination. In the region around the eRASS1 southern ecliptic pole, it drops by 26% from EXT_LIKE > 6 to EXT_LIKE

> 20 (see Fig. 2.10).
We provide area curves as a function of limiting flux built from sensitivity maps in Fig. 2.11. Our measurement is in agreement with the prediction from Merloni et al. (2012), especially at the faint end, but the ability of our method to account for the different exposures guarantees a better sensitivity at the bright flux end. We finally compute the X-ray luminosity of galaxy clusters in the eRASS1 simulation by fitting the surface brightness profile of each object, following the approach described by Ghirardini et al. (2021b). We show that on average we recover the simulated values of X-ray luminosity in Fig. 2.12.
We discuss how to best construct volume-limited samples applying different flux limits in areas covered with varying depth by eROSITA (see Fig. 2.13). This translates into different samples of clusters according to the values of luminosity and redshift. It guarantees an even sampling of the large-scale structure of the Universe also in a case of nonuniform coverage. The selection of these samples and their relative contribution to the cluster number density distribution as a function of redshift and mass is shown in Fig. 2.14.
We presented and analyzed a precise digital twin of the first eROSITA all-sky survey. Performing such a high-level simulation significantly increases our understanding of real data, allowing us to analyze how a realistically complex population of sources is observed by eROSITA. We studied the detection rate of galaxy clusters and AGN, accounting for the fraction of simulated objects that are detected by the eSASS pipeline, together with quantifying the false detection rate and contamination levels in the source catalog for point-like and extended sources. Using these results, one can control the fraction of false detections and or contaminants according to specific cuts of detection and extension likelihood in the real eRASS1 catalog. For example, this is useful for constructing different cluster samples, allowing for a precise contamination fraction. We addressed additional effects impacting the detection of clusters, such as their dynamical state and their physical size. This work is key toward characterizing the population of extragalactic sources in real eROSITA data.

Kapitel 3

The mass function dependence on the dynamical state of dark matter halos

Galaxy clusters are the most massive virialized, gravitationally bound structures in the Universe. Clusters in cosmology are used to construct the halo mass function, which indicates the mass density of halos in a specific volume, in a small mass interval, included between M and M + dM (Weinberg et al., 2013). Early theoretical description of the mass function was given by Press and Schechter (1974), based on the assumption that Gaussian density perturbations overcoming a fixed density contrast collapse into halos. This formally accounts for only half of the total halo mass in the Universe. An alternative approach, employing the excursion set theory, solved these shortcomings by considering the probability of crossing a given barrier with random walks (Bond et al., 1991). This provides a good prediction for high-mass halos, but it predicts too many low-mass objects. The introduction of ellipsoidal collapse corrected these differences between simulations and theory (Sheth and Tormen, 1999, 2002).

The mass function has been extensively studied in more recent works, attempting to find a universal model, independent from cosmology (Jenkins et al., 2001; Tinker et al., 2008; Bhattacharya et al., 2011; Despali et al., 2016; Bocquet et al., 2016; Comparat et al., 2017; Bocquet et al., 2020). A robust way to build such a mass function model is using N-body simulations (Kravtsov et al., 1997; Springel, 2005), e.g. MultiDark (Prada et al., 2012; Klypin et al., 2016). The generalization of such models as a function of cosmological parameters is best handled by emulating the mass function based on large sets of simulations (e.g., McClintock et al., 2019; Nishimichi et al., 2019; Bocquet et al., 2020). It is important to precisely predict the halo mass function to fulfill the potential of current and future X-ray, SZ, or optical cluster surveys, such as eROSITA (Merloni et al., 2012; Predehl et al., 2021), Planck (Zubeldia and Challinor, 2019), SPT-3G (Benson et al., 2014), CMB S4 (Abazajian et al., 2019) SPIDERS/eBOSS (Dawson et al., 2016; Finoguenov et al., 2020), DESI (DESI Collaboration et al., 2016), 4MOST (de Jong, 2011; Finoguenov et al., 2019), Euclid (Laureijs et al., 2011), LSST (LSST Science Collaboration et al., 2009), WFIRST (Spergel et al., 2015). These future surveys will provide tighter

constraints on mass-observable scaling relations. This means that systematic uncertainties due to the accuracy of the halo mass function model and its evolution with redshift will contribute to the total error budget significantly more than in the previous cluster counts experiments (Salvati et al., 2020). Therefore, a detailed prediction of the theoretical dark matter halos statistic is required. In this work, we calibrate a mass function model that additionally includes dynamical properties of dark matter halos. Relaxed dark matter halos can be selected according to multiple diagnostics (Neto et al., 2007; Macciò et al., 2008; Prada et al., 2012; Klypin et al., 2016): (i) the virial parameter $2K/|W|-1$, where K and W are, respectively, the kinetic and potential energy within the virial radius. (ii) The spin parameter $\lambda = J\sqrt{E}/GM^{5/2}$ (Peebles, 1969) traces the dynamical state of the halo. (iii) The fraction of substructures, which is higher in unrelaxed halos. (iv) The offset parameter $X_{\rm off} = |R_{\rm peak} - R_{\rm cm}|/R_{\rm vir}$ (Behroozi et al., 2013; Klypin et al., 2016). It is the difference between the position of the peak of the dark matter density profile and the center of mass, normalized by virial radius. If the halo is perfectly relaxed, the peak of the profile will correspond to the center of mass, so that $\rm X_{off}$ will be small. On the other hand, higher $\rm X_{off}$ values will indicate an unrelaxed halo (e.g. merger, accretion). A combination of these quantities can be used.

To illustrate the respective contribution to the mass function of relaxed and disturbed halos, we divide the sample of halos in HugeMultiDark (hereafter HMD, see Sect. 3.2) at z=0 using $X_{\rm off}$ and λ. We consider the offset parameter in physical scale $X_{\rm off,P} = |R_{\rm peak} - R_{\rm cm}|$, i.e. not normalizing by the virial radius. We consider halos with $X_{\rm off,P} < 100$ kpc/h and $\lambda < 0.007$ as relaxed. In Fig. 3.1, we show the halo mass function of the complete halo population, the relaxed and disturbed halo population. The left panel shows the distribution of redshift zero halos in the $X_{\rm off,P}$ and λ plane. The right panel shows the three multiplicity functions (see formalism in Sect. 3.1) sampled by all halos (green), relaxed halos (blue), and disturbed halos (orange). The bottom panel shows their relative contribution as a function of halo mass. It is clear how at $10^{13.5} M_\odot/h$, the contribution from relaxed structures dominates by a factor of about 0.8 dex. The multiplicity functions of the two samples cross each other at $10^{14.5} M_\odot/h$, then unrelaxed ones take over at the high-mass end. This approach offers a possible connection to selection effects in observations, such as the cool core bias in X-rays (Eckert et al., 2011; Käfer et al., 2019, 2020). It possibly offers a solution to mitigate biases in a cosmological interpretation of clusters abundance. This might improve cosmological constraints using X-ray selected clusters (Ider Chitham et al., 2020). In this chapter, we investigate the variations of the dark matter halo mass function as a function of the dynamical state of the constituting halos. To trace the dynamical state, we use $\rm X_{off}$ and λ. This chapter is structured as follows. We define the formalism in Sect. 3.1. We present the N-body data used in Sect. 3.2. We present the average relations between concentration, $X_{\rm off}$, spin and mass, as well as their distribution around mean values in Sect. 3.3. We define the generalized mass function framework in Sect. 3.4. We present the generalized model of the halo mass function as a function of $\rm X_{off}$, and λ in Sect. 3.5. We present the results of the fit and the best-fit parameters in Sect. 3.6. We summarize our findings in Sect. 3.9, and discuss their implications for cosmological studies.

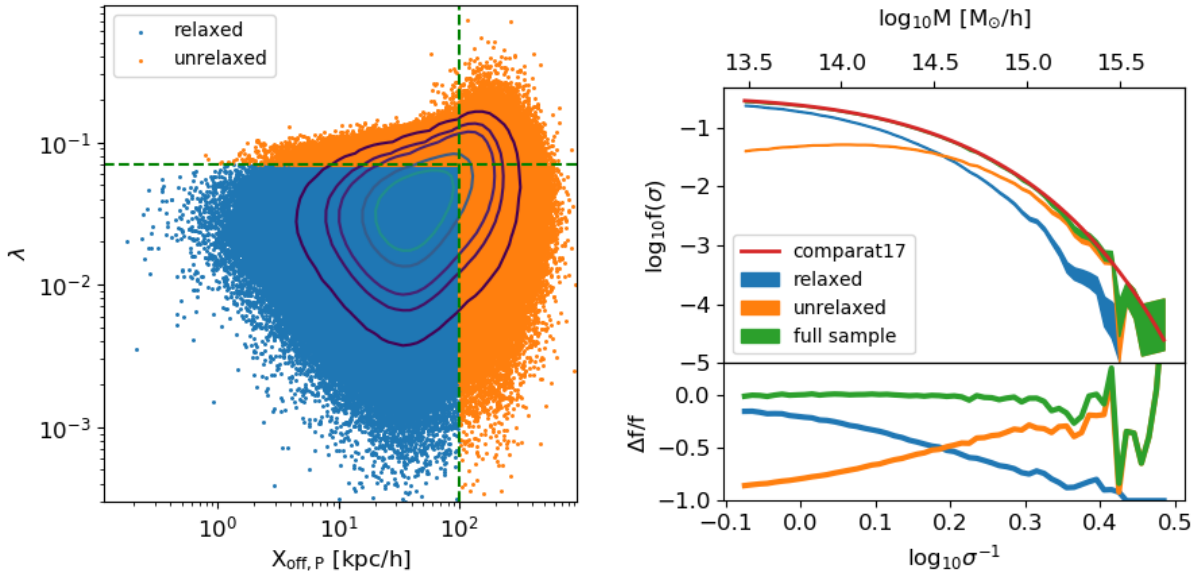

Abbildung 3.1: **Left panel**: Distribution of redshift zero dark matter halos in the $X_{\rm off,P}$ and λ plane. Cuts in $X_{\rm off,P}$ and λ are applied to divide relaxed (blue) and disturbed (orange) structures. The contours contain 1%, 5%, 10%, 30%, 50% of the data. **Right panel**: Halo mass functions v.s. mass (σ) at redshift 0, as defined in Sect. 3.1 and using Equations 3.11 and 3.12. This mass function is built with different subsets of halos from HMD at z=0 (see Sect. 3.2): the red line indicates the model from Comparat et al. (2017), the shaded areas represent the relaxed (blue), unrelaxed (orange) and the full sample (green) of halos. The areas cover 1σ uncertainties. The lower panel shows the residuals fraction of each component compared to the red model in the upper panel.

Tabelle 3.1: Correspondence between mass, peak height, and variance of the linear density field at z=0 and z=0.5.

Mass	z=0		z=0.5	
$\log_{10} M_\odot/h$	$\nu = \delta_c/\sigma$	$\log_{10}(1/\sigma)$	$\nu = \delta_c/\sigma$	$\log_{10}(1/\sigma)$
10	0.446	-0.577	0.579	-0.465
10.5	0.506	-0.523	0.657	-0.41
11	0.58	-0.463	0.752	-0.35
11.5	0.673	-0.399	0.872	-0.286
12	0.79	-0.329	1.02	-0.216
12.5	0.942	-0.253	1.22	-0.14
13	1.14	-0.17	1.48	-0.0568
13.5	1.41	-0.0779	1.83	0.035
14	1.78	0.0236	2.31	0.137
14.5	2.31	0.137	3	0.25
15	3.09	0.263	4.01	0.376
15.5	4.29	0.405	5.56	0.518

These quantities are described by Eq. 1.25 and 1.43.

3.1 Formalism, definitions

The growth of the density perturbations in the matter field is described by the evolution of the overdensity field and its variance as a function of scale. The variance of the smoothed density field σ^2(M,z) is defined in Eq. 1.25. We report explicit values of mass, variance σ (Eq. 1.25), and peak height ν (Eq. 1.43) in Table 3.1 for z=0 and z=0.5. We write the mass function in its differential form as in Eq. 1.41, (see Allen et al., 2011, for a recent review). A comprehensive list of models of the multiplicity function f(σ) is available in Table 1 of Murray et al. (2013). we interchangeably use as mass variable either σ (Equation 1.25) or peak height (Equation 1.43) or mass.

3.2 Simulations

Here we describe the set of gravity-only simulations and the halo finding post-process.

3.2.1 MultiDark

We use the MultiDark simulations (Prada et al., 2012; Riebe et al., 2013; Klypin et al., 2016). They are computed in a Flat ΛCDM Planck (Planck Collaboration et al., 2014a) cosmology (H_0 = 67.77 km s^{-1} Mpc^{-1}, Ω_{m0} = 0.307115, Ω_{b0} = 0.048206, σ_8 = 0.8228) with the GADGET-2 code (Springel, 2005). It is one of the largest sets of high-resolution ($\sim 4000^3$ particles) N-body simulations. We use three MultiDark simulations: HMD, BIGMD, MDPL2, see details in Table 3.2. Alternative simulations which could be used for this

Tabelle 3.2: N-body simulations used in this analysis. L: length of the box in Gpc/h. M_p: mass of the particle in $M_\odot/h$. M_{min}: minimum halo mass considered $M_{vir} > M_{min}$ in $M_\odot/h$. Number of halos in the snapshots at z=0.

Name	L	M_p	M_{min}	N halos
HMD	4.0	7.9×10^{10}	2×10^{13}	13 330 574
BigMD	2.5	2.4×10^{10}	5×10^{12}	27 575 832
MDPL2	1.0	1.51×10^{9}	4×10^{11}	17 036 888

project include Millennium-XXL, DarkSkies, Q Continuum, v^{2}GC simulation, described in Angulo et al. (2012); Skillman et al. (2014); Heitmann et al. (2015); Ishiyama et al. (2015) respectively.

The complete list of simulation outputs (snapshots) utilized are given in Table 3.9, where the expansion parameter a, and the corresponding redshifts are reported for each snapshot.

3.2.2 Halo finding

Finding halos in dark matter simulations is not an easy task (see Knebe et al., 2013; Behroozi et al., 2015, for a review). In this study, halos are identified by the rockstar (Robust Overdensity Calculation using K-Space Topologically Adaptive Refinement), consistent-trees algorithms (Behroozi et al., 2013). It is based on adaptive hierarchical refinement of friends-of-friends (FOF) groups. It works with six phase-space dimensions (halo positions and velocities), and one time dimension. This allows tracking relative motions and merging history between substructures in different snapshots. rockstar computes the halo mass of identified objects by removing the unbound particles inside the virial radius. Virial mass and virial radius are related by

$$M_{vir}(z) = \frac{4}{3}\pi\Delta_{vir}(z)\Omega_M(z)\rho_b(z)R_{vir}^3, \tag{3.1}$$

where R_{vir} encompasses a mean halo density equal to the background matter density multiplied by Δ_{vir}, Ω_M is the matter density parameter, and ρ_b is the matter density of the Universe. The overdensity over the matter background Δ_{vir} is defined according to Bryan and Norman (1998) and Eq. 1.35. Given the reference cosmology adopted in this work, the virial overdensity is equal to 332.5 at z=0 and asymptotically tends to 178 at high redshift. The recovery of main properties such as position, mass, and circular velocity is consistent between different finders. However, derived properties such as spin show a ∼20% scatter (Knebe et al., 2013). This holds especially for low-mass halos with less than 30-40 particles, where the identification of substructures is not straightforward (Knebe et al., 2011). The low-mass limits in table 3.2 are set at more than ∼200 particles per halo in all boxes. This ensures accurate halo properties (Behroozi et al., 2013; Knebe et al., 2013).
In this work, we use the virial overdensity (Equation 1.35). The virial mass function has been shown to be the one that comes closest to universality (Despali et al., 2016).

3.3 Concentration, offset, spin: empirical relations with peak height and redshift

In the footsteps of Klypin et al. (2016); Rodriguez-Puebla et al. (2016), we analyze the average relations linking concentration, λ, and $X_{\rm off}$ to the peak height, respectively in sections 3.3.1, 3.3.2, 3.3.3. We additionally analyze the distributions of these quantities around the mean relations. The mean relations are fitted by models that simultaneously account for the mass and redshift dependence of the relations (Eqs. 3.5, 3.7, 3.9). The probability density functions (PDF) of concentration, spin, and $X_{\rm off}$ are fitted by modified Schechter models, respectively equations 3.6, 3.8, 3.10. PDFs at different redshifts are modeled independently.

3.3.1 Concentration – mass – redshift relation

We study the relation between concentration and mass. Numerical simulations (Navarro et al., 1996) showed that to a good approximation ($\sim 10-20\%$), the density of dark matter halos are described by the profile in Eq. 3.2,

$$\rho(r) = \frac{\rho_s}{(r/R_s)(1+r/R_s)^2}, \tag{3.2}$$

where R_s is the scale radius. ρ_s is the characteristic density of the halo and is equal to

$$\rho_s = \rho_{\rm crit}\frac{\Delta}{3}\frac{c^3}{\ln(1+c) - c/(1+c)}, \tag{3.3}$$

where c is the concentration, Δ the overdensity and $\rho_{\rm crit}$ is the critical density of the Universe. The concentration is a dimensionless quantity defined by

$$c_\Delta = R_\Delta/R_s, \tag{3.4}$$

where Δ can refer to any threshold. In this work, we consider the virial overdensity over the matter background $\Delta_{\rm vir}$ (see Equation 1.35). Therefore, we study the virial concentration $c_{\rm vir} = R_{\rm vir}/R_{\rm s}$.

The concentration – mass relation is an important part of models describing galaxy clusters (e.g. reviews from Allen et al., 2011; Umetsu, 2020) or gravitational lensing (e.g. reviews from Bartelmann, 2010; Kilbinger, 2015).

This relation has been extensively studied in simulations. The concentration anti-correlates with mass with negative redshift trend (e.g. Diemer and Joyce, 2019; Ragagnin et al., 2019). Its detailed trend depends on the measurement method, both in simulations (Meneghetti and Rasia, 2013; Lang et al., 2015; Poveda-Ruiz et al., 2016) and observations (Foëx et al., 2014; Phriksee et al., 2020; Du et al., 2015; Shan et al., 2017). This introduces possibles biases in the measure of concentration (Sereno et al., 2015; Cibirka et al., 2017; van Uitert et al., 2016). Diemer and Kravtsov (2015) found that the relation shows the smaller deviation from universality when adopting the definition c_{200c}. However, it is not

completely universal, meaning that concentration is described not only by mass or ν but also by assembly history. Leaving aside biases in definition and measurement, different models have been proposed to describe it: power-laws (*e.g.* Duffy et al., 2008; Dutton and Macciò, 2014); a combination of power-laws, to describe the high-mass upturn (*e.g.* Klypin et al., 2016; Diemer and Joyce, 2019); semi-analytic models based on Press-Schechter theory (Correa et al., 2015). In this section, we extend the models from Klypin et al. (2016). We adjust a global model that includes a redshift dependence for high-mass halos at relatively low redshift ($z < 1.5$), which is particularly interesting for halos hosting galaxy clusters.

Model

We model the concentration c as a function of the rms of the overdensity field $\sigma(M, z)$ (not as the function of halo mass M). Our parametrization of the concentration – σ relation is a generalization of that of Klypin et al. (2016) and reads:

$$c(\sigma, z) = \frac{b_0}{(1+z)^{0.2}}\left[1 + 7.37\left(\frac{\sigma}{a_0(1+z)^{1/2}}\right)^{3/4}\right]... \\ \left[1 + 0.14\left(\frac{\sigma}{a_0(1+z)^{1/2}}\right)^{-2}\right]. \tag{3.5}$$

The best-fit values are in Table 3.4. We find best-fit values of $a_0 = 0.754091 \pm 0.000004, b_0 = 0.574413 \pm 0.000002$, in agreement with Klypin et al. (2016). This model is fitted using halos with $M > 10^{12.5} M_\odot/h$ ($\nu \sim 0.95$ at z=0). The low-mass end is not sampled due to the particle mass resolution. We do not consider a high redshift regime, where there is a statistical limitation of high-mass halos. We choose the binning limits following Klypin et al. (2016). The strength of our model resides in the ability to predict the concentration - σ relation for a variety of masses and redshifts with a single equation. We recover the hockey stick shape of the relation (see Fig. 3.2). This is consistent with Prada et al. (2012), Klypin et al. (2016), and Diemer and Joyce (2019) (the latter also includes a cosmology dependence). This upturn feature is absent in Duffy et al. (2008); Wang et al. (2019), due to the smaller volumes analyzed (400 Mpc/h, 500 Mpc/h respectively). They are not large enough to obtain a significant number of high-mass objects and probe the upturn.

We describe the distribution of concentration around its mean value and we model its probability density function. We use snapshots from HMD, at four redshifts. Each snapshot is divided into six slices of mass. The PDFs obtained at each redshift are fitted simultaneously. It allows including a σ dependence in our model (see Equation 3.6),

$$P(c, \sigma) = A\left(\frac{c}{x_0\sigma^{e_0}}\right)^{\alpha\sigma^{e_1}} \exp\left[-\left(\frac{c}{x_0\sigma^{e_0}}\right)^{\beta\sigma^{e_2}}\right]. \tag{3.6}$$

The best-fit parameters are included in Table 3.4. Uncertainties are negligible for most parameters. The distribution of concentration around its mean value is well described by a modified Schechter function with mass-dependent terms (Equation 3.6). We use the best-fit model to predict the distribution of concentration in fixed mass slices as a function of redshift. The result is shown in Fig. 3.3. The full distribution of concentration (not sliced in different mass intervals) is well described by a modified Schechter as well.

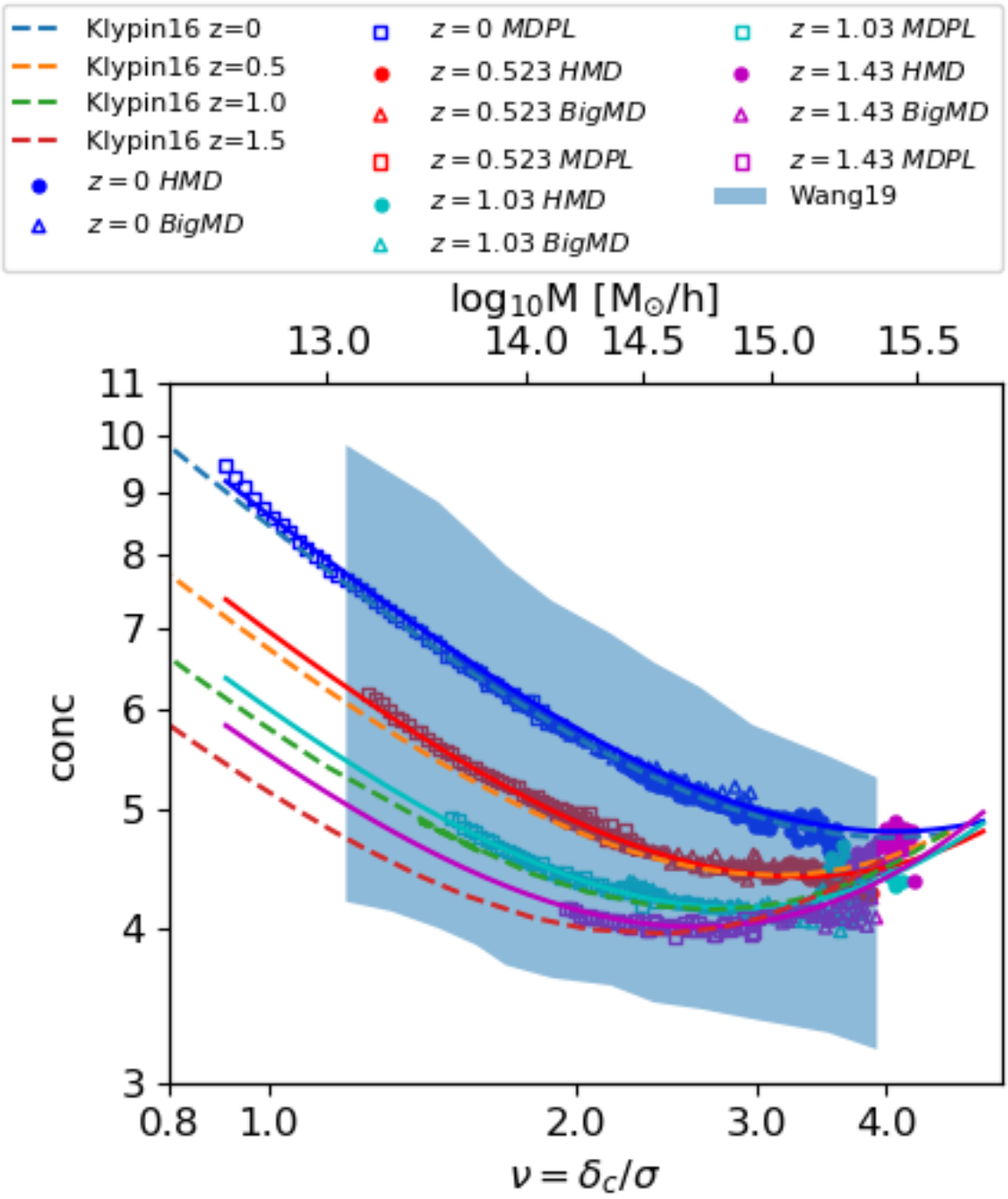

Abbildung 3.2: Concentration - σ relation (Equation 3.5). Circular dots, triangles, and squares represent HMD, BigMD, MDPL2 respectively. They are color-coded by redshift. Straight lines indicate our best-fit model, dotted lines show the model from Klypin et al. (2016), while the shaded blue area indicates the distribution from Wang et al. (2019) at z=0. The upper x-axis converts peaks values into mass at z=0.

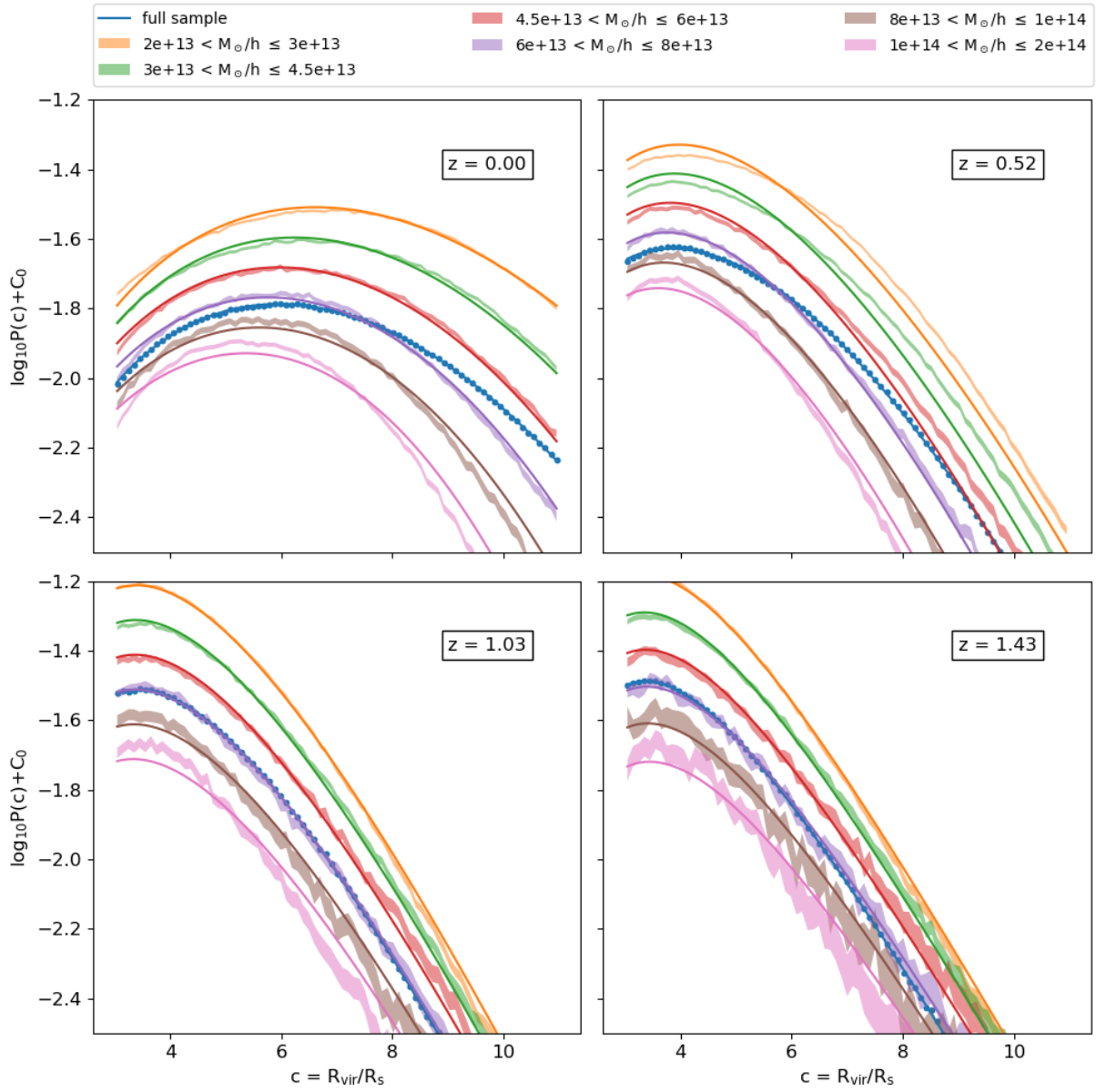

Abbildung 3.3: Probability density function of the concentration (Equation 3.6) at different redshifts, values are reported in the title of each panel. Each set is divided into mass slices and color-coded accordingly. The shaded areas represent the data with 1σ error, while straight lines indicate the best-fit model. The blue points and the line represent the total sample not sliced in mass. For clarity, each line and its fit is shifted by 0.1 dex on the y-axis. This means that the constant C_0 assumes values of (+0.3,+0.2,+0.1,0.0,-0.1,-0.2). The purple line is not shifted, therefore it is the one with the proper normalization.

Discussion

The average value of concentration increases at late times. It goes from 4 at z=1.43 to 5.8 at z=0 for halos corresponding to peak height $\nu = 2$. We confirm the recent discovery of the concentration upturn at high masses (Klypin et al., 2016; Diemer and Joyce, 2019). Klypin et al. (2016) suggest that the high-mass upturn is caused by the tendency of regions with smaller root mean square variance in the overdensity field to be more spherical. These regions are the ones that evolve into high-mass dark matter halos. Because of this aspect, gravitational accretion of matter toward the center is more efficient, result in higher concentration. In addition, very massive halos correspond to the highest peaks of the density field. This means that it is more probable to find them in a phase characterized by strong accretion and therefore higher concentration, as shown by Ludlow et al. (2012).
We find that the concentration of halos with different masses shows a different evolution with redshift (see Fig. 3.3). At z=0 the concentration of high-mass halos gives a smaller contribution to the total probability density function (hereafter PDF) in the high concentration tail. At higher redshift, we see high concentration halos with both high and low mass. The opposite holds for low-mass halos, which contribute more to the total PDF at higher z in the low concentration regime. At high redshift, the distribution of concentration for different mass bins has the same shape. Conversely, at low redshift, this statement does not hold. Indeed, the low-mass distribution is much broader than the high-mass one. There is also a general redshift trend: with decreasing redshift the number of high concentration, low-mass halos increases. More recently in time, the distribution is flatter, in fact, the slope of the power-law is smaller, going from ~ 4.5 at z=1.43 to ~ 1.4 in the present day. Moreover, the mass trend of the exponential decay at z=0 is negative with sigma ($e_2 = -0.959$), which translates into a faster decrease at high-mass. This means that the distribution of concentration of low-mass halos evolves more than that of high-mass objects. For example, at concentration $c = 8$, the PDF of high-mass halos ($1 \times 10^{14} < M < 2 \times 10^{14} M_{\odot}/h$) changes by 0.32 dex between $z = 1.43$ and $z = 0$, while for low-mass objects ($2 \times 10^{13} < M < 3 \times 10^{13} M_{\odot}/h$) it varies by 0.49 dex. There is a net difference of 0.17 dex between the two, meaning that the high end of concentration evolves $\sim$1.5 times faster for lower mass halos compared to high-mass ones. This is in agreement with the fact that these halos evolve in different environments. This causes a different redshift evolution of the distribution of concentration of halos with different masses. Structures with the same mass but different formation histories present different properties. This aspect is related to the notion of assembly bias (Gao et al., 2005; Croton et al., 2007; Angulo et al., 2008). In fact, there is a correlation between halo concentration and the age of the universe when the progenitor reached a fixed fraction of its current mass (Zhao et al., 2009). Moreover, a faster mass accretion causes a slower increase of concentration. Very massive objects reside in highly populated environments, which slows down the evolution of concentration (Zhao et al., 2003). This can be used to predict concentration as a function of assembly history (Giocoli et al., 2012; Ludlow et al., 2013, 2014). The latter is also related to the evolution of cosmological parameters. This allows modeling of the concentration mass relation for cold and warm dark matter halos (Ludlow et al., 2016).

In general, high-mass halos cluster more than low-mass ones. Low-mass objects, part of which live in isolated environments will experience different histories. Isolated low-mass halos will result highly concentrated, while the concentration of low-mass halos in dense environments will stay low. Contreras et al. (2019) found that low concentration halos cluster more than halos with high concentration at high redshift. This is in agreement with our results.

3.3.2 λ – mass – redshift relation

In this subsection, we study the relation between the spin parameter and mass. An accurate theoretical description of the distribution of the halo spin is key to understand possible implications of the systemic rotation on cluster sample definitions. Such rotation is induced by a combination of the initial spin of the halo, the infalling material, and the merging activity.
Measurements of cluster rotation were obtained at low redshift, using member galaxies to infer the rotation movement (Hwang and Lee, 2007; Tovmassian, 2015; Manolopoulou and Plionis, 2017; Bilton et al., 2019). The spin in galaxy clusters has also been studied in X-rays (Bianconi et al., 2013; Eckert et al., 2019a). High-resolution data are needed to explore this topic, both in the optical (Song et al., 2018) and X-ray (Hitomi Collaboration et al., 2018) band. Moreover, cluster cores can be analyzed by distortion of the 6.7 keV line (Sunyaev et al., 2003), fluctuations in the CMB due to rotation (Rephaeli, 1995; Cooray and Chen, 2002), and rotational kSZ effect (Baxter et al., 2019). Inferring motion with the SZ effect was also demonstrated for relaxed clusters with a significant spin on simulated (hydro-dynamically) clusters (Baldi et al., 2018, 2019). Future SZ and X-rays surveys might enable such measurements on a large number of clusters. A statistical description of the halo population as a function of spin is thus of interest and developed in this section.

Model

We analyze the spin–mass–redshift relation and its PDF. The spin is defined as $\lambda = JE^{1/2}/GM^{5/2}$ (Peebles, 1969). We fit the mean relation with mass as a linear relation (Eq. 3.7).

$$\lambda(\sigma) = a_0 + b_0\sigma. \tag{3.7}$$

There is no noticeable dependence on redshift, so we do not consider a redshift evolution in this model. The best-fit parameters are reported in Table 3.5. The relation is shown in Fig. 3.4. We find that the spin correlates weakly with halo mass, as Bett et al. (2007); Rodriguez-Puebla et al. (2016) did. The PDF of the spin parameter is best-fitted by a modified Schechter law with no mass dependence (Equation 3.8).

$$P(\lambda) = A\left(\frac{\lambda}{x_0}\right)^{\alpha} \exp\left[-\left(\frac{\lambda}{x_0}\right)^{\beta}\right]. \tag{3.8}$$

Distributions for different redshift snapshots are shown in Fig. 3.5 (see Table 3.5 for best-fit parameters). The result is consistent with previous findings, (e.g. Rodriguez-Puebla

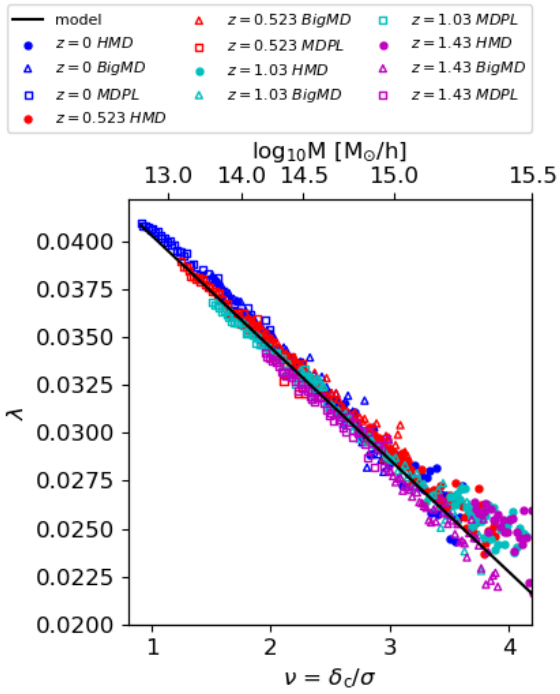

Abbildung 3.4: λ-σ relation (Equation 3.7). The model is a linear relation, with no redshift trend. Data points are color-coded by redshift, while different geometrical shapes refer to different simulations: squares for HMD, triangles for BigMD, and circles for MDPL. The straight line indicates the best-fit model, which considers all simulations and redshift at the same time. The best-fit parameters are given in Table 3.5. The upper x-axis converts peaks values into mass at z=0.

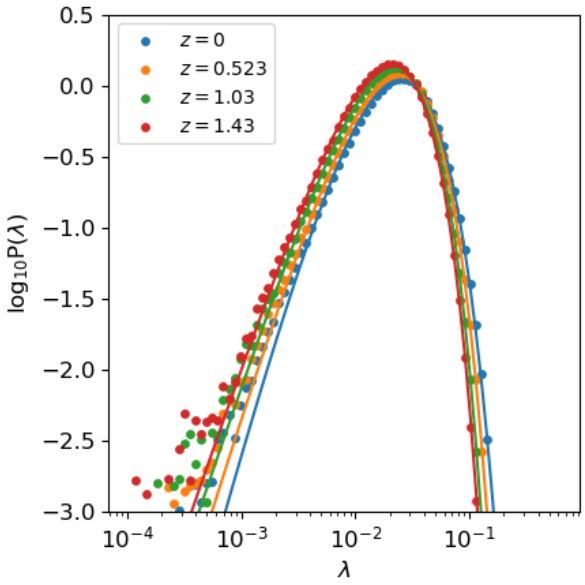

Abbildung 3.5: PDF of λ (Equation 3.8). Individual points represent spin bins used to compute the distribution, straight lines refer to the best-fit modified Schechter model. They are color-coded by redshift. We do not consider mass dependence in this relation. Each redshift slice is fitted independently. The best-fit parameters are given in Table 3.5.

et al., 2016). The evolution with redshift of the PDF shows that with time, halos build up higher spins: the maximum of the PDF shifts to higher spin values when redshift decreases (Fig. 3.5). We also tried a lognormal distribution as a model, unsuccessfully.

Discussion

The small correlation with mass and the well-modeled evolution with redshift makes it a rather simple dependence to account for in statistical studies of the halo population. From the perspective of the measurement of the halo mass function based on a cluster sample, marginalizing over the spin is possible, with limited complications. It also shows the spin cannot be considered as a candidate for assembly bias.

3.3.3 $X_{\rm off}$ – mass – redshift relation

The offset parameter $X_{\rm off}$ traces the relaxation state of the halo (Thomas et al., 2001; Neto et al., 2007; Henson et al., 2017). Hollowood et al. (2019) analyzed the miscentering in SDSS galaxies followed up with Chandra X-ray observations. Provided a link between $X_{\rm off}$ and the brightest cluster galaxy (BCG) to X-ray displacement, an estimation of the bi-variate mass and $X_{\rm off}$ function from observations of clusters is possible. To interpret it, one needs a detailed description of the link between $X_{\rm off}$ and mass, detailed in this section. We find that low-mass halos have smaller offset than high-mass ones at each redshift.

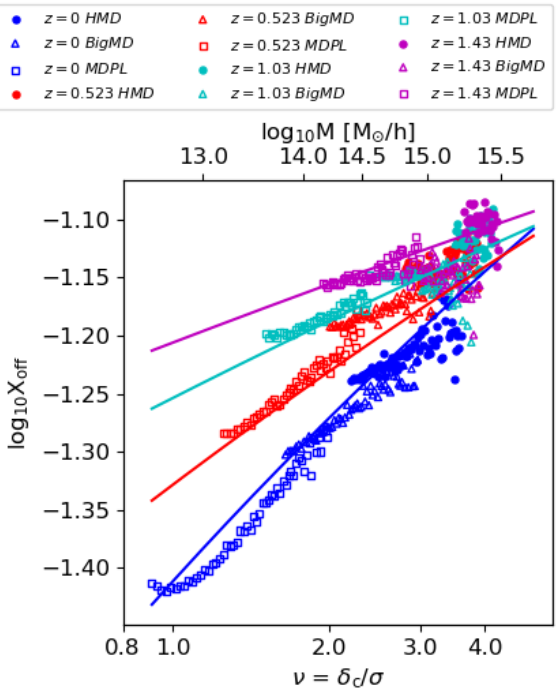

Abbildung 3.6: $X_{\rm off}$-σ relation (Equation 3.9). Circular dots, triangles, and squares represent HMD, BigMD, MDPL2 respectively. They are color-coded by redshift. Straight lines show the best-fit model. The best-fit parameters are given in Table 3.6. The upper x-axis converts peaks values into mass at z=0.

Moreover, the offset parameter is reduced by a factor of ~ 1.5 from $z \sim 1.5$ to $z = 0$. This is in agreement with the fact that structures relax in time.

Model

We model the $\mathrm{X_{off}} - \sigma$ relation with a redshift dependent power-law, see Equation 3.9:

$$\log_{10} X_{\rm off} = \frac{a_0}{E(z)^{0.136}} \sigma^{b_0 E(z)^{-1.11}}, \tag{3.9}$$

where $E(z)$ is the dimensionless Hubble parameter. The best-fit parameters are given in Table 3.6. We find $a_0 = -1.30418 \pm 0.00001$, $b_0 = 0.15084 \pm 0.00001$. We find a significant redshift evolution of the normalization and the slope of the relation. The data obtained from MultiDark simulations and the best-fit models are shown in Fig. 3.6. The best-fit parameters of the $X_{\rm off} - \sigma$ relation and the distribution of $X_{\rm off}$ are given in Table 3.6. As for concentration, the strength of this equation relies upon its ability to predict an average $X_{\rm off}$ value given the mass and the redshift of a dark matter halo.

We describe the distribution of $X_{\rm off}$ around its mean value with a modified Schechter function (Equation 3.10). The PDF of $X_{\rm off}$ does not show mass dependency. Indeed it is

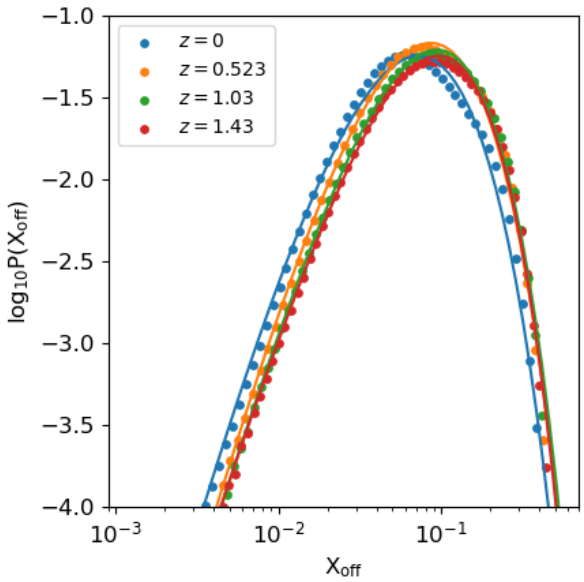

Abbildung 3.7: Probability density function of $X_{\rm off}$. Each panel shows the distribution at a specific redshift. Scatter points indicate the data, while straight lines represent the modified Schechter model. The samples are color-coded by redshift. Each redshift slice is fitted independently by Equation 3.10. The best-fit parameters are given in Table 3.6.

included in the normalization to the virial radius $R_{\rm vir}\alpha M_{\rm vir}^{1/3}$. Therefore, we do not consider any σ dependence.

$$P(X_{\rm off}) = A\left(\frac{X_{\rm off}}{x_0}\right)^{\alpha} \exp\left[-\left(\frac{X_{\rm off}}{x_0}\right)^{\beta}\right]. \tag{3.10}$$

We fit all halos in each redshift snapshot together. Figure 3.7 shows distributions at different redshifts, fitted independently from one another. The parameters are given in Table 3.6.

Discussion

We find a non-zero slope for the relation between $X_{\rm off}$ and mass. On average, high-mass halos have a higher offset parameter. This is described by the negative a_0 parameter. This is in agreement with the hierarchical picture of structure formation. High-mass halos formed recently and have not had time to dynamically relax. A further contribution is given by the environment surrounding these structures. High-mass halos form in the knots of the large scale structure, where more matter is available for inflows and mergers, making these objects more disturbed. In a picture where $X_{\rm off}$ possibly traces the cool core - non-cool core classification of galaxy clusters in X-rays (Eckert et al., 2011), massive structures have a higher fraction of non-cool cores. This influences the high-mass tail of the mass function (see Fig. 3.1). Therefore, given an X-ray flux limit, the mass function of a relaxed galaxy clusters sample will be complete to lower masses than an unrelaxed one. This effect also

evolves with redshift. At high z, halos show a higher offset. In this context, it means that it is more difficult to detect structures at earlier times. More recently, structures have had time to relax and therefore show smaller values of the offset parameter. This is linked to the development of cool cores at low redshift (Ettori and Brighenti, 2008). The PDF shows a power-law growth from low offset values (slope $\alpha = 3.71$ at z=0) and exponential cutoff at high $X_{\rm off}$. It is described by a modified Schechter function (Equation 3.10). Its maximum shifts by a factor ~ 1.5, from $X_{\rm off} \sim 0.09$ to $X_{\rm off} \sim 0.06$, between redshift 1.4 and 0, confirming that halos have more time to relax. The shape of the distribution does not show a significant redshift trend. The width of the probability density functions, measured at $\log_{10} P(X_{\rm off}) = -2.5$, at $z = 0$ and $z = 1.4$ agree with 2.8% accuracy. At z=0 these values span from $X_{\rm off} \sim 0.01$ to $X_{\rm off} \sim 0.3$. This impedes the offset to be a possible assembly bias candidate.

3.4 Generalized mass function

Generalizations of the mass function have been made in a number of directions: cosmological parameters, angular momentum, and friction (Achitouv and Corasaniti, 2012; Achitouv et al., 2014; Del Popolo et al., 2017) via detailed modeling of the collapse barrier. Nevertheless, it is technically demanding to connect such parameters to observations. In this section, we generalize the mass function formalism to include additional variables ($X_{\rm off}$,λ) in its formulation. These two quantities describe properties of dark matter halos and will hopefully become connectable to observational properties.

3.4.1 Definitions

We compute the mass function as follows:

$$\frac{dn}{dlnM} = \frac{\Delta N_M}{V \Delta \ln M}, \tag{3.11}$$

where ΔN_M is the number of halos in each mass bin $\Delta \ln M$ and V is the total volume of the simulation. It corresponds to Eq. 1.41. The $M(\sigma)$ relation and its first derivative are computed with COLOSSUS (Diemer, 2018). By convention, the relation between a certain mass and its corresponding scale is normalized with the matter density at z=0.

By combining the previous equations with Eq. 1.41, we estimate a multiplicity function:

$$f(\sigma) = \frac{dn}{dlnM} \frac{M}{\rho_m} \left(\frac{dln\sigma^{-1}}{dlnM} \right)^{-1}. \tag{3.12}$$

3.4.2 Generalization

We include the offset parameter and spin by generalizing the approach described in the previous section. The number density of halos as a function of mass and dynamical state

is described by Equation 3.13:

$$\frac{dn}{\mathrm{dln}M\ \mathrm{dlog}X_{\mathrm{off}}\ \mathrm{dlog}\lambda} = \frac{\Delta N_{M,X_{\mathrm{off}},\lambda}}{V\ s_M\ s_{X_{\mathrm{off}}}\ s_\lambda}, \tag{3.13}$$

where $\Delta N_{M,X_{\mathrm{off}},\lambda}$ is the number halos in each mass, X_{off} and λ bin, V is the total volume of the simulated cube and s_M, $s_{X_{\mathrm{off}}}$, s_λ are the natural (base 10) logarithm of mass $(X_{\mathrm{off}}, \lambda)$ binning. Equivalently to equation 3.12, we calculate

$$h(\sigma, X_{\mathrm{off}}, \lambda) = \frac{dn}{\mathrm{dln}M\ \mathrm{dlog}X_{\mathrm{off}}\ \mathrm{dlog}\lambda}\frac{M}{\rho_m}\left(\frac{dln\sigma^{-1}}{dlnM}\right)^{-1}. \tag{3.14}$$

We consider a single integration of $h(\sigma, X_{\mathrm{off}}, \lambda)$, which results in the set of Eqs. 3.15. The notation g_X designates the marginalization of $h(\sigma, X_{\mathrm{off}}, \lambda)$ over the variable X.

$$\begin{aligned} g_\lambda(\sigma, X_{\mathrm{off}}) &= \int h(\sigma, X_{\mathrm{off}}, \lambda)d\lambda, \\ g_{X_{\mathrm{off}}}(\sigma, \lambda) &= \int h(\sigma, X_{\mathrm{off}}, \lambda)dX_{\mathrm{off}}, \\ g_\sigma(X_{\mathrm{off}}, \lambda) &= \int h(\sigma, X_{\mathrm{off}}, \lambda)d\sigma, \end{aligned} \tag{3.15}$$

Integrating again, we obtain

$$\begin{aligned} f_{X_{\mathrm{off}},\lambda}(\sigma) &= \int g_\lambda(\sigma, X_{\mathrm{off}})dX_{\mathrm{off}} = \int g_{X_{\mathrm{off}}}(\sigma, \lambda)d\lambda, \\ f_{\sigma,\lambda}(X_{\mathrm{off}}) &= \int g_\lambda(\sigma, X_{\mathrm{off}})d\sigma = \int g_\sigma(X_{\mathrm{off}}, \lambda)d\lambda, \\ f_{\sigma,X_{\mathrm{off}}}(\lambda) &= \int g_{X_{\mathrm{off}}}(\sigma, \lambda)d\sigma = \int g_\sigma(X_{\mathrm{off}}, \lambda)dX_{\mathrm{off}}, \end{aligned} \tag{3.16}$$

The functions f, g, h are thus linked by derivatives as follows:

$$\begin{aligned} g(X, Y) &= \frac{\partial f(X)}{\partial Y}, \\ h(X, Y, Z) &= \frac{\partial^2 f(X)}{\partial Y \partial Z} = \frac{\partial g(X, Y)}{\partial Z}, \end{aligned} \tag{3.17}$$

where X, Y, Z are permutations of the variables $\sigma, X_{\mathrm{off}}, \lambda$. With this method we recover the multiplicity function f(σ), which in this notation is $\mathrm{f}_{X_{\mathrm{off}},\lambda}(\sigma)$. This allows studying the behavior of the dark matter halo mass function according to different variables, making sure that in the end, our analysis provides an accurate multiplicity function.

Tabelle 3.3: Cosmic variance in different MD simulations.

simulation	cosmic variance
HMD	0.02
BigMD	0.03
MDPL	0.04

3.4.3 Mass – Xoff – λ function, $h(\sigma, X_{\rm off}, \lambda)$

Here we present a model for the generalized mass function $h(\sigma, X_{\rm off}, \lambda)$ introduced in the previous subsection. The relaxation state of a dark matter halo is related to the values of $X_{\rm off}$ and λ parameters. We consider HMD, BigMD, MDPL2 to build a 3D histogram of halo counts in bins of $\sigma, X_{\rm off}, \lambda$, according to equations 3.13 and 3.14. Mass functions are expressed as multiplicity functions $f(\sigma)$. It allows the inclusion of part of the redshift evolution in σ. We focus on the high-mass end of the mass function, using halos with $M > 2.7 \times 10^{13} M_{\odot}/h$. So, we measure the halo number density in bins of $\log_{10} \sigma^{-1}$ instead of mass. We consider linear spaced bins from -0.09 to 0.6, with 0.01 width, corresponding to values close to 2.8×10^{13} and 1.3×10^{16} $M_{\odot}/h$. We consider 50 bins spanning logarithmically from $10^{-3.8}$ to $10^{-0.2}$ for $X_{\rm off}$ and 50 bins spanning logarithmically from $10^{-4.5}$ to $10^{-0.1}$ for λ. This is almost three orders of magnitudes higher than the mass resolution in HMD ($7.9 \times 10^{10} M_{\odot}/h$). Therefore, we will not be impacted by the mass resolution of the simulations. The total sample consists of 8,051,654 halos for HMD, 2,103,896 for BigMD and 142,527 for MDPL. First, we estimate directly $f(\sigma)$ for each simulation, according to Equations 3.11 and 3.12. Then, we estimate h by computing a 3D histogram in the same mass bins and different $X_{\rm off}$ and λ bins. Uncertainties on histogram values are computed considering a Poisson-number count term and a cosmic variance term, according to Equation 3.18.

$$\begin{aligned} \delta h(\sigma, X_{\rm off}, \lambda) &= h(\sigma, X_{\rm off}, \lambda) \sqrt{\frac{1}{N_{M,Xoff,\lambda}} + C^2}, \\ \delta \log_{10} h(\sigma, X_{\rm off}, \lambda) &= \frac{1}{\ln 10} \frac{\delta h(\sigma, X_{\rm off}, \lambda)}{h(\sigma, X_{\rm off}, \lambda)}, \end{aligned} \tag{3.18}$$

where C is a term accounting for cosmic variance, which is set differently according to the type of simulation. Values for the cosmic variance are given in Table 3.3. These values are estimated by Comparat et al. (2017), using a jackknife method for variance at low masses. We fitted bins containing more than 50 halos, which according to equation 3.18 means an uncertainty around 15% and thus avoid being dominated by Poisson uncertainty.

3.5 Model

We create a single model for the $h(\sigma, X_{\rm off}, \lambda)$ function. We describe in detail its features at z=0 and its redshift evolution.

3.5.1 Redshift zero

We consider that both λ and $X_{\rm off}$ probability density functions are described by a modified Schechter function (Eq. 3.8 and 3.10), and combine these two functions with the multiplicity function along the mass axis. We obtain the model described in Eq. 3.19.

$$\begin{aligned} &h(\sigma, X_{\rm off}, \lambda, z, A, a, q, \mu, \alpha, \beta, \gamma, \delta, e) = ... \\ &A\sqrt{\frac{2}{\pi}}\left(\sqrt{a}\frac{\delta_c}{\sigma}\right)^q \exp\left[-\frac{a}{2}\frac{\delta_c^2}{\sigma^2}\right]\left(\frac{X_{\rm off}}{\mu'}\right)^\alpha ... \\ &\exp\left[-\left(\frac{X_{\rm off}}{\mu'}\right)^{0.05\alpha}\right]\left(\frac{\lambda}{\mu}\right)^\gamma \exp\left[-\left(\frac{X_{\rm off}}{\mu'\sigma^e}\right)^\beta\left(\frac{\lambda}{\mu}\right)^\delta\right], \end{aligned} \tag{3.19}$$

with $\mu' = 10^{1.83\log_{10}\mu}$, to disentangle the degeneracy between the two knees of the modified Schechter functions. This model recalls the Bhattacharya et al. (2011) formulation along the mass axis and considers the combination of a power-law and an exponential cutoff (i.e. a modified Schechter function) along the $X_{\rm off}$ and λ axis. The last exponential contains crossed terms between $X_{\rm off}$ and λ, which allows taking into account their correlation. Both $X_{\rm off}$ and λ modified Schechter functions do not have mass dependency (as suggested by Equations 3.8, 3.10). In the Bhattacharya et al. (2011) formulation, there is a double power-law, here we consider a single σ power-law. Additional σ dependencies are described by crossed mass-dependent terms in the exponential cutoff, which relates $X_{\rm off}$ and λ modified Schechter functions. The position of the knee of the $X_{\rm off}$ function, i.e. μ', is the same in its two exponential cutoffs, but in the second one, we introduce the scaling with mass, through the parameter e. We correlate directly the knees of the modified Schechter functions for λ and $X_{\rm off}$ (μ, $\mu' = 10^{1.83\log_{10}\mu}$), and the slopes of the power-law and exponential cutoff of $X_{\rm off}$ (α, 0.05α), to write the model in the most compact possible way.

3.5.2 Evolution with redshift

Since structures accrete matter with time and grow, the mass function depends on redshift (Springel et al., 2005). Part of this redshift evolution is in the mass-σ relation through the matter power spectrum (Equation 1.25). However, this does not make the mass function completely universal at different times. Tinker et al. (2008) showed that a spherical overdensity mass function evolves up to 30% from z=0 to z=2.5. Despali et al. (2016) highlighted how only virial overdensity nears the universality for the mass function. Crocce et al. (2010) considered a friend-of-friends mass function and found a 10% evolution up to z=2. In this work we use ROCKSTAR to identify halos, its base is a FOF algorithm as well. Departures from universality are partially explained by the cosmology dependence of the mass function on the power spectrum and growth rate (Ondaro-Mallea et al., 2022). We find that the distribution functions of $X_{\rm off}$ and λ show a redshift trend as well (see Sect. 3.3 and Equations 3.8, 3.10). We provide a detailed description of the evolution of $h(\sigma, X_{\rm off}, \lambda)$ with redshift. Further investigation, using simulations in different cosmologies, is needed

to assess a possible relation between parameters in Equation 3.20 and cosmological parameters.

Given our model at z=0, we use the latter as the benchmark model to fit the halo mass$-X_{\rm off}-\lambda$ function at a higher redshift. For this goal, we concatenate again samples from HMD, BigMD, and MDPL2 at redshift 0.045, 0.117, 0.221, 0.425, 0.523, 0.702, 0.779, 1.032, 1.425. We note that BigMD is not tabulated at the same exact redshift snapshots as the other two simulations. Nonetheless, we use snapshots as close as possible, resulting in a 1.3% difference for the worst-case scenario at z=1.425. Further details about the snapshots are available in Appendix 3.7. For all these snapshots we consider the same $X_{\rm off}$ and λ binning as we did for z=0. However, we slightly shift upwards the σ binning compared to the z=0 case, allowing us to reach masses of $7\times10^{12}\ M_\odot/h$ at z=0.702 and $10^{12}\ M_\odot/h$ at z=1.425.

We include a redshift evolution for all the parameters A, a, q, μ, α, β, γ, δ, and e. We stress that we did not consider an evolving critical density contrast with redshift, fixing it at z=0. So δ_c in Equation 3.19 is fixed at the value of 1.68647. Considering its evolution, even if tiny, introduces the need for additional evolution of the parameters, as pointed out by Bhattacharya et al. (2011). We model the redshift evolution for these parameter using exponents k_0, k_1, k_2, k_3, k_4, k_5, k_6, k_7, k_8 as follows in Equation 3.20:

$$\begin{aligned}
\log_{10} A(z) &= \log_{10} A_0(1+z)^{k_0},\\
a(z) &= a_0(1+z)^{k_1},\\
q(z) &= q_0(1+z)^{k_2},\\
\log_{10} \mu(z) &= \log_{10} \mu_0(1+z)^{k_3},\\
\alpha(z) &= \alpha_0(1+z)^{k_4},\\
\beta(z) &= \beta_0(1+z)^{k_5},\\
\gamma(z) &= \gamma_0(1+z)^{k_6},\\
\delta(z) &= \delta_0(1+z)^{k_7},\\
e(z) &= e_0(1+z)^{k_8},
\end{aligned} \tag{3.20}$$

3.6 Results

We present the result of the fits to the data (Sect. 3.2) and the parameters of the model (Sect. 3.5). We fit directly $\log_{10} h(\sigma, X_{\rm off}, \lambda)$, which allows better modeling of the high-mass end. We consider a Gaussian likelihood

$$\log \mathcal{L} = -0.5 \sum \left(\frac{D-M}{E}\right)^2, \tag{3.21}$$

where D is $\log_{10} h(\sigma, X_{\rm off}, \lambda)$ computed from Equation 3.13, M is the base 10 logarithm of the model (Equation 3.19) and E is the uncertainty of $\log_{10} h(\sigma, X_{\rm off}, \lambda)$ (see log error in Equation 3.18).

3.6.1 Redshift zero

The best-fit parameters at z=0 are obtained maximizing the likelihood in equation 3.21. We derive posterior probability distributions and the Bayesian evidence with the nested sampling Monte Carlo algorithm MLFriends (Buchner, 2014, 2019), using the UltraNest [1] software. The results are shown in Fig. 3.12. We used flat priors. The description of the parameters, priors, and posteriors is summarized in Table 3.7. We obtain $\log_{10} A = -22.004 \pm 0.006$, $a = 0.885 \pm 0.004$, $q = 2.284 \pm 0.016$, $\log_{10} \mu = -3.326 \pm 0.001$, $\alpha = 5.623 \pm 0.002$, $\beta = -0.391 \pm 0.001$, $\gamma = 3.024 \pm 0.003$, $\delta = 1.209 \pm 0.001$, $e = -1.105 \pm 0.005$. The parameter a is in agreement with Bhattacharya et al. (2011). Our q parameter shows higher values, but this is expected because an additional mass trend is described by the knee of the exponential cutoff with mixed $X_{\rm off}, \lambda$ terms. α, γ describe the power-law increment from small $X_{\rm off}, \lambda$ values, respectively. The second one is similar to the values computed for the spin distribution in Sect. 3.3, while the first one is bigger than almost a factor two. This is expected because an additional offset trend is described by the negative β parameter. Moreover, together with β, the parameter e allows accounting for the relation between offset and spin, including mass dependency as well. The negative e allows the shifting of the peak along the $X_{\rm off}$ axis to higher values with mass, according to the findings in Sect. 3.3.

Since this is a 3D model, we show in Fig. 3.8 all six combinations of h(σ,$X_{\rm off}$,λ) integrated in 1D (Equation 3.15). To perform the integrals we exploit the Simpson's rule method for numerical integration [2]. We show each 2D distribution in five different slices of a single quantity. We recover the typical exponential cutoff along the mass axis, as well as the modified Schechter shapes for offset and spin. Our model describes σ and λ evolution very well, $g_{X_{\rm off}}(\sigma, \lambda)$ agrees with the data also in the tails of the distribution. We notice small deviations in the distribution of $X_{\rm off}$ when $X_{\rm off}$ values approach the spatial resolution limit. For low redshift samples, each version of MultiDark has its own resolution limit: 25 kpc/h for HMD, 10 kpc/h for BigMD, 5 kpc/h for MDPL2. The mass trend of the low $X_{\rm off}$ slice is slightly underpredicted by the model at low mass. The same holds for the spin trend: there is a 0.1 dex, 3.1σ tension between the peak values of model and data. This is expected within the resolution limit. An improvement toward efficient computation and future generation of N-body simulations will be needed to probe the kpc scales of Dark Matter halos in simulation cubes of the Gpc scale. All the panels in Fig. 3.8 involving σ show different uncertainties between succeeding mass bins as a result of the concatenation of bins from different MultiDark versions. MDPL2 bins contain fewer halos than BigMD, which contain fewer halos than HMD. This translates into smaller

[1] https://johannesbuchner.github.io/UltraNest/
[2] https://docs.scipy.org/doc/

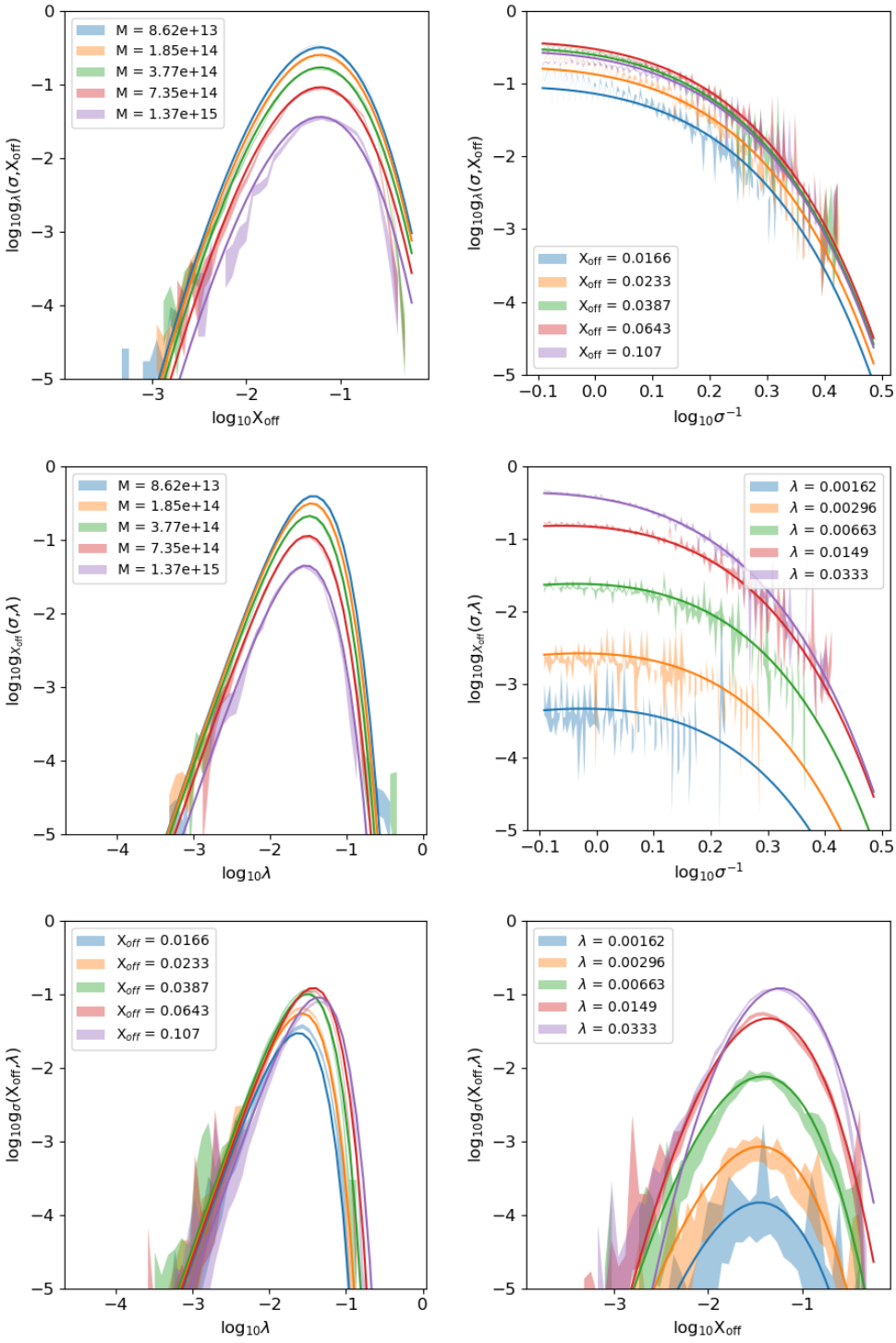

Abbildung 3.8: Single integration of the 3D model. In each panel straight lines indicate the best-fit model, while shaded areas represent the data with 1σ uncertainties. Top left panel: $g_\lambda(\sigma, X_{\rm off})$ as a function of $X_{\rm off}$ in different mass slices. Top right panel: $g_\lambda(\sigma, X_{\rm off})$ as a function of σ in different $X_{\rm off}$ slices. Middle left panel: $g_{X_{\rm off}}(\sigma, \lambda)$ as a function of λ in different mass slices. Middle right panel: $g_{X_{\rm off}}(\sigma, \lambda)$ as a function of σ in different λ slices. Bottom left panel: $g_\sigma(X_{\rm off}, \lambda)$ as a function of λ in different $X_{\rm off}$ slices. Bottom right panel: $g_\sigma(X_{\rm off}, \lambda)$ as a function of $X_{\rm off}$ in different λ slices. The integrals are defined in Equations 3.15.

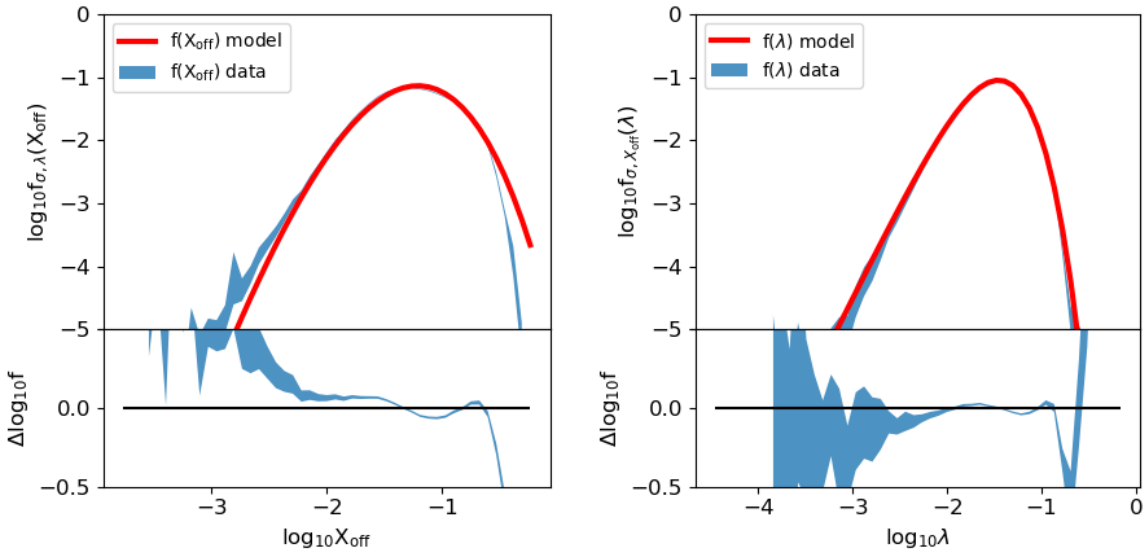

Abbildung 3.9: $f_{\sigma,\lambda}(X_{\rm off})$ and $f_{\sigma,X_{\rm off}}(\lambda)$ comparison between data and model. In top panels straight red lines indicate the integral on the best-fit model, while shaded blue areas represent the integral on the 3d $h(\sigma, X_{\rm off}, \lambda)$ data with 1σ uncertainties. Each bottom panel shows the residual trend with σ error, the straight black line represents the perfect match between data and model, with null residual. Top left panel: $f(X_{\rm off})$ as a function of $X_{\rm off}$. Bottom left panel: residual between $f_{\sigma,\lambda}(X_{\rm off})$ data and model in logarithmic scale. Top right panel: $f(\lambda)$ as a function of λ. Bottom right panel: residual between $f_{\sigma,X_{\rm off}}(\lambda)$ data, and model in logarithmic scale.

uncertainties for bins containing a higher number of halos. Overall, the model provides an excellent representation of the data.

We further integrate the model in Equation 3.19, obtaining the distributions of $X_{\rm off}$ $f_{\sigma,\lambda}(X_{\rm off})$ and λ $f_{\sigma,X_{\rm off}}(\lambda)$ (Equation 3.16). Figure 3.9 shows the result. The distributions around the peaks are well described by the model. This is important because these functions are dominated by objects described by the peak of the PDF. The spin distribution is better behaved than the $X_{\rm off}$ in the tails. This is expected due to spatial resolution limits. Moreover, this is a further confirmation of the fact that $X_{\rm off}, \lambda$ scatter around their mean value with modified Schechter distributions (see Sect. 3.3).

We obtain the multiplicity function $f_{X_{\rm off},\lambda}(\sigma)$ marginalizing $h(\sigma, X_{\rm off}, \lambda)$ on $X_{\rm off}, \lambda$, i.e. performing the double integral

$$f_{X_{\rm off},\lambda}(\sigma) = \int\int h(\sigma, X_{\rm off}, \lambda) dX_{\rm off} d\lambda.$$

The result is shown in Fig. 3.10. In the top panel, we show a comparison between our model, the data obtained from simulations, and the Comparat et al. (2017) model, fitted on these same simulations at z=0. The multiplicity functions computed directly on each simulation cube, without taking $X_{\rm off}$ and λ into account, are shown by three shaded regions in different

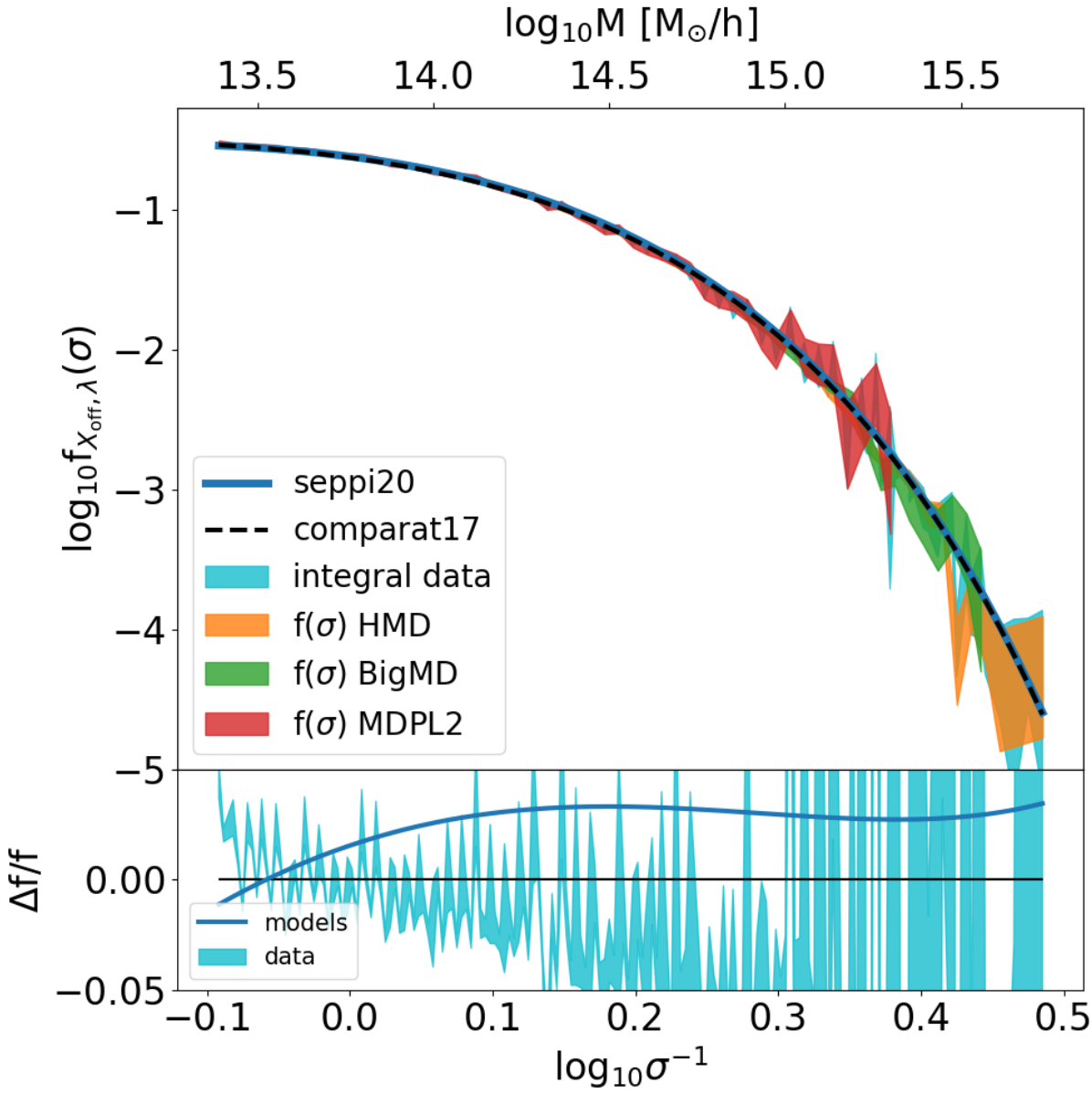

Abbildung 3.10: Multiplicity functions comparison. Top panel: three shaded regions show the 1σ contours of $f(\sigma)$ data directly computed on different simulations (orange for HMD, green for BigMD, red for MDPL2), the light blue shaded region is the 1σ contour of the 2D integral computed on the concatenated sample containing all three simulations, the dashed pink line indicates the mass function from Comparat et al. (2017), while the blue solid line is the $f(\sigma)$ we recover integrating our model along $X_{\rm off}$ and λ. Low panel: the blue thick line is the fractional difference between our $f(\sigma)$ and the Comparat et al. (2017) one. The light blue shaded area denotes the 1σ contours of the residual between the integrated data and our the best-fit model, the black horizontal line indicates the perfect match with null residual.

colors: red for MDPL2, green for BigMD, and orange for HMD. Bigger simulation boxes extend to higher mass values. The light blue shaded region represents the 2D integral computed on the concatenated sample of all three simulations. The solid blue line is the integral of our model and the dashed pink line is the Comparat et al. (2017) model. In the lower panel, we show the percentage difference between the multiplicity function $f(\sigma)$ we recover and the Comparat et al. (2017) model, obtained on the same MultiDark simulations. This difference is always under 3.3% in the mass range of interest. It is also compatible with uncertainty on the data. Our model is able to recover the halo mass distribution in the simulations, with the advantage of taking into account parameters that describe their dynamical state as well. Once again, we stress that our model is adapted to masses higher than $2.7 \times 10^{13} M_{\odot}/h$ at z=0 ($10^{12} M_{\odot}/h$ at z=1.4).

3.6.2 Evolution with redshift

To study the redshift evolution we start from the fiducial model at z=0. We add the redshift dependence (Equation 3.20) to the best-fit parameters in Equation 3.19 at z=0. We concatenate samples for 10 redshift values as described in 3.5.2.

We obtain the values of the exponents in equation 3.20 fitting the z trend of each parameter for all the concatenated snapshots simultaneously. We obtain $k_0 = -0.0441 \pm 0.0001$, $k_1 = -0.161 \pm 0.001$, $k_2 = 0.041 \pm 0.002$, $k_3 = -0.1286 \pm 0.0002$, $k_4 = 0.1081 \pm 0.0002$, $k_5 = -0.311 \pm 0.001$, $k_6 = 0.0902 \pm 0.0004$, $k_7 = -0.0768 \pm 0.0004$, $k_8 = 0.612 \pm 0.002$. The full result is shown by the triangular plot in Fig. 3.13. Priors and posteriors for each parameter are given in Table 3.8.

Redshift dependence is shown in Fig. 3.11. The shaded areas include the uncertainty on both the best-fit parameter at z=0 and on the z evolution, according to equation 3.22

$$\delta P = \left[\left(\frac{\partial P}{\partial P_0} \delta P_0 \right)^2 + \left(\frac{\partial P}{\partial k} \delta k \right)^2 \right]^{1/2}, \tag{3.22}$$

where P is each parameter in equation 3.19, P_0 is its value at z=0 and k indicates each parameter describing the evolution in equation 3.20.

The parameters A,q,μ,α,γ,β show an increasing redshift trend. On the other hand a,β, and e decrease with redshift. This means that with increasing redshift, the modified Schechter functions need to increase quicker and decrease slower. The knee describing the mass trend (parameter a) decreases with redshift, in agreement with Bhattacharya et al. (2011). They find no redshift dependence for the slope of mass trend (parameter q). This is not true for this work, where q increases with z. However, this is mitigated by the mass trend of the position of the knee in the crossed $X_{\rm off}, \lambda$ exponential cutoff, which decreases at early times. The fact that the position of the $X_{\rm off}$ knee (parameter μ) moves to higher values at high z confirms the results of Sect. 3.3, with the higher average value of $X_{\rm off}$ early in time (Figures 3.6, 3.7).

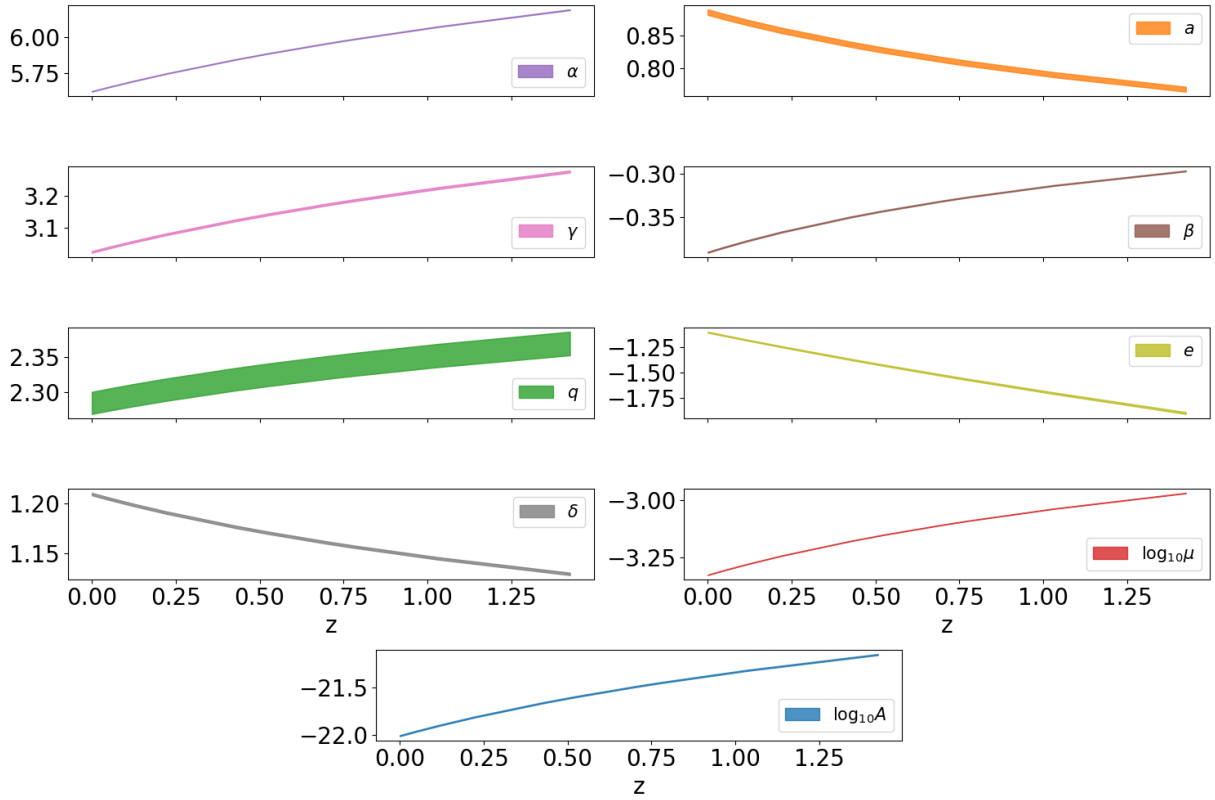

Abbildung 3.11: Redshift evolution of the best-fit parameters of our model. Each panel shows a single parameter. The values at z=0 are reported in Table 3.7. The redshift evolution is described by Equation 3.20, the slopes are given in Table 3.8.

Tabelle 3.4: best-fit parameters for concentration-σ relation and its PDF P(c).

		a_0			b_0		
		0.754091 ± 0.000004			0.574413 ± 0.000002		
	A	α	β	x_0	e_0	e_1	e_2
z=0	0.041 ± 0	1.397 ± 0.001	2.604 ± 0.001	7.225 ± 0.002	0.089 ± 0.001	0.776 ± 0.001	-0.959 ± 0.001
z=0.52	0.044 ± 0	2.501 ± 0.001	1.283 ± 0.001	2.394 ± 0.001	0.579 ± 0.001	-0.325 ± 0.001	0.0334 ± 0.001
z=1.03	3.37e-3 ± 0	4.14 ± 0.12	0.927 ± 0.002	0.688 ± 0.025	0.188 ± 0.002	0.081 ± 0.001	0.0813 ± 0.001
z=1.43	2.54e-3 ± 0	4.45 ± 0.13	0.924 ± 0.002	0.623 ± 0.031	0.198 ± 0.003	0.095 ± 0.001	0.103 ± 0.001

The models are described by Eqs. 3.5 and 3.6. Uncertainties are of the percentage accuracy. In order to have compact information, when uncertainties are smaller than 4 order of magnitudes with respect to the parameter, a value of 0 is written.

Tabelle 3.5: best-fit parameters for λ-σ relation and its PDF P(λ) at different redshifts.

a_0			b_0	
4.5357e-2 ± 2e-6			-5.4328e-3 ± 1e-7	
	A	α	β	x_0
z=0	0.274 ± 0.009	3.002 ± 0.013	0.773 ± 0.001	4.33e-3 ± 0
z=0.52	1.01 ± 0.02	2.623 ± 0.004	0.911 ± 0.001	7.46e-3 ± 0
z=1.03	1.769 ± 0.009	2.409 ± 0.002	1.006 ± 0.001	9.32e-3 ± 0
z=1.43	2.089 ± 0.007	2.351 ± 0.001	1.031 ± 0.001	9.34e-3 ± 0

The models are described by Eqs. 3.7 and 3.8. Uncertainties are the percentage accuracy. In order to have compact information, when uncertainties are smaller than 4 order of magnitudes with respect to the parameter, a value of 0 is written.

3.7 Figures and Tables

In this Appendix, we collect Figures and Tables relative to this work. They describe the mean relations between concentration, offset parameter, spin, and mass; as well as the full probability density functions of these quantities. Moreover, we show additional plots describing the halo $\sigma - X_{\rm off} - \lambda$ function.

3.8 Offset in physical units

Here we present the results of the same analysis elaborated in Sections 3.3 and 3.6 to the offset parameter in physical units $X_{\rm off,P}$, measured in kpc/h. This approach allows the comparison between the physics of dark matter simulations and observations.

3.8.1 $X_{\rm off,P}$ - mass - redshift relation

The relation between $X_{\rm off,P}$, mass and redshift is modeled by

$$\log_{10} X_{\rm off,P}(\sigma, z) = \frac{b_0}{E(z)^{0.06}}\left[1 + 2.39\left(\frac{\sigma}{a_0 E(z)^{0.8}}\right)^{c_0\sigma}\right]. \tag{3.23}$$

Tabelle 3.6: best-fit parameters for $X_{\rm off}$-σ relation and its PDF P($X_{\rm off}$).

	a_0	b_0		
	-1.30418 ± 0.00001	0.15084 ± 0.00001		
	$\log_{10} A$	α	β	$\log_{10} x_0$
z=0	-3.09 ± 0.26	3.71 ± 0.12	0.64 ± 0.03	-2.31 ± 0.09
z=0.52	-2.72 ± 0.17	3.69 ± 0.03	0.69 ± 0.01	-2.11 ± 0.01
z=1.03	-2.18 ± 0.05	3.44 ± 0.02	0.77 ± 0.01	-1.88 ± 0.01
z=1.43	-1.79 ± 0.02	3.19 ± 0.01	0.84 ± 0.01	-1.70 ± 0.01

The models are described by Eqs. 3.9 and 3.10. Uncertainties on the mean relation are under the percentage level accuracy.

Tabelle 3.7: Model parameters with priors and posterior constraints at redshift zero.

Parameter	Prior	Posterior
$\log_{10} A$	(-23,-20)	$-22.004^{+0.006}_{-0.006}$
a	(0.5,1.0)	$0.885^{+0.004}_{-0.004}$
q	(1.5,2.5)	$2.284^{+0.016}_{-0.016}$
$\log_{10} \mu$	(-3.5,-3.0)	$-3.326^{+0.001}_{-0.001}$
α	(5.4,5.8)	$5.623^{+0.002}_{-0.002}$
β	(-0.5,-0.3)	$-0.391^{+0.001}_{-0.001}$
γ	(2.8,3.2)	$3.024^{+0.003}_{-0.003}$
δ	(1.0,1.4)	$1.209^{+0.001}_{-0.001}$
e	(-1.2,-0.8)	$-1.105^{+0.005}_{-0.005}$

The full distribution of the posteriors is shown in Fig. 3.12.

Tabelle 3.8: Model parameters with prior and posterior constraints for the redshift evolution of the halo $\sigma - X_{\rm off} - \lambda$ function.

Parameter	Prior	Posterior
k_0	(-0.08,0.07)	-0.0441±0.0001
k_1	(-0.25,0.05)	-0.161±0.001
k_2	(-0.05,0.15)	0.041±0.002
k_3	(-0.18,0.02)	-0.1286±0.0002
k_4	(-0.02,0.18)	0.108±0.0002
k_5	(-0.7,0.1)	-0.311±0.001
k_6	(-0.1,0.2)	0.0902±0.0004
k_7	(-0.2,0.1)	-0.0768±0.0004
k_8	(-0.05,0.85)	0.612±0.002

In the posteriors, when the 1σ uncertainty on the 1d distribution is smaller than 3 order of magnitudes with respect to the parameter, we write null error. The redshift evolution of each parameter is shown in Fig. 3.11.

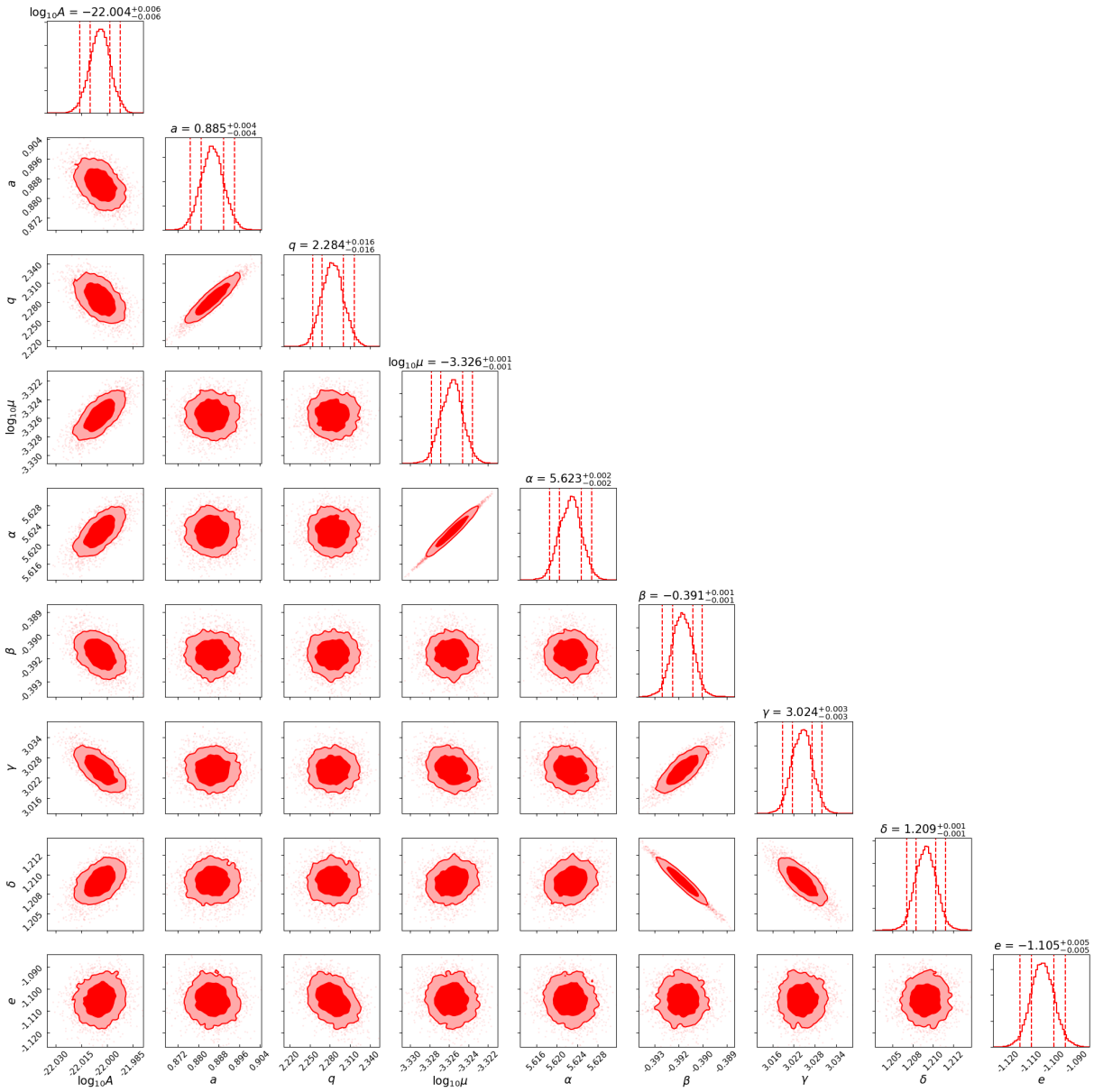

Abbildung 3.12: Marginalized posterior distributions of the best-fit parameters of the halo $\sigma - X_{\rm off} - \lambda$ function. The 0.68 and 0.95 confidence levels of the posteriors are shown as filled 2D contours. The 2.5th, 16th, 84th and 97.5th percentile of the 1-d posterior distributions are indicated by the vertical lines on the diagonal plots. The model is given by Equation 3.19. The parameters are also given in Table 3.7.

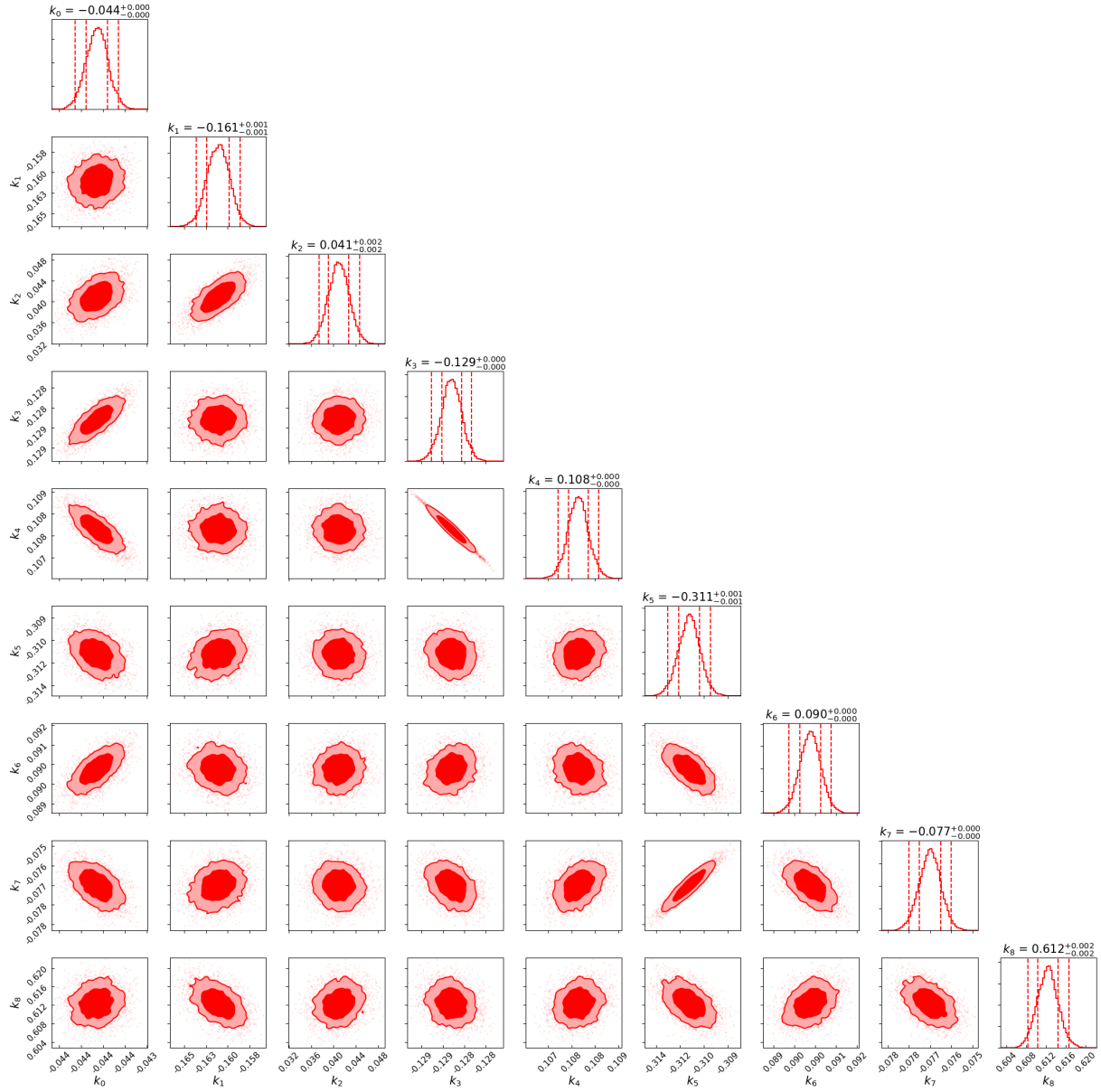

Abbildung 3.13: Marginalized posterior distributions of the best-fit parameters describing the redshift evolution of $h(\sigma, X_{\rm off}, \lambda)$. The 0.68 and 0.95 confidence levels of the posteriors are shown as filled 2D contours. The model is given by Equation 3.20. The parameters are reported in Table 3.8.

Tabelle 3.9: Full list of snapshots used, available in HMD, BigMD, MDPL2.

a	z	T(Gyr)	HMDPL	BigMD	MDPL2
1.0	0	13.82	x	x	x
0.9567	0.04526	13.19	x		x
0.956	0.04603	13.18		x	
0.8953	0.1169	12.27		x	
0.8951	0.1172	12.26	x		x
0.8192	0.2207	11.09	x		x
0.8173	0.2235	11.06		x	
0.7016	0.4253	9.198	x		x
0.7003	0.428	9.177		x	
0.6583	0.5191	8.487		x	
0.6565	0.5232	8.458	x		x
0.5876	0.7018	7.319	x		x
0.5864	0.7053	7.299		x	
0.5623	0.7785	6.90		x	
0.5622	0.7787	6.89	x		x
0.5	1	5.88		x	
0.4922	1.032	5.753	x		x
0.4123	1.425	4.482	x		x
0.409	1.445	4.431		x	

Tabelle 3.10: best-fit parameters for $X_{\rm off,P}$-σ relation and P($X_{\rm off,P}$).

		a_0	b_0	c_0	
		0.16523 ± 0.00004	0.74872 ± 0.00001	-0.39607 ± 0.00003	
	A	α	β	x_0	e_0
z=0	-5.45 ± 0.01	10.56 ± 0.87	1.23 ± 0.09	2.34 ± 0.19	-0.57 ± 0.05
z=0.52	-5.45 ± 0.01	10.58 ± 0.87	1.23 ± 0.11	2.35 ± 0.19	-0.43 ± 0.04
z=1.03	-5.35 ± 0.01	10.54 ± 0.86	1.24 ± 0.11	2.36 ± 0.19	-0.33 ± 0.03
z=1.43	-5.32 ± 0.01	10.56 ± 0.86	1.32 ± 0.11	2.36 ± 0.19	-0.46 ± 0.04

The models are described by Eqs. 3.23 and 3.24. Uncertainties on the mean relation are under the percentage level accuracy.

The distribution of $X_{\rm off,P}$ around its mean value is described by a modified Schechter function, but $X_{\rm off,P}$ is not normalized to the virial radius. Therefore, a mass dependence has to be included in the relation.

$$P(X_{\rm off,P}) = A\left(\frac{X_{\rm off,P}}{x_0\sigma^{e_0}}\right)^{\alpha}\exp\left[-\left(\frac{X_{\rm off,P}}{x_0\sigma^{e_0}}\right)^{\beta}\right]. \tag{3.24}$$

The best-fit parameters are given in Table 3.10 and the results are shown in Figures 3.14 and 3.15.

3.8.2 Mass $-$ $X_{\rm off,P}$ $-$ λ function

In this section, we collect figures and tables that describe h(σ,$X_{\rm off,P}$, λ). The analysis is similar to h(σ,$X_{\rm off}$, λ) explained in Sections 3.4 and 3.5. The only difference is that the modified Schechter function that describes $X_{\rm off,P}$ needs a mass-dependent term. Therefore, we introduce an additional parameter and model the distribution according to

$$\begin{aligned} &h(\sigma, X_{\rm off,P}, \lambda, z, A, a, q, \mu, \alpha, \beta, e_0, \gamma, \delta, e_1) = ... \\ &A\sqrt{\frac{2}{\pi}}\left(\sqrt{a}\frac{\delta_c}{\sigma}\right)^{q}\exp\left[-\frac{a}{2}\frac{\delta_c^2}{\sigma^2}\right]\left(\frac{X_{\rm off,P}}{\mu\sigma^{e_0}}\right)^{\alpha}... \\ &\exp\left[-\left(\frac{X_{\rm off,P}}{\mu\sigma^{e_0}}\right)^{0.05\alpha}\right]\left(\frac{\lambda}{0.7\mu}\right)^{\gamma} exp\left[-\left(\frac{X_{\rm off,P}}{\mu\sigma^{e_1}}\right)^{\beta}\left(\frac{\lambda}{0.7\mu}\right)^{\delta}\right], \end{aligned} \tag{3.25}$$

We recover the fiducial mass function at the $\sim 3.9\%$ level.

3.9 Summary

In the context of the hierarchical model of structure formation, the evolution of the number density of galaxy clusters is a powerful cosmological probe. In order to achieve precision cosmology with the next generation of galaxy clusters samples (such as eROSITA All Sky

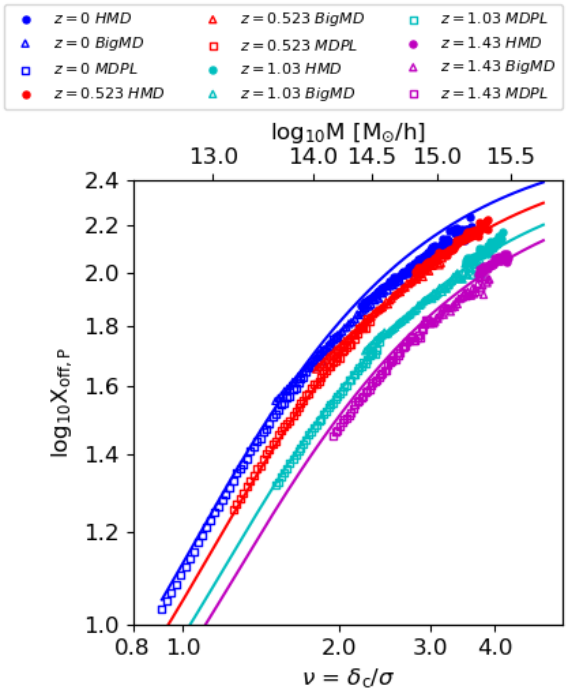

Abbildung 3.14: $X_{off,P}$-σ relation (Equation 3.23). Circular dots, triangles and squares represent HMD, BigMD, MDPL2 respectively. They are color-coded by redshift. Straight lines show the best-fit model. Parameters are given in Table 3.10.

Tabelle 3.11: h(σ,$X_{off,P}$, λ) model parameters with priors and posterior costraints.

Parameter	Prior	Posterior
$\log_{10} A$	(-23,-20)	$-22.004^{+0.009}_{-0.009}$
a	(0.5,1.0)	$0.878^{+0.004}_{-0.004}$
q	(1.5,2.5)	$2.257^{+0.013}_{-0.013}$
$\log_{10} \mu$	(-3.5,-3.0)	$-3.149^{+0.002}_{-0.002}$
α	(5.4,5.8)	$5.624^{+0.002}_{-0.002}$
β	(-0.4,-0.3)	$-0.365^{+0.001}_{-0.001}$
e_0	(-2.0,-1.4)	$-1.606^{+0.002}_{-0.002}$
γ	(2.8,3.2)	$3.095^{+0.003}_{-0.003}$
δ	(1.0,1.4)	$1.168^{+0.002}_{-0.001}$
e_1	(-3.0,-2.5)	$-2.270^{+0.005}_{-0.005}$

The full distribution of the posteriors is shown in the triangular plot in Fig. 3.20.

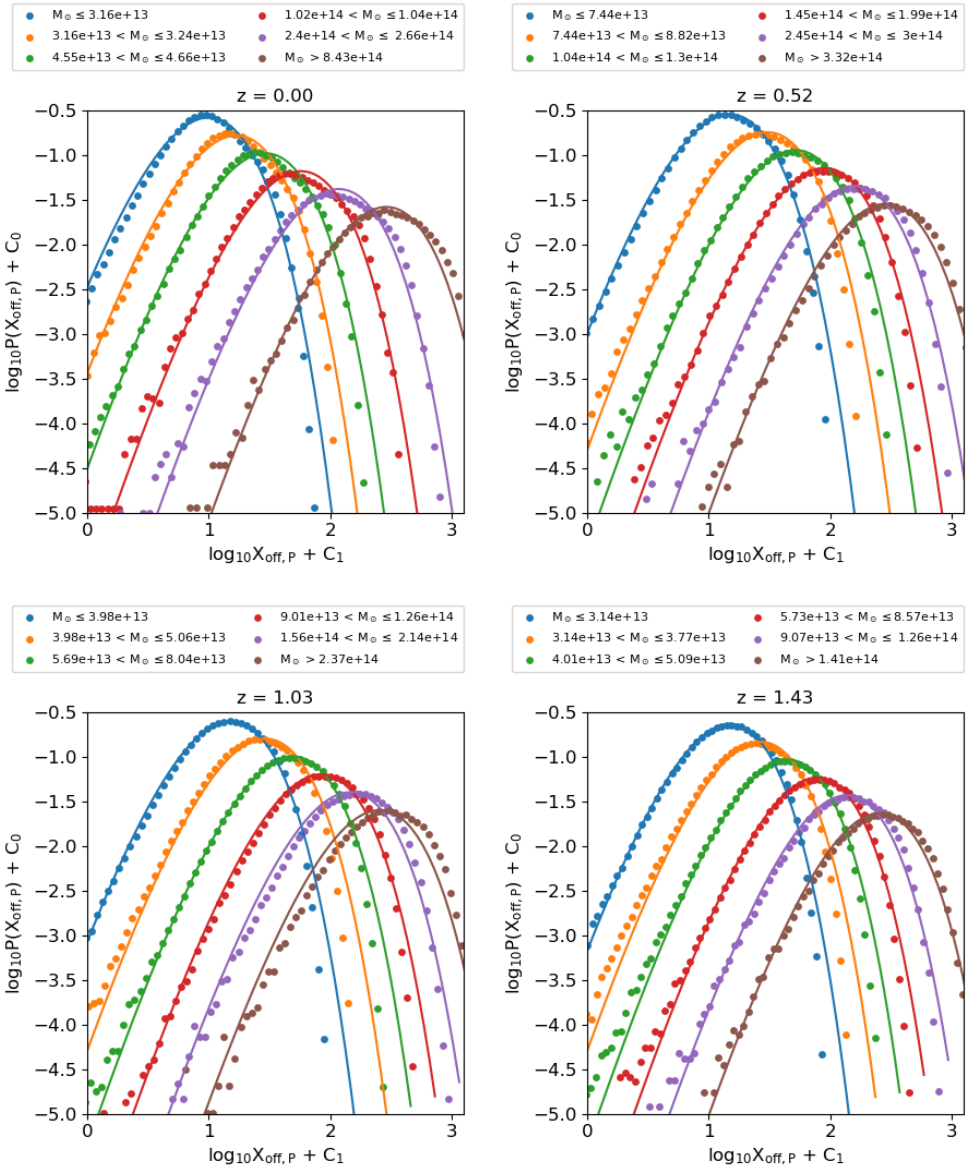

Abbildung 3.15: Probability density function of $X_{\mathrm{off,P}}$ (Equation 3.24). Each panel shows the distribution at a specific redshift. Each set is divided in mass slices, identified by color. Scatter points indicate the data, while straight lines represent the modified Schechter model 3.10. For clarity, each line and its fit are shifted by 0.2 dex along both axis. This means that both coefficients C_0 assumes values (+0.6,+0.4,+0.2,0.0,-0.2,-0.4), while C_1 is (-0.6,-0.4,-0.2,0.0,+0.2,+0.4). The red line is not shifted, therefore it is the one with the correct normalization.

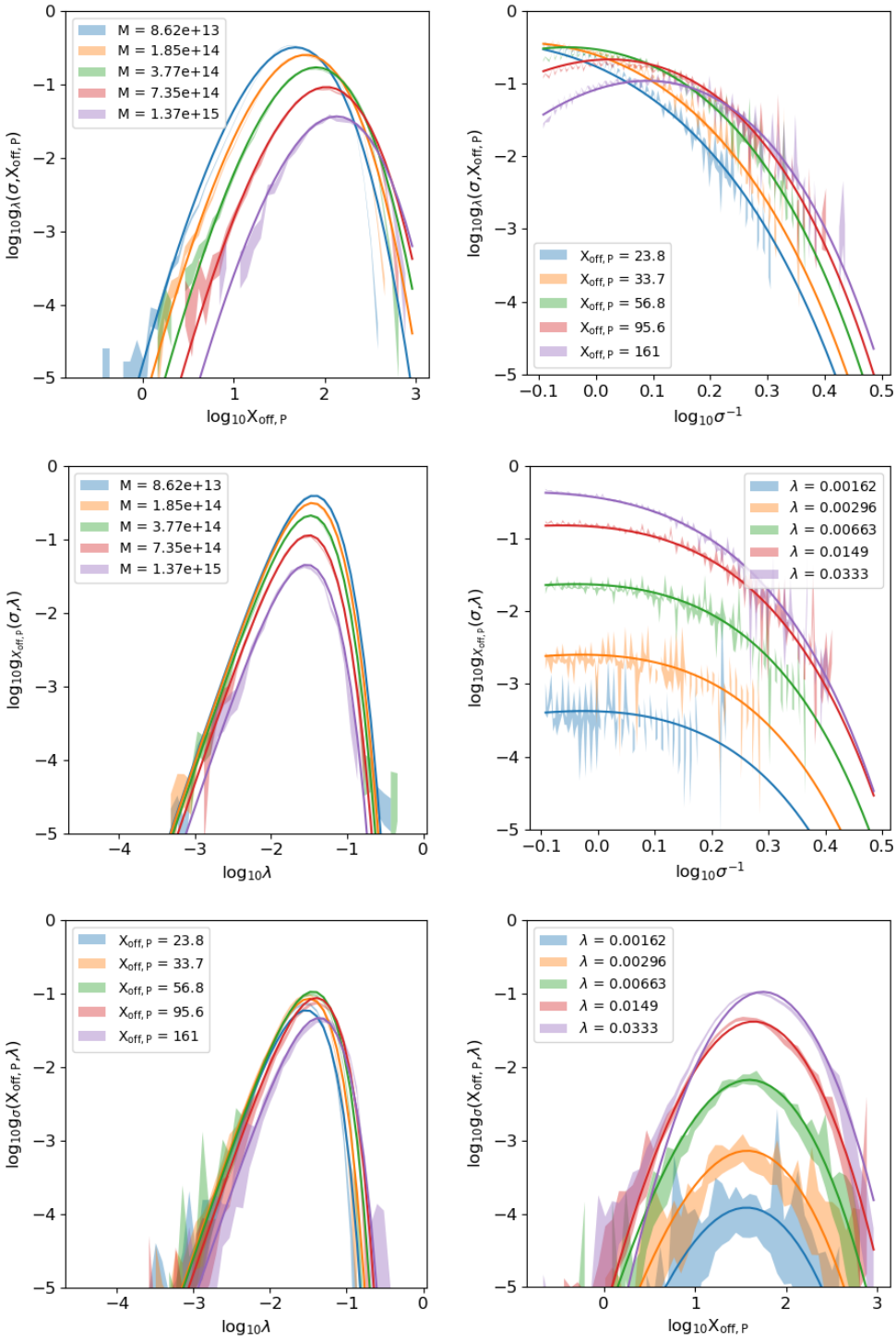

Abbildung 3.16: Single integration of the 3D model. In each panel straight lines indicate the best-fit model, while shaded areas represent the data with 1σ uncertainties. Top left panel: $g_\lambda(\sigma, X_{\rm off,P})$ as a function of $X_{\rm off,P}$ in different mass slices. Top right panel: $g_\lambda(\sigma, X_{\rm off,P})$ as a function of σ in different $X_{\rm off,P}$ slices. Middle left panel: $g_{X_{\rm off,P}}(\sigma, \lambda)$ as a function of λ in different mass slices. Middle right panel: $g_{X_{\rm off,P}}(\sigma, \lambda)$ as a function of σ in different λ slices. Bottom left panel: $g_\sigma(X_{\rm off,P}, \lambda)$ as a function of λ in different $X_{\rm off,P}$ slices. Bottom right panel: $g_\sigma(X_{\rm off,P}, \lambda)$ as a function of $X_{\rm off,P}$ in different λ slices.

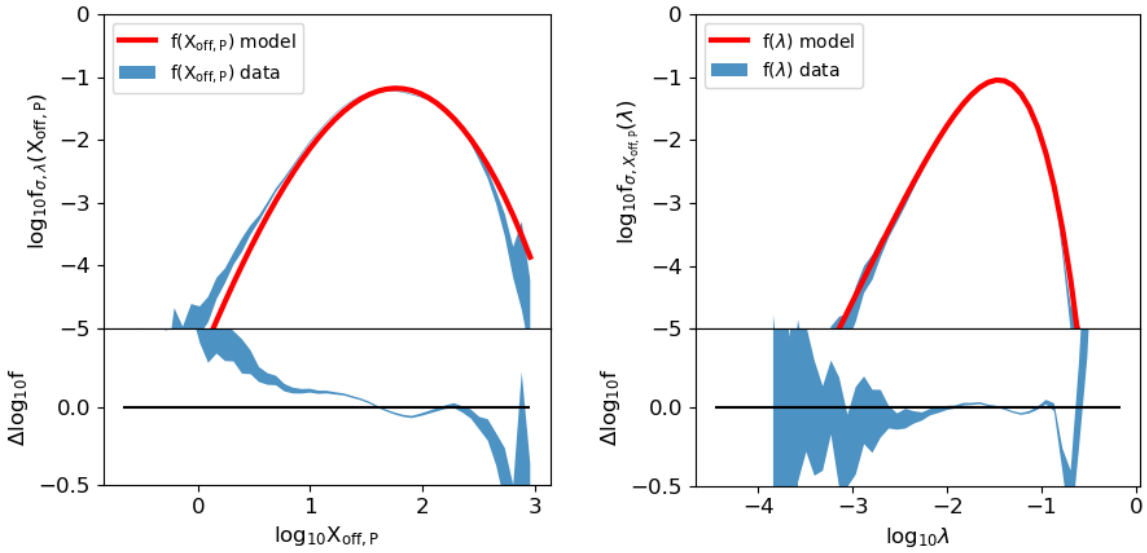

Abbildung 3.17: $f(X_{\rm off,P})$ and $f(\lambda)$ comparison between data and model. In top panels solid red lines indicate the integral on the best-fit model, while shaded blue areas represent the integral on the 3d $h(\sigma, X_{\rm off,P}, \lambda)$ data with 1σ uncertainties. Each bottom panel shows the residual trend with σ error, the straight black line represents the perfect match between data and model, with null residual. Top left panel: $f(X_{\rm off,P})$ as a function of $X_{\rm off,P}$. Bottom left panel: residual between $f(X_{\rm off,P})$ data and model in logarithmic scale. Top right panel: $f(\lambda)$ as a function of λ. Bottom right panel: residual between $f(\lambda)$ data and model in logarithmic scale.

Tabelle 3.12: Model parameters with prior and posterior constraints for the redshift evolution of h(σ,$X_{\rm off,P}$, λ).

Parameter	Prior	Posterior
k_0	(-0.08,0.07)	-0.0131±0.0001
k_1	(-0.25,0.05)	-0.146±0.001
k_2	(-0.1,0.1)	0.04±0.002
k_3	(-0.15,0.05)	-0.0716±0.0003
k_4	(-0.05,0.15)	0.0789±0.0001
k_5	(-0.7,0.1)	-0.4199±0.0005
k_6	(-0.15,0.05)	-0.0554±0.001
k_7	(-0.05,0.25)	0.1526±0.0002
k_8	(-0.25,0.05)	-0.1834±0.0004
k_9	(-0.05,0.35)	0.235±0.001

In the posteriors, when the 1σ uncertainty on the 1d distribution is smaller than 3 order of magnitudes with respect to the parameter, we write null error. The redshift evolution of each parameter is shown in Fig. 3.19.

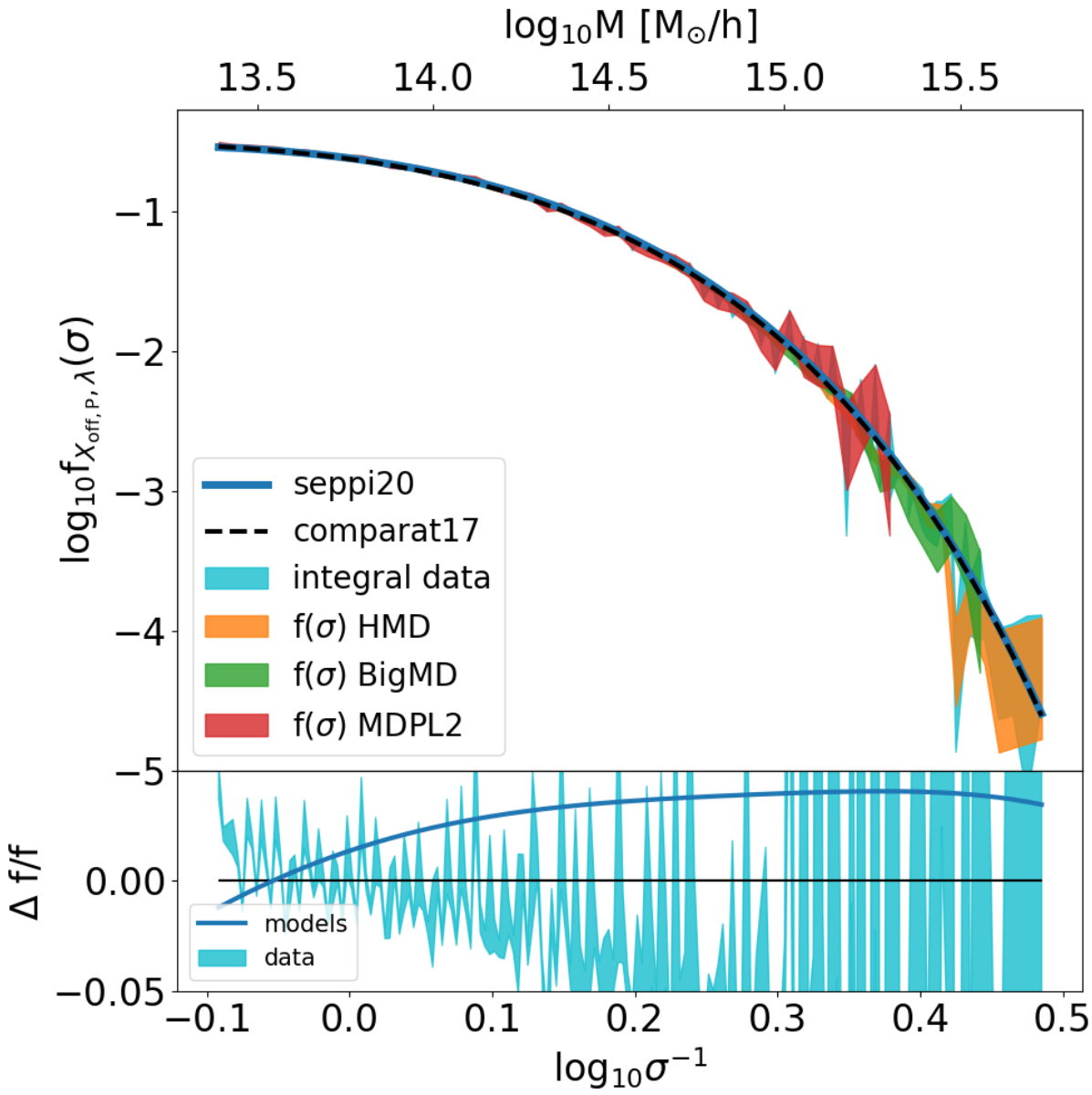

Abbildung 3.18: Multiplicty functions comparisons. Top panel: three shaded regions show the 1σ contours of $f(\sigma)$ data directly computed on different simulations (orange for HMD, green for BigMD, red for MDPL2), the light blue shaded region is the 1σ contour of the 2D integral computed on the concatenated sample containing all three simulations, the dashed pink line indicates the mass function from Comparat et al. (2017), while the blue solid line is the $f(\sigma)$ we recover integrating our model along $X_{\rm off,P}$ and λ. Low panel: the blue thick line is the fractional difference between our $f(\sigma)$ and the Comparat et al. (2017) one. The light blue shaded area represent the 1σ contours of the residual between the integrated data and our the best-fit model, the black horizontal line indicates the perfect match with null residual.

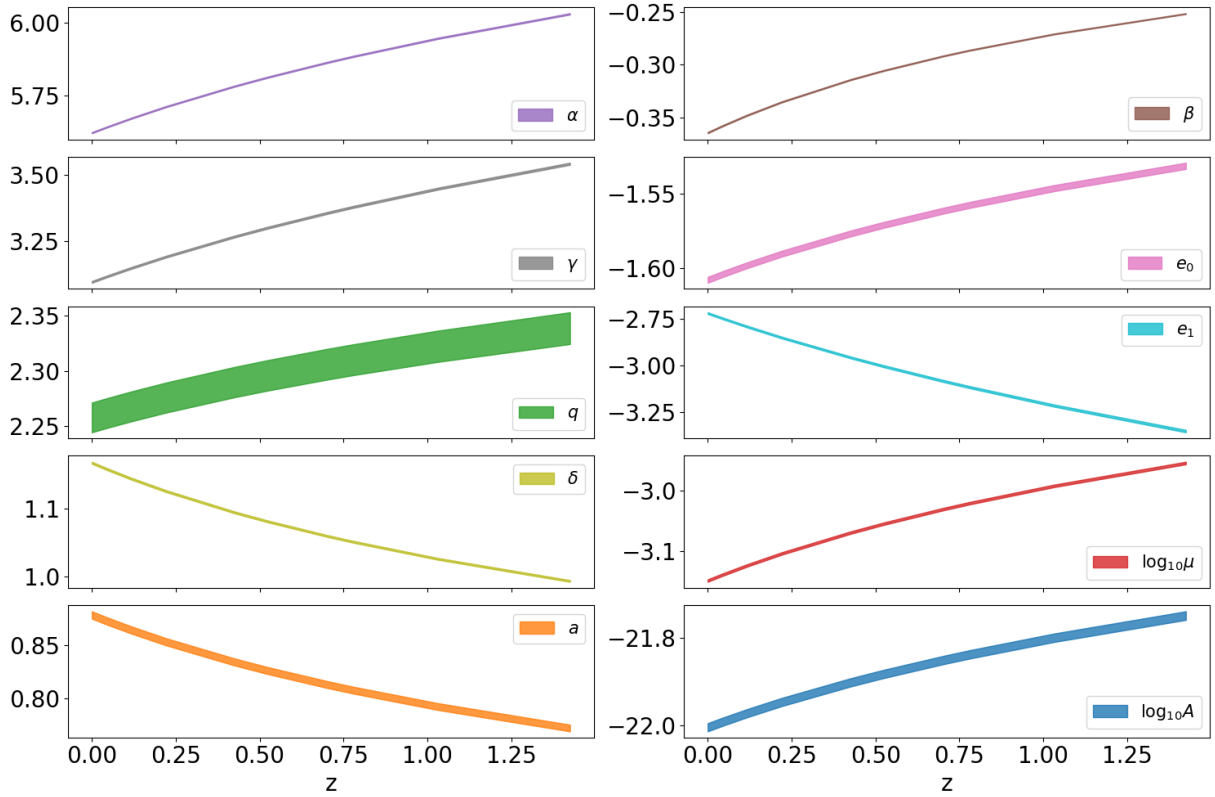

Abbildung 3.19: Redshift evolution of the best-fit parameters for h(σ,$X_{\rm off,P}$, λ). Each panel shows a single parameter. The values at z=0 are reported in Table 3.11. The slopes of the redshift trends are given in Table 3.12.

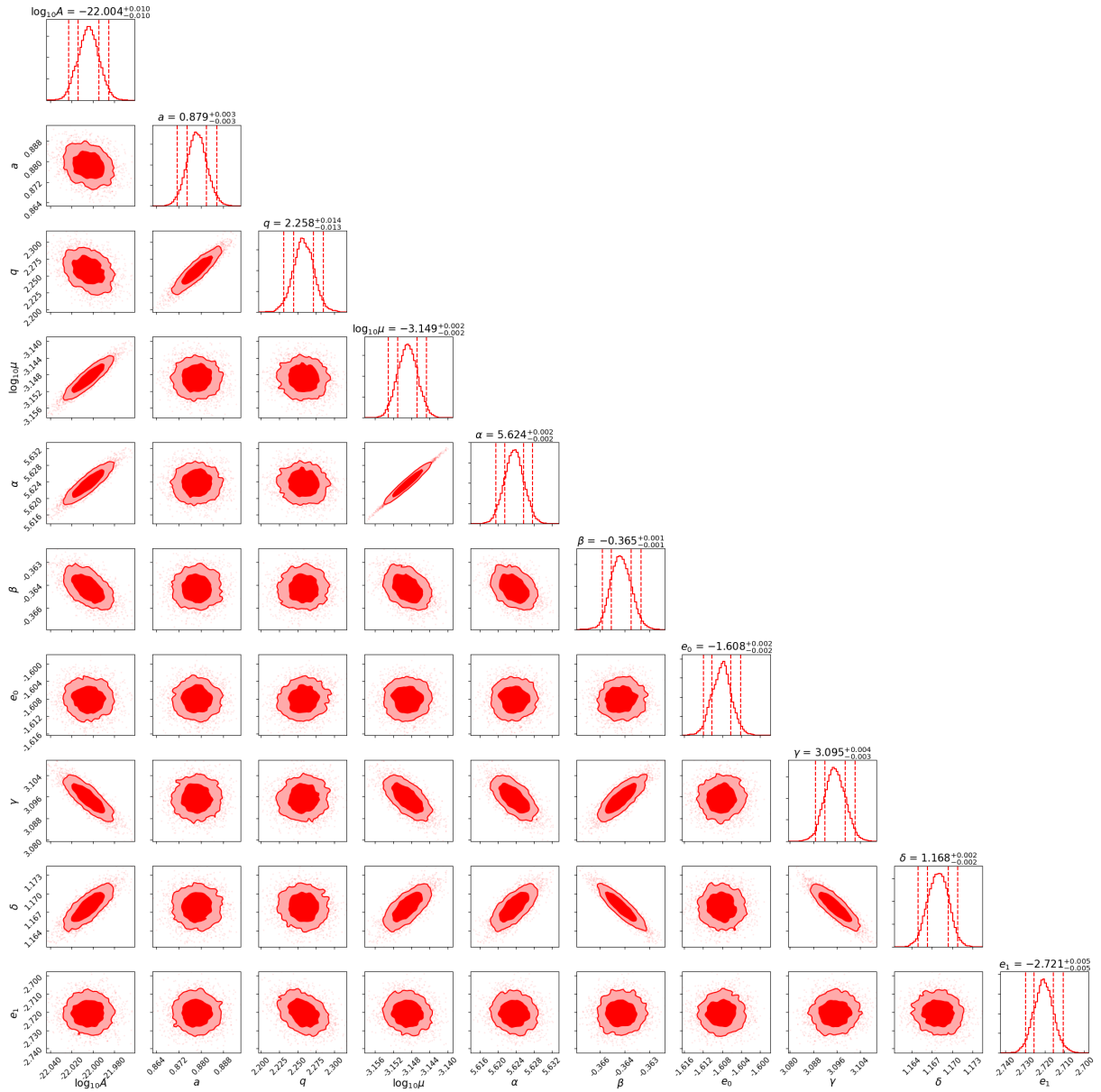

Abbildung 3.20: Marginalized posterior distributions of the h(σ,$X_{\mathrm{off,P}}$, λ) best-fit parameters at redshift 0. The 0.68 and 0.95 confidence levels of the posteriors are shown as filled 2D contours. The 2.5th, 16th, 84th and 97.5th percentile of the 1-d posterior distributions are indicated by the vertical lines on the diagonal plots.

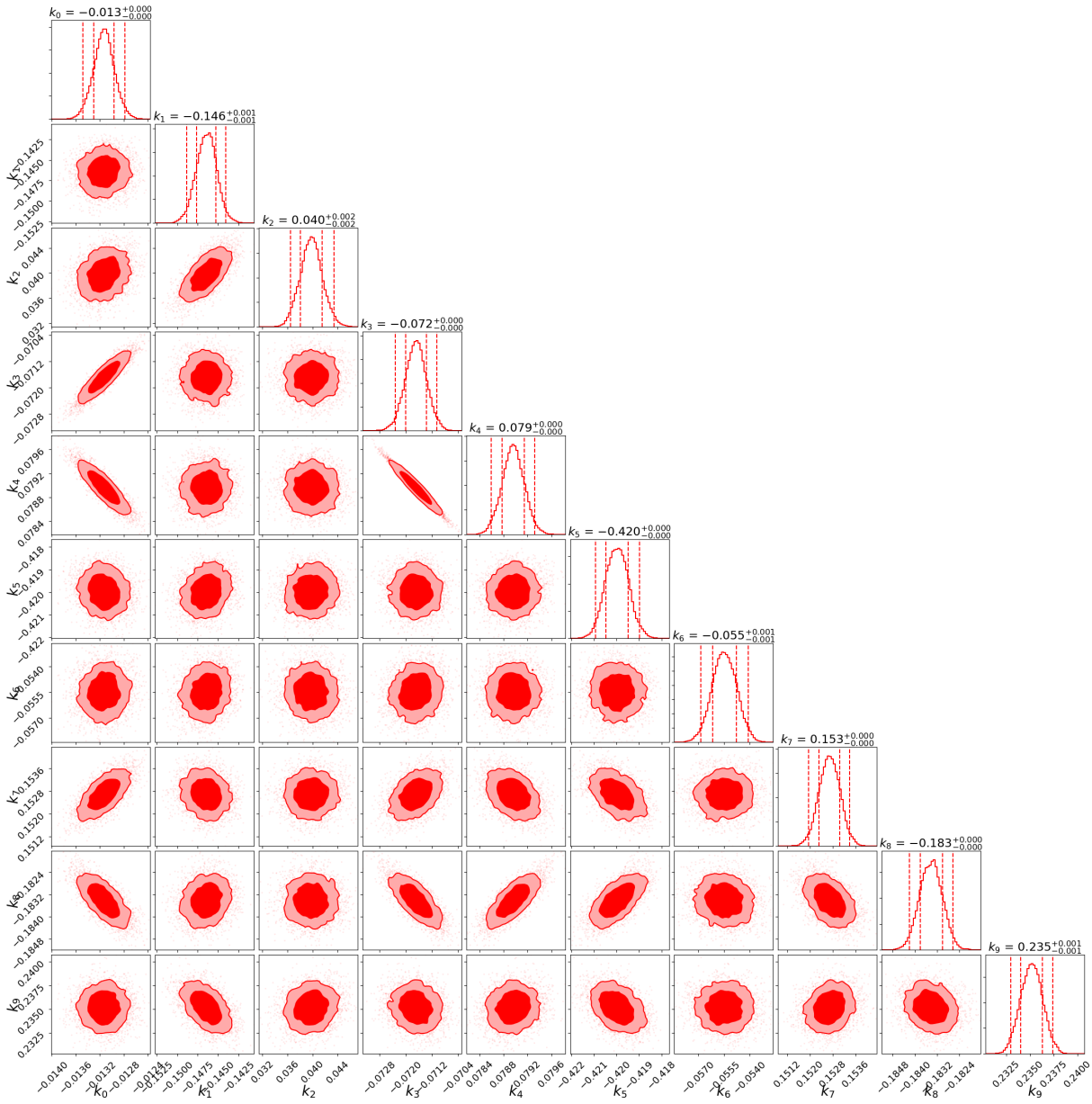

Abbildung 3.21: Marginalized posterior distributions of the best-fit parameters describing the redshift evolution of $h(\sigma, X_{\mathrm{off,P}}, \lambda)$. The 0.68 and 0.95 confidence levels of the posteriors are shown as filled 2D contours. The 2.5th, 16th, 84th and 97.5th percentile of the 1-d posterior distributions are indicated by the vertical lines on the diagonal plots.

Survey Merloni et al. (2012)), precise modeling of the theoretical mass function is necessary. We calibrate a model that includes a dynamical description of dark matter halos. Using the formalism described here and the MultiDark simulations, we quantify the impact on the mass function of the lack of unrelaxed structures (see Fig. 3.1). We explore relations between quantities that describe different aspects of dark matter halos, including their dynamical state. We investigate the concentration-mass relation. We confirm the recent discovery of the concentration upturn at high masses, in agreement with previous results from Prada et al. (2012) and Klypin et al. (2016), based on a similar set of MultiDark simulations. In addition, our model provides a prediction of concentration according to mass and redshift with one single equation (Equation 3.5). The probability density function of concentration is a modified Schechter law, with mass dependency (Equation 3.6). We find that the concentration of low-mass halos has a faster redshift evolution than high-mass objects, especially in the high concentration regime. For concentration $c = 8$, the PDF for high-mass halos shifts by 0.32 dex from $z = 1.43$ to $z = 0$, while the low-mass one changes by 0.49 dex. We find the spin parameter λ to be modeled by a linear relation with mass and a probability density function well described by a modified Schechter function (Equation 3.8), in agreement with Rodriguez-Puebla et al. (2016). The offset parameter evolves with mass and redshift according to Equation 3.9. The negative slope of the relation suggests that low-mass halos are typically more relaxed compared to high-mass objects. This is true at every redshift. The offset distribution around the mean value is well described by a modified Schechter function (Equation 3.10). The peak of the distribution shifts by a factor of 1.5 between $z \sim 1.4$ and $z = 0$. This is in agreement with halos relaxing with time and the recent formation of cool cores in galaxy clusters.

We define a general mass function framework, where dark matter halos are not only described as a function of mass but also by the two additional variables $X_{\rm off}, \lambda$. This approach allows considering mass, offset parameter, and spin of each halo at the same time in a $\sigma - X_{\rm off} - \lambda$ function (Equation 3.14). We model it in Sect. 3.5 combining terms of a fiducial mass function (Bhattacharya et al., 2011) with modified Schechter functions for $X_{\rm off}, \lambda$, as obtained in Sect. 3.3. This new approach allows accounting for the dynamical state of dark matter halos directly in the context of the halo mass function, providing 2D and 1D distributions at the same time.

We describe the fitting procedure and results in Sect. 3.6. Our result at z=0 recovers the Comparat et al. (2017) mass function, which is fitted on the same set of simulations, with 3.3% accuracy. This means that our model is able to account for the dynamical state of dark matter halos simultaneously with mass and to describe the multiplicity function with great precision. In addition, our model includes the redshift evolution, according to Equation 3.20. The result is shown in Fig. 3.11.

Kapitel 4

Offset between X-ray and optical centers in clusters of galaxies: connecting eROSITA data and simulations

Galaxy clusters can be identified at different wavelengths, given their distinctive observational features. These include an over-density of red sequence galaxies (Gladders and Yee, 2005; Yang et al., 2007); a distortion of the images of background galaxies by strong and weak gravitational lensing (Maturi et al., 2005; Miyazaki et al., 2018b); X-ray emission due to thermal bremsstrahlung from the hot intra-cluster gas (Ebeling et al., 1998; Böhringer et al., 2000; Vikhlinin et al., 2009b; Pierre et al., 2016); and the distortion of the cosmic microwave background (CMB) spectrum due to the Sunyaev-Zel'dovich effect (Sunyaev and Zeldovich, 1972; Planck Collaboration et al., 2014b).

The definition and identification of the cluster center is a key aspect of their analysis. From a purely dark matter standpoint, it is natural to consider the deepest point in the potential well of the dark matter halo hosting the cluster. This is traced best by lensing observations (Zitrin et al., 2012). Other possibilities involve the peak or the centroid of the gas emission in the X-ray and millimeter bands (Rossetti et al., 2016; Gupta et al., 2017). Finally, a cluster center can be identified using optical and infrared data (Ota et al., 2020), by considering the position of the central galaxy (CG), for example, the brightest cluster galaxy (BCG).

Agreement between these definitions is expected if the dark matter halo and different baryonic components are completely relaxed and in equilibrium within the potential well of the cluster. However, galaxy clusters are rarely in complete dynamical equilibrium. They assembled at late times, undergoing mergers. This leads to disturbed mass distribution and an offset between different definitions of the cluster center.

A deeper insight into this topic is now possible thanks to eROSITA. The goal of this chapter is to exploit the optical follow-up of eROSITA clusters to gain a deeper knowledge of the offset between X-ray and optical centers, and compare the result to predictions from

hydrodynamical simulations and N-body models. We summarize the key aspects of various definitions of the cluster center, and their implications in terms of the dynamical state.

Historically, it is often assumed that the BCG is the central galaxy of the cluster, i.e. the one closest to the deepest point of the halo potential well (van den Bosch et al., 2004; Weinmann et al., 2006). Therefore, the BCG is used to define the cluster center in the optical band. However, recent works show that this is not always the case (Einasto et al., 2011; Lange et al., 2018). In particular, Skibba et al. (2011) find that the BCG is not the central galaxy in ∼25% of galaxy group-like halos. This fraction increases to ∼45% for clusters of galaxies. The development of new techniques to identify clusters in the optical band reduced the mis-centered fraction. The red-sequence Matched-filter Probabilistic Percolation (redMaPPer, Rykoff et al., 2014, 2016) is a cluster finding algorithm for photometric surveys, such as the Sloan Digital Sky Survey (SDSS, York et al., 2000), the Dark Energy Survey (DES, The Dark Energy Survey Collaboration, 2005), and the Large Synoptic Survey Telescope (LSST) at the Rubin Observatory (LSST Science Collaboration et al., 2009). It locates the cluster optical center using additional information such as redshift and local galaxy density. Hoshino et al. (2015) analyzed the occupation of luminous red galaxies with redMaPPer centering probabilities and show that the BCG is not the central galaxy in 20-30% of the clusters. The centering algorithm of redMaPPer is based on assigning a probability to each member of being the central galaxy and provides a more consistent definition of the optical center.

Rozo and Rykoff (2014) studied the performance of redMaPPer on SDSS data by comparing the optical catalog to overlapping X-ray and SZ data. They find that about 80% of the clusters are well centered, with offsets smaller than 50 kpc. The remaining 20% consist of mergers, which exhibit much larger offsets even up to 300 kpc. The displacement decreases at low redshift (Gozaliasl et al., 2019). This is in agreement with the hierarchical scenario, where structures relax at late times. The offset between peaks in various bands has been exploited to identify relaxed and disturbed systems (Mann and Ebeling, 2012; Rossetti et al., 2016, 2017; Oguri et al., 2018; Ota et al., 2020, 2022).

A detailed description of these offsets and their link to the cluster dynamical state is also important to assess possible biases and selection effects, especially in the current era of precision cosmology. For instance, a partial knowledge of the baryon physics affecting the evolution of galaxy clusters biases scaling relations between observables and clusters masses (Bahar et al., 2022; Chiu et al., 2022), which ultimately impact cosmological results (Chisari et al., 2019b; Genel et al., 2019; Salvati et al., 2020; Debackere et al., 2021; Castro et al., 2021). The disturbance and morphological diversity of these extended objects make the understanding of selection effects non-trivial (Weißmann et al., 2013; Cao et al., 2020). In addition, baryonic properties potentially affect the selection of clusters in astronomical surveys. They might alter the values of a specific observable, which ends up affecting the number of objects in the sample compared to an unbiased theoretical prediction.

X-ray observations of galaxy clusters suffer from the cool core bias (Eckert et al., 2011; Käfer et al., 2019). The largest dark matter halos hosting massive clusters of galaxies assemble at late times. Some clusters will not have had enough time to dynamically relax and develop a cool core. The resulting peak in the X-ray surface brightness profile can

bias the detection towards relaxed structures with a cool core, affecting the completeness of X-ray-selected samples of galaxy clusters. At fixed mass and redshift, cool core clusters are therefore more probable to be detected compared to non cool core ones. The cool core bias is expected to play a role in the characterization of clusters as extended sources. Cool core clusters possibly have a higher probability of being confused for point sources, because the peaked emissivity in the central region dominates over the extended emission in the cluster outskirts (Somboonpanyakul et al., 2021; Bulbul et al., 2022). This has an impact on cosmological studies using the halo mass function (Seppi et al., 2021). Therefore, it is necessary to take this selection effect into account. The evidence of the cool core bias in X-ray-selected samples has been highlighted by different works, especially when comparing X-ray to SZ selected samples, which are not affected by such bias due to the lower sensitivity to the central gas density (Hudson et al., 2010; Eckert et al., 2011; Rossetti et al., 2017; Andrade-Santos et al., 2017; Lovisari et al., 2017; Giulia Campitiello et al., 2022). However, other studies do not find a significant preference for relaxed clusters (e.g. Mantz et al., 2015a; Nurgaliev et al., 2017; McDonald et al., 2017; De Luca et al., 2020). This topic has been analyzed with eROSITA data by Ghirardini et al. (2021a), who did not find a clear bias towards relaxed structures. In addition, Bulbul et al. (2022) find a preference for cool cores only when looking for clusters cataloged as point sources. Strong evidence for the cool core bias in the point-like sample is is also predicted by eROSITA simulations (Seppi et al., 2022) (see Fig. 2.9).
The fraction of mass in substructures, central entropy, spin, and offset parameters give additional insight into the dynamical state (Meneghetti et al., 2014; Biffi et al., 2016; Henson et al., 2016; De Luca et al., 2020; Seppi et al., 2021).
A precise knowledge of the cluster center would benefit various studies, such as the measure of weak lensing profiles (Chiu et al., 2022), where the error in the measurement may be reduced with a better comprehension of the miscentering (George et al., 2012; Zhang et al., 2019; Yan et al., 2020; Ota et al., 2022); or detailed comparison of cluster density profiles with simulations (Zhuravleva et al., 2013; Diemer, 2022).

Here, we measure the offset between the position of the X-ray and the optical centers for eROSITA clusters. We use two samples: eFEDS (Liu et al., 2022), and eRASS1. The optical follow-up is performed with a modified version of redMaPPer tailored to eROSITA, making use of the prior knowledge of the X-ray position (Ider Chitham et al., 2020). We study the distribution of the offsets and different physical effects affecting them. We look for a link between observations and the dynamical state of dark matter halos in N-body simulations (Klypin et al., 2016; Seppi et al., 2021). We consider the offset parameter ($X_{\rm off}$), that is the displacement between the peak of the mass profile and the center of mass of dark matter halos. In Kap. 3, we calibrated a mass function model that allows marginalization over variables related to the dynamical state. Instead, we marginalize on mass, predict the distribution of $X_{\rm off}$, and compare it to the displacement between X-ray and optical centers. A correlation between $\rm X_{off}$ and the observable offset is expected, as the central galaxy may be trapped in the deepest point of the halo potential, while the distribution of the X-ray emitting gas depends on the gravitational potential also on larger scales. We exploit hydrodynamical simulations to develop this connection. We use the Magneticum

(Biffi et al., 2013; Hirschmann et al., 2014; Biffi et al., 2018; Ragagnin et al., 2017) and the Illustris-TNG (Pillepich et al., 2018c; Nelson et al., 2019) simulations.

In the rest of this chapter, we refer to the offset between the X-ray and the optical center in eROSITA clusters as *observed offset*, to the one measured in hydrodynamical simulations as *predicted offset*, and to X_{off} as the *offset parameter*. This chapter is structured as follows. We summarize the eROSITA data, its processing, the optical follow-up, and the hydrodynamical simulations in Sect. 4.1. We describe our method for computing the observed offsets in eROSITA data, in simulations, and using the N-body model from Seppi et al. (2021) in Sect. 4.2. We present the distributions of the offsets and our results in Sect. 4.3. We discuss our findings and how to use the offsets in a cosmological framework in Sect. 4.4. We finally summarize our results in Sect. 4.5.

4.1 Data

In this Section, we describe the X-ray observations, the optical follow-up, and the hydrodynamical simulations used in this chapter.

4.1.1 eROSITA

We use X-ray data from the eROSITA X-ray telescope. The observations are processed with the eROSITA Standard Analysis Software System (eSASS, Brunner et al., 2022). A detailed description of the detection process and a discussion on cluster detection with eROSITA is given in Kap. 2. In addition, we study the probability of membership for all galaxy members in each cluster using redMaPPer (Rykoff et al., 2014; Ider Chitham et al., 2020) in scanning mode, making use of the prior knowledge of the X-ray position.

eFEDS

We use the eFEDS cluster catalog from Liu et al. (2022). It includes 542 clusters with $\mathcal{L}_{\rm DET}>5$ and $\mathcal{L}_{\rm EXT}>6$. The clusters are confirmed in the optical band and the redshifts are measured with redMaPPer (Rykoff et al., 2014), combining optical data from different surveys such as the Hyper Suprime-Cam (HSC) Subaru Strategic Program (HSC-SSP, Oguri et al., 2018), the Dark Energy Camera Legacy Survey (DECaLS, Dey et al., 2019)), the Sloan Digital Sky Survey (SDSS, Blanton et al., 2017), the 2MASS Redshift Survey (2MRS, Huchra et al., 2012), and the Galaxy And Mass Assembly Survey (GAMA, Driver et al., 2011). A detailed weak-lensing study on HSC observations by Chiu et al. (2022) provides halo masses for a subsample of 434 eFEDS clusters.

eRASS1

eROSITA performed its first scan of the whole sky during the first six months of the survey phase, from December 13th 2019 until June 11th 2020, completing the first all-sky survey

Tabelle 4.1: Numerical and physical parameters describing the Magneticum and Illustris-TNG simulations.

	Magneticum-Box2/hr	**TNG-300-1**
Box size [Mpc/h]	352	205
$\Omega_{\rm M}$	0.272	0.3089
$\Omega_{\rm B}$	0.0456	0.0486
Ω_{Λ}	0.728	0.6911
σ_8	0.809	0.8159
$\rm H_0$	70.4	67.74
$\rm n_s$	0.963	0.9667
N particles	$2\times\ 1584^3$	2500^3
$\rm M_{DM}$ [$\rm M_\odot/h$]	6.9×10^8	5.9×10^7

Notes. Volume: total comoving volume covered by the simulation, $\Omega_{\rm M}$: total matter density parameter, $\Omega_{\rm B}$: baryonic matter density parameter, Ω_{Λ}: dark energy density parameter, σ_8: normalization of the linear matter power spectrum, H_0: Hubble constant, $\rm n_s$: initial slope of the linear matter power spectrum, N particles: total number of dark matter particles in the simulation, $\rm M_{DM}$: mass of the dark matter particles.

(eRASS1[1]). Given the scanning strategy of the telescope, the exposure time depends on the angular position on the sky. Shallow regions around the ecliptic equator are covered for less than 100 seconds, while deep areas around the ecliptic poles are observed for more than 1.2 ks (see Predehl et al., 2021, for more details). The average exposure time of the eRASS1 survey is ∼250 s. We use the German half of the sky (eROSITA_DE). The majority of the area overlaps with different optical surveys, such as the Dark Energy Camera Legacy Survey (DECaLS, Dey et al., 2019), the Dark Energy Survey (DES, Sevilla-Noarbe et al., 2021), and the Kilo-Degree Survey (KiDS, Kuijken et al., 2019). The optical identification, the measurement of redshifts and optical properties is carried out by redMaPPer (Rykoff et al., 2014; Ider Chitham et al., 2020).
We use the X-ray position measured by eSASS, the optical centers (see Sect. 4.2), and the redshift provided by redMaPPer.

4.1.2 Simulations

In this work, we use the Illustris-TNG and the Magneticum simulations. The main numerical and cosmological parameters for the two simulations are shown in Table 4.1.

Magneticum

The Magneticum simulation suite[2] is a set of cosmological hydrodynamical and dark-matter-only (DMO) simulations (Biffi et al., 2013; Hirschmann et al., 2014; Dolag, 2015;

[1] https://www.mpe.mpg.de/7461950/erass1-presskit
[2] http://www.magneticum.org

Steinborn et al., 2015; Ragagnin et al., 2017; Dolag et al., 2017; Singh et al., 2020), spanning different ranges of resolution and box size. These simulations are run with the TreePM-SPH code `P-GADGET3` (Springel, 2005). Multiple processes regulated by baryonic physics are taken into account in the simulation, such as radiative cooling (Wiersma et al., 2009), heating due to star formation, supernovae, galactic winds (Springel and Hernquist, 2003), chemical enrichment (Tornatore et al., 2007), and AGN feedback processes (Fabjan et al., 2010). The Magneticum simulations are successful at reproducing the black hole mass density (Di Matteo et al., 2008), the AGN luminosity function (Hirschmann et al., 2014; Steinborn et al., 2016; Biffi et al., 2018), morphological properties of galaxies (Teklu et al., 2015; Remus et al., 2017), and the pressure profiles of galaxy clusters (Gupta et al., 2017). This set of simulations has been used to quantify the impact of baryons on the halo mass function (Bocquet et al., 2016; Castro et al., 2021), and for dedicated studies of the Large Scale Structure around merging galaxy clusters with eROSITA (Biffi et al., 2022).
We focus on the **Box2/hr** simulation. This is computed assuming a WMAP cosmology (Komatsu et al., 2011). Given our interest in clusters of galaxies, it provides a great compromise between the size of the box and the resolution of the dark matter halos. The side of the simulated cube is 352 Mpc/h (500 Mpc). The box contains 475 halos more massive than M_{500c}=1×10^{14} $M_\odot$ at z=0[3]. The resolution of the dark matter particles is 6.9×10^8 $M_\odot/h$, which allows detailed properties of the most massive halos hosting clusters and groups to be measured. A summary of the key parameters for the simulation is reported in Table 4.1.

Illustris-TNG

The Illustris-TNG project[4] is a collection of 18 complementary hydrodynamical simulations coupled with dark-matter-only runs (Weinberger et al., 2017; Pillepich et al., 2018c; Barnes et al., 2018; Nelson et al., 2019). It spans different box sizes, resolutions, and treatment of baryons. The simulations are run with the quasi-Lagrangian code `AREPO` (Weinberger et al., 2020). It includes gas radiative mechanisms, star formation, stellar evolution, supernovae explosions, the formation and accretion of supermassive black holes, and the amplification of magnetic fields. The TNG project successfully reproduces the galaxy color distribution as a function of stellar mass (Nelson et al., 2018), the stellar mass function at recent epochs, the distribution of stellar mass inside galaxy clusters (Pillepich et al., 2018a), the scaling relation between radio power and X-ray emission in galaxy clusters (Marinacci et al., 2018), the low redshift quasar luminosity function (Weinberger et al., 2018), the chemical evolution of gas in galaxies (Naiman et al., 2018), and the galaxy two-point correlation function (Springel et al., 2018).
The TNG project assumes a Planck cosmology (Planck Collaboration et al., 2016b). We use the TNG-300-1 simulation. It is the largest available box, with a side of 205 Mpc/h (300 Mpc). It is smaller than Magneticum Box2/hr, and contains therefore fewer halos:

[3] M_{500c} is the total mass of the cluster encompassed by a radius containing an average density that is 500 times larger than the critical density of the Universe.

[4] `https://www.tng-project.org`

159 objects more massive than M_{500c}=1×10^{14} $M_\odot$ at z=0. However, it has a higher particle resolution (see Table 4.1).

4.2 Method

In this Section, we describe how we processed and analyzed the data, and how we compared it to theoretical models and hydrodynamical simulations.

4.2.1 Offset for eROSITA clusters

Our sample consists of eFEDS and eRASS1 clusters with EXT_LIKE > 6. We additionally require a measure of the uncertainty on the X-ray position by eSASS (RADEC_ERR > 0). For eRASS1, we exclude clusters that are not covered by optical surveys and are therefore lacking a measure of the optical center. We determine the X-ray center using the cluster position defined by eSASS. It is the best fit position of the β-model fit convolved with the PSF, that determines the X-ray centroid.
We consider two definitions of the optical center. The first one is given by the centering algorithm of redMaPPer, that uses a Bayesian classification algorithm to locate the cluster optical center. It is based on a model of the color of red sequence galaxies, and a local red galaxy density filter that provides consistency between the photometric redshift of the central galaxy and the cluster. It also matches the central galaxy luminosity to an expected value given the cluster richness, which is closely related to the total number of galaxy members hosted by the cluster (Rykoff et al., 2014). The optical center is not always coincident with the brightest cluster galaxy. In fact, Rykoff et al. (2016) find that the optical center from redMaPPer is not the brightest member for about 20% of the clusters. This approach provides the probability for each member to be the central cluster galaxy P_{cen} (see Eq. 56 in Rykoff et al., 2014).
Secondly, we explore an alternative definition of the optical center and consider the position of the galaxy member with the largest membership probability p_{mem}. It is the probability that a galaxy near a cluster is a cluster member and should not be confused with the probability of being the central galaxy P_{cen}. It is computed for each galaxy by combining different filters. The most important one is a model of the color evolution of red-sequence galaxies as a function of redshift (see Eq. 1 in Rykoff et al., 2014).
Given the angular positions of the X-ray and optical centers, we compute the angular separation between them and convert it to the comoving physical kiloparsec scale based on the cluster redshift, according to

$$\Delta_{X-O} = \frac{c}{H_0} \int_0^z \frac{dz}{\sqrt{\Omega_M (1+z)^3 + \Omega_\Lambda}} \times \theta, \qquad \text{[kpc]} \tag{4.1}$$

where z is the cluster redshift, c is the speed of light, H_0 is the Hubble constant, and θ is the angular separation between X-ray and optical centers in radians. Similarly, we measure the observed offset between the X-ray center and the position of the galaxy with the largest

membership probability, following again Eq. 4.1, but considering the galaxy with largest p_{mem} instead of the optical center from redMaPPer. We name this quantity $\Delta_{X-P_{mem}}$.
We estimate the error on the X-ray center by multiplying the uncertainty on the angular X-ray position by the physical scale per unit angle. We estimate a systematic error on the optical center accounting for the separation between the optical center identified by redMaPPer and the position of the five members with the largest P_{cen} weighted by their centering probability, according to Equation 4.2:

$$\delta_O = \sqrt{\sum_{i=1}^{N=5} (P_{cen,i} \times \Delta_{O-O_i})^2}, \tag{4.2}$$

where the index i runs on the five most probable members and Δ_{O-O_i} is the separation in kpc scale between the optical center and the position of the i galaxy member. In this manner, the uncertainty δ_O accounts for the fact that the definition of a center is complicated when there are many bright galaxies with similar probability of being the central galaxy.
We finally compute the cumulative distribution function (CDF) of Δ_{X-O} and $\Delta_{X-P_{mem}}$. We do this first for the whole eRASS1 and eFEDS samples, by restricting to secure clusters with more than 20 counts and richness λ>20. We obtain 182 (4564) clusters from eFEDS (eRASS1) satisfying these conditions. We then focus on a more specific sub-sample of 87 eFEDS clusters between redshift 0.15 and 0.4 and M_{500c} between 1×10^{14} and 8×10^{14} $M_\odot$. The mean values of mass and reshift are $M_{500c}=2.16\times10^{14}$ $M_\odot$ and z=0.30. We use this well-defined sample to do a comparison with simulations.

4.2.2 Analytical DMO model

We propose a link between the observed X-ray to optical offset in observations and the theoretical model developed in Kap. 3. There we calibrated a model for the halo mass function, that additionally includes variables describing the dynamical state of dark matter halos. We are particularly interested in the offset parameter X_{off}, that is the displacement between the center of mass of a dark matter halo and the peak of its density profile, normalized to the virial radius. Such mass function model predicts the dark matter halo abundance as a function of mass, offset parameter, and spin, offering the possibility of integrating out one or more of these variables. We marginalize on mass and spin and obtain the analytical prediction of the 1D distribution for the offset parameter (see Eqs. 3.14, 3.15, and 3.16).
We compare the observed offsets between the X-ray and optical centers to the offset parameter in physical scales $X_{off,P}$ (in kpc, i.e. not normalized to the virial radius) from N-body simulations (see Eq. 3.25). We compute the halo multiplicity function dependent on the offset parameter by marginalizing on mass and detection probability according to Equation 4.3:

$$f(X_{off,P}) = \int_{M_{low}}^{M_{up}} g(\sigma(M), X_{off,P}) P(M) dM, \tag{4.3}$$

where M_{low}=1×10^{14}M$_\odot$ and M_{up}=8×10^{14}M$_\odot$. In addition, we marginalize over the detection probability as a function of mass $\hat{P}$(M). The mass trend is encoded in the variance of the density field σ. We calibrate the detection probability by dividing the eFEDS multiplicity function (see Eq. 4.4) computed from the cluster number density by the theoretical prediction and model it with an error function. We refer to the detection probability model as P(M), computed and modeled according to Eq. 4.4:

$$\begin{aligned} f(\sigma) &= \frac{d\mathrm{n}}{d\ln\mathrm{M}}\frac{M}{\rho_m}\left(\frac{d\ln\sigma^{-1}}{d\ln\mathrm{M}}\right)^{-1}, \\ \hat{P}(M) &= f_{\mathrm{eFEDS}}(\sigma)/f_{\mathrm{SEPPI+21}}(\sigma), \\ P(M) &= \frac{1}{2}erf[A(\log_{10} M_{500c} - M_0)] + \frac{1}{2}, \end{aligned} \tag{4.4}$$

where the parameters have been fit with the *curve_fit* software[5]. The values are M_0=14.314±0.001 and A=2.30±0.03. Finally, we account for projection effects by projecting the theoretical three dimensional $X_{\mathrm{off,P}}$ on the sky according to Eq. 4.5:

$$S_{X_{\mathrm{off,P}}} = X_{\mathrm{off,P}}\frac{1}{\pi}\int_0^\pi \sin\theta d\theta = \frac{2}{\pi}X_{\mathrm{off,P}}. \tag{4.5}$$

We use the corrected $S_{X_{\mathrm{off,P}}}$ to compute the theoretical cumulative distribution function of the projected offset parameter.

4.2.3 Prediction from hydrodynamical simulations

We process the Magneticum and TNG simulations in a similar way to each other. For each halo in the simulation, we relate the optical center to the position of the main subhalo identified by the SubFind algorithm, which contains the central galaxy. We relate the X-ray center to the gas center computed from an emission measure weighted center of mass. The position of each gas particle contained by the halo is weighted by its mass and local density. We consider particles within the virial radius of each halo. We restrict to X-ray emitting gas particles with temperature between 0.1 and 10 keV. We finally compute the gas center according to Equation 4.6:

$$\begin{aligned} w_{g,i} &= \rho_{g,i} \times m_{g,i}, \\ CM_g &= \frac{\sum_{\mathrm{i=1}}^{N} w_{g,i} \times x_{g,i}}{\sum_{\mathrm{i=1}}^{N} w_{g,i}}, \end{aligned} \tag{4.6}$$

where $w_{g,i}$ is the weight assigned to each gas particle, $\rho_{g,i}$ is the local gas density, $m_{g,i}$ is the gas particle mass, and $x_{g,i}$ is the particle position. The index i runs on the N gas particles contained by a halo. We compute the predicted offset between the CG and the gas center

[5] https://scipy.org

as their relative distance on the x–y projected cartesian plane.
In addition, we run the ROCKSTAR halo finder (Behroozi et al., 2013) on both Magneticum and TNG. We process the hydro simulation as well as the respective parent dark-matter-only (DMO) runs. ROCKSTAR provides a measure of X_{off} for each identified halo. We focus on distinct main halos. We perform a positional matching between our ROCKSTAR halo catalogs from the hydro and DMO runs to the SUBFIND catalogs provided together with the particle data from the Magneticum and TNG projects. The agreement between these catalogs is excellent. We discard halos where the location of the center disagrees by more than 500 kpc and the measure of M_{500c} differs more than 10% between the three catalogs. Because the SUBFIND catalogs provide the total mass, including stars and gas, we correct it by the gas fraction from Pratt et al. (2009), see Fig. 8 therein, when comparing it to ROCKSTAR masses. The gas fraction from Pratt et al. (2009) provides a robust measure compared to a variety of other samples (see Eckert et al., 2021, for a review). In total, we lose about 2% (1%) of the halos with $M_{500c}>1\times10^{13}$ (1×10^{14}) $M_\odot$. The matched halos allow us to compare the displacement between the gas center and the central galaxy to the offset parameter for common objects between the three catalogs. To match the average redshift of the high mass–low redshift eFEDS sample, we study the snapshot at z=0.30 for TNG-300. For Magneticum, we use the closest snapshot available at z=0.25, where the particle data has been stored for the full hydro and the DMO run. We verify that this does not bias our results in Sect. 4.3.1. The TNG-300 snapshot at z=0.30 contains 2232 (107) halos with $M_{500c}>1\times10^{13}$ (1×10^{14}) $M_\odot$. The Magneticum Box2/hr snapshot at z=0.25 contains 9293 (314) halos with $M_{500c}>1\times10^{13}$ (1×10^{14}) $M_\odot$.

4.3 Results

In this Section, we present our main findings: the distribution of the displacement between the X-ray and optical center in eROSITA, its comparison to Magneticum and TNG, and to the N-body model from Seppi et al. (2021).

4.3.1 Offset distributions and comparison to simulations for eFEDS

We focus first on the sample over which we have the most control: a subsample of 87 eFEDS clusters with $0.15 < z < 0.4$ and $1\times10^{14} < M_{500c} < 8\times10^{14}$ $M_\odot$. The redshift and the X-ray positional uncertainty have been measured for all these clusters. Given the mean values of $M_{500c}=2.16\times10^{14}$ $M_\odot$ and z=0.30 for this sample, the eFEDS selection function yields an average completeness of about 80% (Liu et al., 2022). The observed offsets Δ_{X-O} and $\Delta_{X-P_{mem}}$ are shown by the blue and orange shaded areas in Fig. 4.1. The green line denotes the CDF of the projected offset parameter $S_{X_{off,P}}$, the red (violet) line shows the CDF of the predicted offset in the Magneticum (TNG-300) simulation at z=0.25 (0.30). The corresponding dashed and dotted lines account for maximum and minimum projection effects. For each cluster in the simulations, we consider the largest possible displacement in the case where the two centers lay on a plane that is perpendicular to the line of sight, and

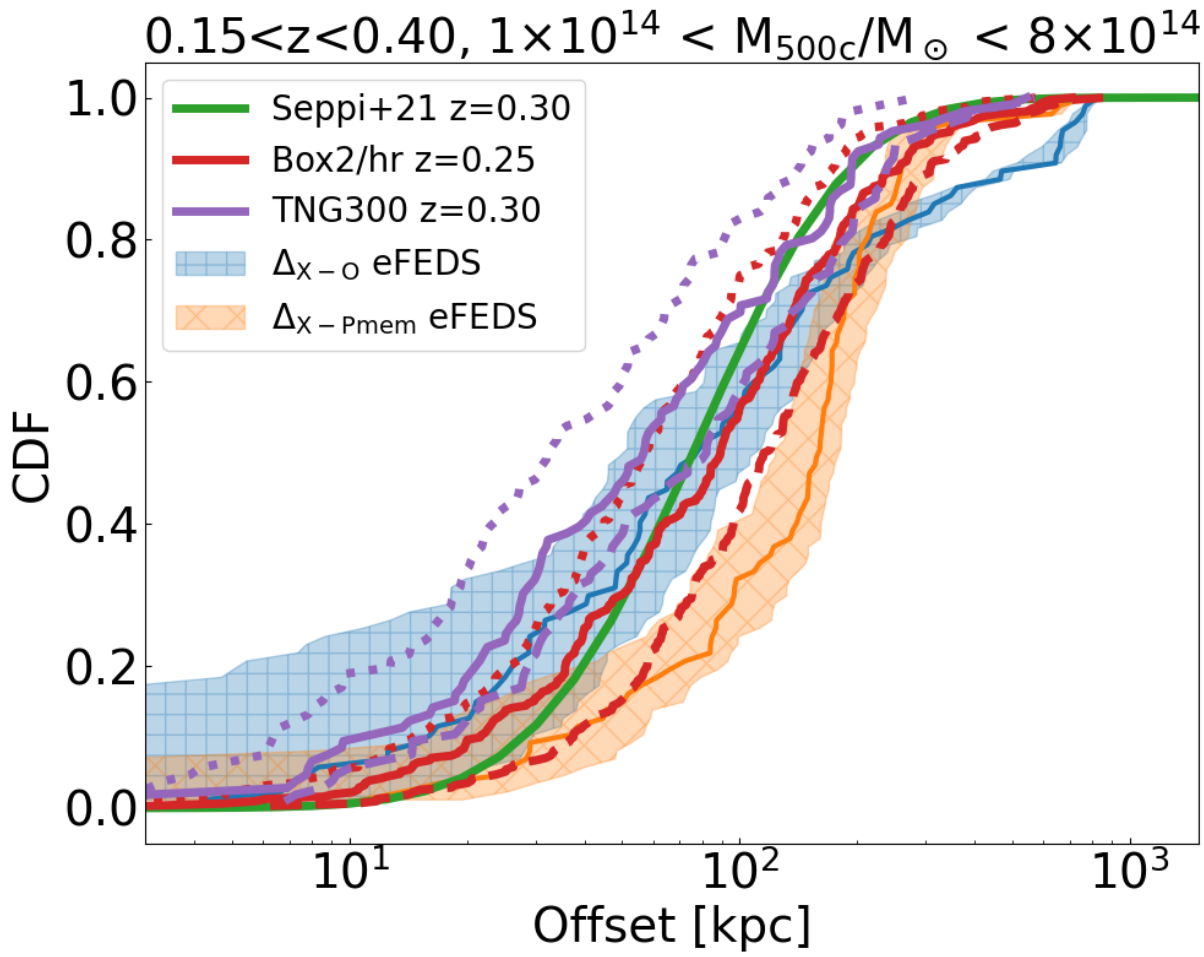

Abbildung 4.1: Comparison between the offsets measured in eROSITA, the prediction of the theoretical model, and hydrodynamical simulations. The cumulative distribution functions of the observed offsets between X-ray and optical centers for eFEDS clusters between redshift 0.15 and 0.4, and mass between 1×10^{14} and 8×10^{14} $M_\odot$ are denoted by the blue and orange lines. The first one refers to the optical center identified by the redMaPPer centering algorithm, the latter to the position of the galaxy with the largest membership probability. The shaded areas identify the uncertainty on the distributions. The green line shows the prediction obtained from the Seppi et al. (2021) model described in Sect. 4.2.2. The red (violet) curve denotes the CDF of the predicted offsets between the gas center and the CG position in the Magneticum (TNG) simulation described in Sect. 4.2.3. The corresponding dashed and dotted lines account for the maximum and minimum projection effects. There is a broad agreement between the data, the prediction of the simulations, and the N-body model. However, the tails of the distributions are different. The N-body model predicts larger (smaller) displacements compared to data and hydrodynamical simulations at the low (high) offset end.

the minimum one by choosing the smallest displacement after projecting the same clusters on the x-y, y-z, and x-z planes.

Average offsets

For the eFEDS sub-sample, we study the average observed offset at the 50% percentile point of the CDF. We measure Δ_{X-O}=$76.3^{+30.1}_{-27.1}$ kpc and $\Delta_{X-P_{mem}}$=$157.4^{+20.6}_{-34.0}$ kpc. The flattening of Δ_{X-O} at large offsets is given by a tail of recent mergers and disturbed clusters. The average predicted offset in hydrodynamical simulations is equal to 57.2 kpc for TNG-300 and 87.1 kpc for Magneticum Box2/hr. We see that both simulations predict offsets that are on average compatible with the distribution of Δ_{X-O}, but they are in disagreement with the observed offset between the X-ray center and the position of the galaxy with the largest membership probability $\Delta_{X-P_{mem}}$. Therefore, the displacement between the hot gas and the CG in hydrodynamical simulations is a good prediction of the offset between the optical center from redMaPPer and the X-ray position from eSASS in eROSITA data. To assess whether the different redshift considered for the Magneticum simulation impacts our findings, we do the same analysis for the snapshot at z=0.25 of the TNG-300 simulation, where the particle data are available also for the parent DMO run. We find that the predicted offset is on average smaller by about 4 kpc compared to the snapshot at z=0.30. This is much smaller than the typical uncertainties on the data. We conclude that studying the snapshot at z=0.25 in the Magneticum simulation does not bias our results.
The average value of the projected offset parameter is $S_{X_{off,P}}$=75.8 kpc. Similarly to the offsets predicted by TNG and Magneticum, it is in agreement with Δ_{X-O}, but disagrees with $\Delta_{X-P_{mem}}$. On average, we conclude that there is good agreement between the offsets observed in eROSITA clusters, the ones predicted by hydrodynamical simulations, and by the N-body model from Kap. 3.

Tails of the distributions

The tails of the Δ_{X-O} and $S_{X_{off,P}}$ distributions, where the CDF is smaller than about 0.2 and larger than 0.7, are different. We attribute this discrepancy to baryonic effects that are not present in the DMO simulations, such as dragging, cooling, and the disruption of the gas by AGN feedback and recent mergers. These effects tilt the shape of this distribution. This is in agreement with previous works, where the shape of the offset distribution changes from a modified Schechter function in N-body simulations (Rodriguez-Puebla et al., 2016; Seppi et al., 2021) to a lognormal distribution in data (Mann and Ebeling, 2012). We further discuss this result in Sect. 4.4.

Fraction of relaxed clusters

To separate relaxed and disturbed clusters, we follow the example of Ota et al. (2022) and apply an observed offset cut according to

$$\Delta_{X-O} < 0.05 \times R_{500c}. \tag{4.7}$$

We find that 27 clusters out of 87 are classified as relaxed. Our relaxed fraction of 31% is in agreement with the upper limit of <39% found by Ota et al. (2022). We additionally estimate an upper limit of the relaxed fraction by accounting for the uncertainty on the measure of the observed offset (as explained in Sect. 4.2.1), assuming the lower limit of Δ_{X-O} within the error. We apply again the same cut in Eq. 4.7 and obtain a relaxed fraction of $< 59\%$. Our results show that there is not a strong preference for relaxed objects compared to unrelaxed ones in this eFEDS cluster sample. This is in agreement with previous work on eFEDS data. Ghirardini et al. (2021a) combined eight different morphological parameters (central density, concentration, centroid shift, ellipticity, cuspiness, power ratios, photon asymmetry, and Gini coefficient) into the single relaxation score parameter. They did not find a clear preference for cool core clusters over disturbed ones and showed that the transition from a relaxed to a disturbed cluster population is smooth. In addition, Bulbul et al. (2022) analyzed the clusters hidden in the point-like sample of eFEDS sources, identifying them using optical data. They found that only the clusters in the point-like sample show a peaked profile in the central region. Finally, predictions from eROSITA simulations show that there is a significant preference for the detection of relaxed systems just in the point-like sample (see Fig. 2.9).

4.3.2 Full eROSITA samples

We now expand upon the eFEDS subsample by measuring the position of the X-ray and optical centers for the full sample of 182 eFEDS and 4564 eRASS1 clusters as explained in Sect. 4.2.

We show the cumulative distribution function of the observed offsets in Figure 4.2. The blue (green) line shows the offset between the X-ray center and the redMaPPer center for the eFEDS (eRASS1) sample. The shaded areas denote the 1σ error on the observed offset. We measure an average value of $\Delta_{X-O}=92.6^{+44.3}_{-35.1}$ kpc in eFEDS and $\Delta_{X-O}=158.5^{+53.0}_{-57.5}$ kpc in eRASS1. On average, the eRASS1 sample shows larger offsets compared to eFEDS. Since eRASS1 is a shallow survey compared to the deeper and more uniform eFEDS, and the detection probability for a given cluster grows as a function of exposure time (Clerc et al., 2018; Seppi et al., 2022), it contains a larger fraction of high-mass, high-offset objects compared to eFEDS.

The trends of the displacement between the X-ray center and the position of the galaxy with the largest membership probability are more similar. They are shown by the orange (red) line in Fig. 4.2 for eFEDS (eRASS1). The shaded areas denote the 1σ error on the observed offset. We measure an average value $\Delta_{X-P_{mem}}=160.9^{+40.5}_{-45.6}$ kpc in eFEDS and $\Delta_{X-P_{mem}}=162.7^{+45.8}_{-46.3}$ kpc in eRASS1. The full samples mix clusters with different mass and redshift, which dilutes the intrinsic differences of $\Delta_{X-P_{mem}}$ between the eFEDS and the eRASS1 samples.

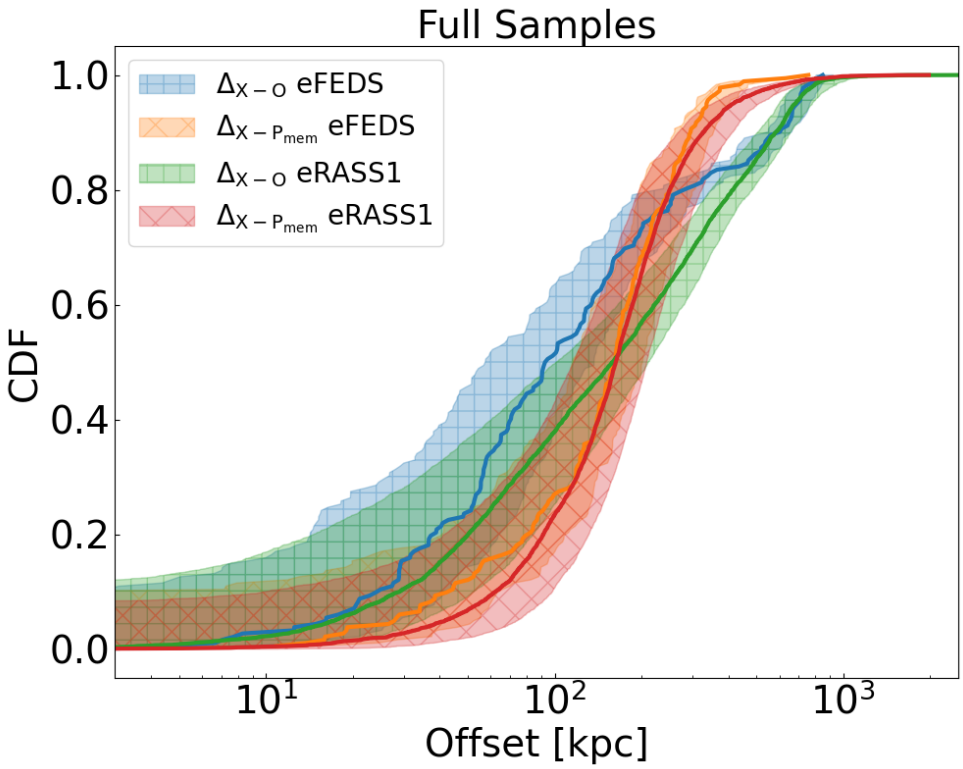

Abbildung 4.2: Cumulative distribution functions of the observed offsets between X-ray and optical centers for eROSITA clusters between redshift 0.15 and 0.8, more than 20 counts, and richness λ>20. These cuts yield 182 (4564) clusters from eFEDS (eRASS1). The shaded areas denote the uncertainty on the distributions. Different colors denote distinct definitions of the optical center: the one identified by the redMaPPer centering algorithm and the position of the galaxy with the largest membership probability (blue and orange for eFEDS, green and red for eRASS1).

4.4 Discussion

In this section, we discuss the different offsets presented in Sect. 4.3. The physical effects affecting the offsets play a key role in understanding the cause behind the smaller (larger) displacements measured in eROSITA data and hydrodynamical simulations compared to the N-body model in the low (high) offset regime. The origin of the observed offsets in clusters of galaxies is related to the different response of each cluster component to different astrophysical phenomena.

The effect of mergers and AGN feedback on the offset distribution is discussed in Sect. 4.4.1 and 4.4.2. Their combination and the transition from the DMO scenario to the observed offset distribution are presented in Sect. 4.4.3. The discrepancies between the offsets predicted by different hydrodynamical simulations are discussed in Sect. 4.4.4. Finally, the potential use of the offsets in a cosmological experiment is presented in Sect. 4.4.5.

4.4.1 Mergers

Very large observed offsets likely originate from mergers between smaller objects into massive clusters. Cluster mergers are one of the most energetic processes in the Universe, as the total kinetic energy involved reaches values up to 10^{65} erg (Markevitch et al., 1999; Sarazin, 2002; Markevitch and Vikhlinin, 2007). In this context, it is particularly interesting to explore the differences between the central galaxy and the gas in relation to the dark matter distribution.

The dark matter is mostly sensitive to gravitational interaction, while the central galaxy and the gas are additionally subject to a variety of effects such as electromagnetic forces, ram pressure, scattering, and cooling (Merten et al., 2011). When two clusters merge, the dark matter components stream through each other according to the evolving gravitational field, without being slowed down by the dragging experienced by baryonic components because of the additional interactions. The result is that after a merger, when the newly formed halo relaxes and the gas cools down, the displacement between the dark matter profiles of the merging clusters is larger than the separation between the gas distributions. In fact, clusters undergoing mergers typically show large offsets up to hundreds of kpc between different components (Menanteau et al., 2012; Mann and Ebeling, 2012; Dawson et al., 2012; Monteiro-Oliveira et al., 2017). Large observed offsets provide therefore a hint of merger activity, compared to small offsets that characterize the pre-merger phase (Jauzac et al., 2015; Ogrean et al., 2015). An extreme case is the famous 1E 0657–56, also known as the bullet cluster (Markevitch et al., 2002; Clowe et al., 2006), where the total mass distribution traced by weak lensing extends to larger radii compared to the emission of the hot gas imaged with the Chandra X-ray observatory. This is in agreement with our result in Fig. 4.1, where we find a larger amount of clusters showing an observed offset of tens of kpc compared to the DMO prediction.

In addition, since the gas trails the dark matter during a merger because of ram pressure and friction, the gas starts sloshing within the cluster potential. This causes large observed offsets when the gas approaches the point of null velocity and positive acceleration during the sloshing process (Ascasibar and Markevitch, 2006; Markevitch and Vikhlinin, 2007; Sanders et al., 2020; Pasini et al., 2021). The complex behavior of the gas during the merging process is not easily mappable to the dark matter-only scenario. In these cases, the large offsets seen in data are not compatible with simple theoretical models, as described by De Propris et al. (2021). The authors find that the BCG is generally aligned with the cluster mass distribution, showing that even if being displaced by a merger or if the dark matter halo is not relaxed, the central galaxy does follow the cluster potential. Hikage et al. (2018) tested the performance of the redMaPPer centering algorithm using galaxy-galaxy lensing and confirmed that the central galaxy is not always the brightest member. A similar result was presented by Hoshino et al. (2015), who studied the distribution of luminous red galaxies in clusters. Therefore, the BCG is possibly a biased tracer of the deepest point of the halo potential well, especially for unrelaxed systems where the definition of the BCG is not trivial, and the brightest cluster galaxy may belong to a satellite merging halo. The identification of the central galaxy using centering probabilities with redMaPPer

provides a better tracer of the center of the dark matter halo. This is consistent with the agreement between the eROSITA data and hydrosimulations only when using Δ_{X-O}. In fact, the median of the $\Delta_{X-P_{mem}}$ distribution does not agree with the DMO model, with hydrodynamical simulations, nor with the median of Δ_{X-O}. The additional information from the whole galaxy population encoded in Δ_{X-O} provides an optical center that is on average closer to the X-ray center. However, in complex mergers galaxies in the cluster outskirts may shift the optical center away from the X-ray one compared to the galaxy with largest membership probability. This explains the extension to large values for Δ_{X-O} in the most disturbed clusters.
Compared to the definition of the optical center, the X-ray emitting gas may not properly trace the center of the cluster potential after being disrupted by complex mergers. This was studied by Cui et al. (2016). The authors analyzed the location of different centers of galaxy clusters in simulations and found that the BCG shows a better correlation with the center of the gravitational potential compared to the X-ray gas. They measure an average separation between the BCG and the potential center smaller than 10 kpc, while the displacement between X-ray and potential centers reaches average values of tens of kpc. This is also in agreement with Fig. 4.1, where the data shows larger offsets compared to the DMO prediction in the high offset regime. We conclude that on one hand, the central galaxy is on average more likely to be trapped in the vicinity of the deepest point of the cluster potential. On the other hand, the X-ray center, being related to gas permeating the whole halo, is more sensitive to the overall variations of the potential during merger activity and is altered by AGN feedback (see Sect. 4.4.2). This is in concordance with previous work on observations (George et al., 2012) and simulations (Cui et al., 2016).

4.4.2 AGN feedback

AGN feedback plays an additional role in this context. Efficient accretion onto the SMBH of the central galaxy is known to impact the gas on very large scales inside the dark matter halo hosting a galaxy cluster (see Eckert et al., 2021, for a review). The central AGN does not only reorganize the gas on large scales but the presence of jets digging cavities in the gas distribution produces a significant diversity of the gas morphology. The structure of the gas can be disrupted by AGN feedback, which pushes the gas away from the cluster center (Gaspari et al., 2012; Gitti et al., 2012; McNamara and Nulsen, 2012; Li et al., 2015). This contributes to the larger predicted offsets measured in the presence of baryons compared to the N-body simulations. However, the AGN impact on the offsets is not immediate. In fact, the central galaxy becomes active when there is enough gas supply to the central region of the cluster, which means that a cluster is more likely to be relaxed shortly before the beginning of AGN activity (Fabian, 2012; Pinto et al., 2018). It is also reasonable to expect a correlation between the AGN impact on the gas distribution and redshift. Dark matter halos are smaller at early times and the feedback may distribute and reorganize the gas on large scales more easily.
The notion that the gas is displaced compared to the dark matter distribution has been explored by Cui et al. (2016). They compare simulations with and without AGN feedback

and find that its activation enhances the offset between the gas and the dark matter centers, especially for clusters with predicted offsets between around 10 and 30 kpc in the simulation without AGN. The predicted offset reaches an average value of about 70 kpc in the run with active AGN. This is in agreement with our measurement of the observed offset Δ_{X-O} is Sect. 4.3. The authors also demonstrate that the X-ray centroid is more consistent than the X-ray peak between hydrodynamical runs with different baryonic physics. This supports our way of locating the X-ray center with eSASS, which accounts for the overall distribution of the emission, rather than simply choosing the brightest pixel.

4.4.3 Physical interpretation of the offset distribution

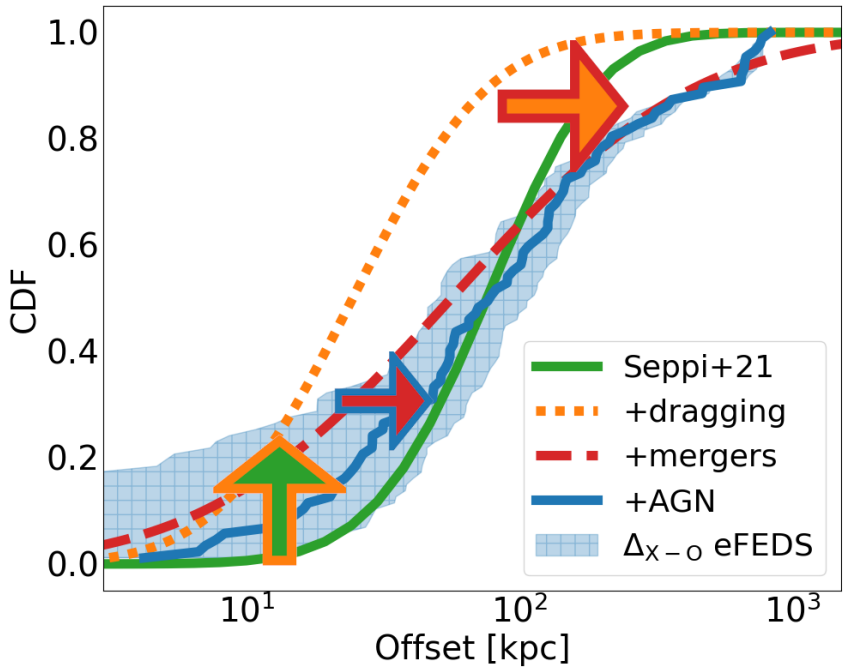

Abbildung 4.3: Illustration showing the interpretation of the impact of different astrophysical effects on the observed offset distribution. The green line refers to the dark matter only scenario (see Fig. 4.1). The dragging due to ram pressure and baryon friction increases the number of clusters with a small offset and is displayed in orange. This causes a shift of the CDF, which is highlighted by the green arrow with an orange edge. Major and minor mergers are responsible for the largest offsets, which further shift the right-hand tail of the distribution (in red) compared to the DMO case. This second transition is highlighted by the orange line with a red edge. Finally, AGN feedback increases the observed offsets in a medium regime, reducing the number of clusters with a small offset. The final result is the Δ_{X-O} distribution measured in eFEDS (see Fig. 4.1). It is shown in blue and the third transition is displayed by the red arrow with a blue edge.

We combine the discussion from the previous paragraphs and formulate a physically

motivated interpretation of the observed offset distributions presented in Fig. 4.1. We use the illustration in Fig. 4.3 to qualitatively guide the discussion. The green line showing the DMO analytical model, and the blue line with shaded area denoting the eFEDS result are the same as in Fig. 4.1. We interpret the shift of the distribution from the DMO scenario to the observations due to different astrophysical phenomena. First, the addition of small scale baryonic effects such as dragging, ram pressure, and friction reduces the observed offsets compared to the DMO case. This is likely to happen in minor mergers, where the gas distribution is not catastrophically disrupted, but the gas ends up trailing the dark matter component of the merging objects. The baryon dragging is reflected in an increment of the CDF at small offsets (orange line), shown by the green arrow with an orange edge. In addition, complex and major mergers can significantly disrupt the gas distribution or even strip the central galaxy from the bottom of the potential well, resulting in larger offsets compared to the DMO scenario. This causes a shift of the right-hand side of the CDF towards larger values, from the green and orange lines to the red one. The transition is highlighted by the orange arrow with a red edge. Furthermore, the AGN feedback alters the gas distribution, reducing the number of clusters with a small predicted offset, as shown by Cui et al. (2016). The final CDF is therefore more skewed towards larger offsets, following the red arrow with a blue edge. The final result is the offset distribution measured in the eFEDS subsample. It includes all these contributions and is shown by the blue line. The final CDF grows less rapidly compared to the dark matter only case, which is what we find when comparing $\Delta_{\mathrm{X-O}}$ to the analytical DMO model (see Fig. 4.1). Very large samples in future eROSITA all-sky surveys will allow a more detailed study of cluster relaxation and observed offsets at fixed cluster properties such as mass and redshift.

4.4.4 Discrepancy between offsets in simulations

The different offsets predicted by TNG and Magneticum may have different causes. First, the Magneticum and TNG simulations are run assuming different cosmologies (see Table 4.1). In particular, the WMAP cosmology assumed for Magneticum is slower in producing collapsed structures, due to the smaller Ω_{M} and σ_8 compared to the Planck cosmology in TNG. Therefore, at a fixed redshift, the merger rate is different between the two simulations: halos have merged more recently in Magneticum because the growth factor is proportional to the matter density in the Universe. It makes them more disturbed, which may additionally contribute to the larger offsets predicted by Magneticum compared to TNG.

An additional factor is the AGN feedback scheme (Hirschmann et al., 2014; Weinberger et al., 2018). The implementation of the accretion onto SMBHs is similar in these two simulations. It is based on an Eddington-limited Bondi accretion rate, following the Bondi–Hoyle–Lyttleton approximation (Bondi and Hoyle, 1944; Bondi, 1952), and accounts for a two-way accretion mode, transitioning from a high accretion state (quasar mode), characterized by the presence of a thin disk, where the feedback is inefficient and released into the surrounding gas as thermal energy, to a low accretion state (radio mode), characterized by the quiescent infall of gas from the hot halo in quasi-hydrostatic equilibrium. In this

case, the feedback is more efficient, and powerful radio jets are produced, that heat the gas kinetically (see Croton et al., 2006; Fanidakis et al., 2011).
AGN feedback models reproduce the majority of AGN observations but struggle to perfectly grasp the full wealth of observed properties (Biffi et al., 2018; Comparat et al., 2019). Detailed predictions should therefore be taken with caution. Nonetheless, different choices of the parameters in the feedback prescription may explain the larger offsets predicted by Magneticum compared to TNG. For example, the transition between the quasar mode and the radio mode, based on a choice of the Eddington ratio between accretion rate and Eddington limit, follows different thresholds. This is fixed at 1% in Magneticum. In TNG instead, a black hole mass-dependent threshold is chosen, such that its value is smaller than 1% for $M_{BH} \lessapprox 10^{8.4}$ $M_\odot$, and can reach larger values of 10% only for the most massive black holes of 10^9 $M_\odot$. Therefore, the radio mode where the feedback is more efficient is active for longer accretion phases in Magneticum compared to TNG. The gas may be ultimately pushed out to smaller distances in TNG, causing the lower values of the predicted offsets. In addition, other differences may impact the modeling of AGN feedback in relation to the offsets. For example, the feedback efficiency in the thermal mode is slightly larger in Magneticum (0.03) than TNG (0.02). The efficiency in the kinetic mode is fixed in Magneticum (0.1), while in TNG it depends on the local density of the environment, which makes the coupling between AGN feedback and gas weaker in low density regions. In both cases, the gas may experience a push out to larger distances in Magneticum. These different prescriptions lead to a redshift-dependent switch of the feedback mode in TNG. Moreover, a black hole seed of mass 1.18×10^6 $M_\odot$ is generated at the center of dark matter halos more massive than 7.38×10^{10} $M_\odot$ in TNG (Weinberger et al., 2017). In Magneticum instead, the assignment is based on the stellar mass of the halo, as a seed black hole of 4.55×10^6 $M_\odot$ is placed at the position of the most bound stellar particle in halos with stellar mass $M_\star \gtrapprox 1.4\times10^{10}$ $M_\odot$. Finally, Magneticum also allows the accretion of fractions of each gas particle onto the SMBH, providing a more continuous representation of the accretion process.

4.4.5 Cosmology with offsets

The halo mass function model developed by Seppi et al. (2021) allows marginalization of the halo abundance on variables related to the dynamical state of dark matter halos, such as $X_{\rm off}$. This mitigates related selection effects. For example, some X-ray-selected cluster samples are affected by the cool core bias (Eckert et al., 2011). Relaxed clusters where the gas has cooled in the central region exhibit a peaked emission in the core. It potentially biases the detection towards such objects, compared to non-cool core ones, where the emissivity profile is flatter. We propose to use the offset between X-ray and optical centers as an observable to link real data to $X_{\rm off}$. This has the potential to enable a cosmological cluster count experiment as a function of mass and observed offset, unbiased by selection effects related to the cluster dynamical state.
We study the relation between $\Delta_{\rm X-O}/R_{\rm 500c}$ in simulations and the offset parameter $X_{\rm off}$. We use the value of $X_{\rm off}$ measured with ROCKSTAR on the halos in the DMO parent

Tabelle 4.2: Best fit parameters for the relation between the displacement between the gas and the central galaxy and the offset parameter in simulations.

	A	B
Magneticum Box2/hr	1.19±0.24	0.46±0.27
TNG-300-1	1.34±0.59	0.56±0.63

simulation. This connects observable properties related to the gas and stars in clusters to intrinsic properties of halos in the model calibrated on N-body simulations (Seppi et al., 2021). There is a positive correlation between $\Delta_{\mathrm{X-O}}$ and $\mathrm{X_{off}}$. Disturbed clusters with a large offset parameter in N-body simulations also exhibit a large displacement between the gas and the central galaxy in the respective hydro run. We model the correlation between the offset predicted by Magneticum and TNG to $\mathrm{X_{off}}$ with a power-law relation (Eq. 4.8):

$$\log_{10} \frac{\Delta_{\mathrm{X-O}}}{R_{500c}} = A \times \log_{10} X_{\mathrm{off}} + B. \tag{4.8}$$

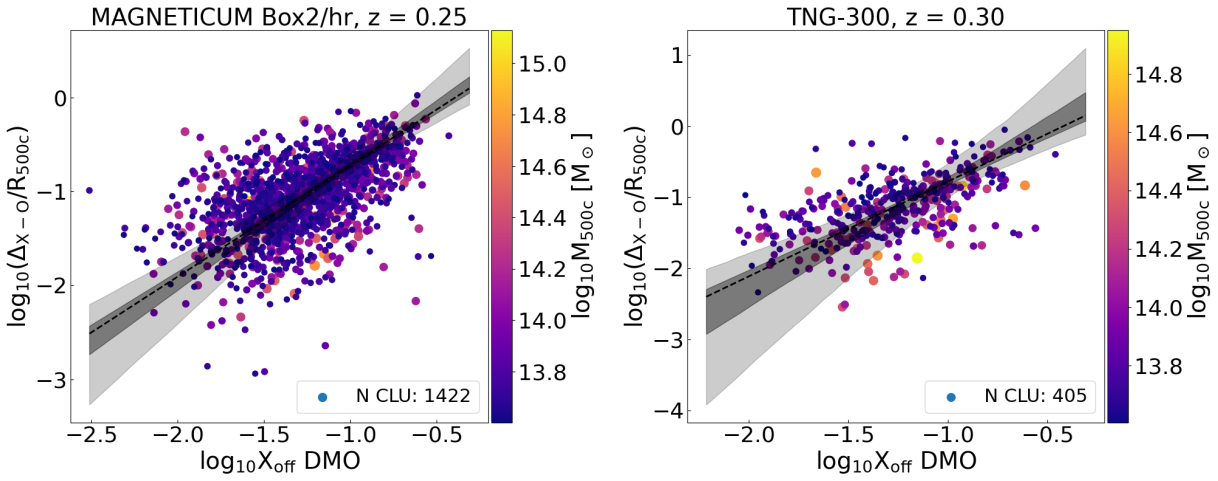

Abbildung 4.4: Average relation between the displacement $\Delta_{\mathrm{X-O}}$ predicted by hydrodynamical simulations and the offset parameter $\mathrm{X_{off}}$ measured on their parent DMO run. It allows linking an observable to an intrinsic property of dark matter halos. The upper panel shows the Magneticum Box2/hr simulation, the bottom panel refers to TNG-300. Each color denotes a specific mass range, the number of clusters in each bin are written in the legend. The shaded areas denote the standard deviation in each bin. The black dashed line shows the best-fit model, the black shaded areas denote the 1σ uncertainty on the model. The bottom-right legend shows the total number of clusters in each mass bin.

There is no explicit trend as a function of mass in Eq. 4.8. Indeed the mass dependence is in the normalization of $\Delta_{\mathrm{X-O}}$ to $\mathrm{R_{500c}}$ and of the offset parameter to the virial radius. Therefore, we fit halos of different masses together. We perform the fitting of Eq. 4.8 using

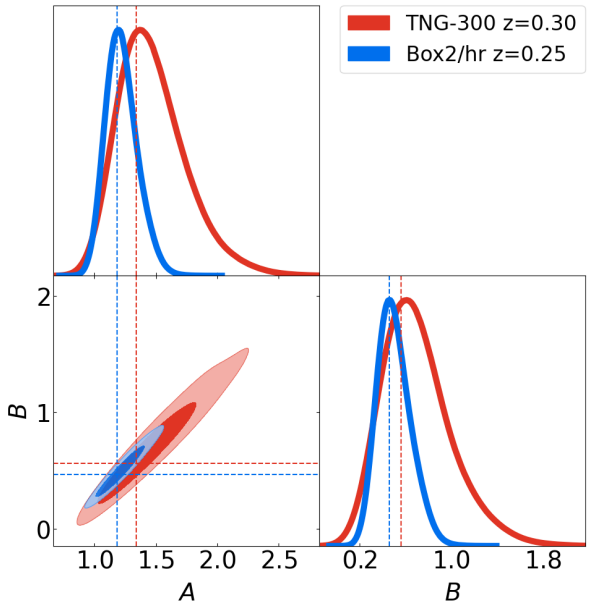

Abbildung 4.5: Triangular plot showing the posterior distributions of the best-fit parameters relating $X_{\rm off}$ to the offset between the gas center and the central galaxy in simulations (see Eq. 4.8). The red (blue) lines and contours show the TNG-300 (Magneticum-Box2) simulation. The shaded areas of the bottom-left panel denote the 1σ and 2σ confidence level contours.

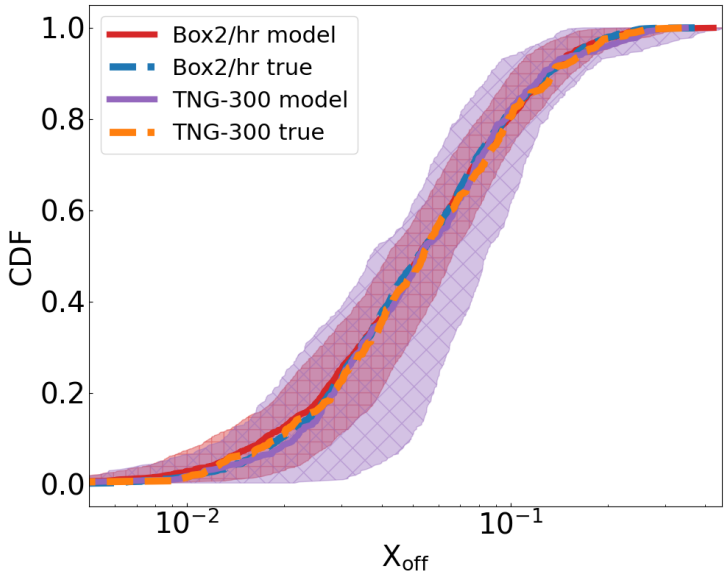

Abbildung 4.6: Recovery of the theoretical $X_{\rm off}$ distribution using the $\Delta_{\rm X-O}$–$X_{\rm off}$ relation. The application of Eq. 4.8 on the offset predicted by Magneticum and TNG is shown in red and violet, respectively. The shaded areas contain the 1σ uncertainty on the model of the $\Delta_{\rm X-O}$–$X_{\rm off}$ relation. The blue (orange) dashed line refer to the direct measure of $\rm X_{off}$ in the DMO counterpart of Magneticum (TNG). We find excellent agreement between the distribution of the true $\rm X_{off}$ and the prediction of Eq. 4.8.

the UltraNest[6] package (Buchner, 2019, 2021). We fit all individual halos more massive than $\rm M_{500c}$>4×10^{13} $\rm M_\odot$. We assume a Poisson likelihood of the form $\log \mathcal{L} = -\sum M + \sum D \times \log M$, where M, D represent the model and the data, respectively. The $\Delta_{\rm X-O}/\rm R_{500c}$ to $\rm X_{off}$ relation is presented in Fig. 4.4. The left-hand (right-hand) panel shows the Magneticum (TNG) simulation. The figure is color-coded according to mass, spanning from low mass groups with $\rm M_{500c}$=4×10^{13} $\rm M_\odot$ to the most massive clusters with $\rm M_{500c}$>1×10^{15} $\rm M_\odot$. The black dashed lines show the best-fit model and the black shaded areas contain the 1σ and 2σ uncertainties on the model. The Magneticum simulation shows a tighter constraint on the relation compared to TNG. Indeed, Magneticum contains a larger amount of halos than TNG thanks to its larger volume, which enables more precise modeling of the $\Delta_{\rm X-O}$–$X_{\rm off}$ relation. The best fit parameters are shown in Table 4.2. The slope and the normalization of the relation (parameters A and B) are compatible between Magneticum and TNG. The full 2D and the marginalized 1D posterior distributions are shown in Fig. 4.5. The blue contours denoting the Magneticum simulation span a smaller area on the A–B plane compared to the red ones, showing the TNG-300 box, because of the different amount of halos in the two simulations.

Starting from the measure of $\Delta_{\rm X-O}/\rm R_{500c}$ in the simulations, we predict the $\rm X_{off}$ distribution in Magneticum and TNG by inverting the model in Eq. 4.8. We find an excellent

[6] https://johannesbuchner.github.io/UltraNest/

agreement between the distribution of X_{off} measured on the DMO counterparts of the hydrodynamical simulations with the prediction obtained from the X-ray to optical offset and inverting Eq. 4.8. The result is shown in Fig. 4.6. The true X_{off} CDF is shown in blue for Magneticum and in orange for TNG. The red and violet lines denote the prediction of X_{off} from Δ_{X-O}. The shaded areas include the 1σ uncertainty on the model in Eq. 4.8. Our prediction of X_{off} from the X-ray to optical offset is able to recover the true X_{off} distribution with great precision. It enables the direct mapping of an observable offset to the offset parameter in DMO simulations, providing a reliable estimator of the mass–X_{off} function $g(\sigma(M), X_{off})$ (Seppi et al., 2021).
In a full end-to-end cosmological study as a function of cluster mass and observed offset, one can marginalize over the parameters of the relation in Eq. 4.8, similarly to the standard way of marginalizing over the mass observable scaling relation parameters in recent cosmological analysis with clusters of galaxies (Mantz et al., 2015b; Bocquet et al., 2019; Ider Chitham et al., 2020).

4.5 Summary

The eROSITA X-ray telescope is detecting clusters of galaxies at an unprecedented rate. It provides a large sample of clusters and groups to study astrophysical properties and constrain cosmological parameters. A key aspect of galaxy cluster studies is the definition of their center.

In this chapter, we measure the offset between the X-ray and optical centers for clusters observed by eROSITA. We consider two cluster catalogs: the eFEDS and the eRASS1 samples. We study two possible definitions of the optical center: the one provided by the centering algorithm of redMaPPer, and the position of the member galaxy with the largest membership probability. On average, the offsets measured in eRASS1 are larger compared to eFEDS (see Fig. 4.2). This is a consequence of the shallower exposure of eRASS1. It does not allow detecting lower mass clusters with smaller observed offsets, that are instead present in the eFEDS sample. However, interpreting the offsets distributions for the whole samples is complicated, because they include clusters at different evolutionary phases, with various mass and redshift.
Therefore, we select a well controlled subsample of eFEDS ($0.15 < z < 0.4$ and $1\times10^{14} < M_{500c} < 8\times10^{14}$ $M_\odot$), where the masses have been measured using weak lensing. We measure an average observed offset $\Delta_{X-O}=76.3^{+30.1}_{-27.1}$ kpc. Using a threshold of $\Delta_{X-O} < 0.05 \times R_{500c}$ (Ota et al., 2022) to select relaxed systems, we measure a relaxed fraction of 31%. According to this criterion, the eFEDS subsample does not show a preference for relaxed clusters, in agreement with Ghirardini et al. (2021a); Bulbul et al. (2022).
We compare the offsets measured in this sample to those predicted by the Magneticum Box2/hr and TNG-300 hydrodynamical simulations, and the N-body model of the offset parameter from Seppi et al. (2021). In the hydrodynamical simulations, we locate the optical center at the position of the main subhalo and the X-ray center at the center of mass of the hot gas particles weighted by their mass and density. The result is shown in Fig. 4.1.

We find a broad agreement between them, especially for the median of the distribution (i.e., the 50% percentile point of the CDF). However, the tails of the distributions are different. We find that the offsets measured in eROSITA data and predicted by Magneticum and TNG are smaller (larger) compared to the N-body model in the low (high) offset regime. This inconsistency is caused by baryons, that reduce the offsets for relaxed systems due to cooling and dragging and increase it for disturbed ones, mainly due to mergers and secondly AGN feedback . This scenario agrees with other work on observations (Churazov et al., 2003; George et al., 2012; Zenteno et al., 2020) and simulations (Molnar et al., 2012; Zhang et al., 2014; Cui et al., 2016). We also find that considering the optical center provided by redMaPPer instead of the galaxy with the largest membership probability provides a better comparison with simulations.

We explore the possibility to introduce the offsets in a cosmological cluster count experiment. We use the displacement between X-ray and optical centers as a proxy for $X_{\rm off}$. We fit a power-law relation between them (see Figs. 4.4 and 4.5). We find a relation that is independent of mass, thanks to the normalization to the cluster size ($R_{\rm 500c}$ for $\Delta_{\rm X-O}$ and $\rm R_{vir}$ for $\rm X_{off}$). Our model of the $\Delta_{\rm X-O}$–$\rm X_{off}$ relation provides a precise prediction of the true $\rm X_{off}$ distribution in Mangeticum and TNG. It is then possible to measure the cluster abundance as a function of mass and observed offset and use the model from Seppi et al. (2021) to constrain cosmological parameters. The best-fit parameters of the $\Delta_{\rm X-O}$–$X_{\rm off}$ relation can be marginalized over similarly to the common mass observable scaling relation. This allows marginalizing over selection effects related to the cluster dynamical state directly in the measure of the halo mass function.

4.6 Images

Figure 4.7 shows six eFEDS clusters. The panels show clusters as seen in the optical band by HSC, using g, r, and z bands. The green dashed lines denote the 3σ contours of the X-ray emission. The green cross identifies the X-ray center from eSASS. Two definitions of the optical center are shown: the galaxy with the largest $\rm p_{mem}$ (blue cross), and the optical center identified by the centering algorithm of redMaPPer (pink plus sign). In some cases, different characterizations of the center are in agreement. The cluster in the bottom-left panel is an example. In other cases, the offsets are larger than hundreds of kpc, especially when the X-ray morphology has a complex structure (central-left panel in Figure 4.7).

Figure 4.8 shows three of the most massive clusters in the snapshot at z=0.25 of the Magneticum simulation. Each row displays one cluster. The left-hand column shows the hot gas component, the central one the stars, and the right-hand one the total matter distribution. The first cluster is relaxed: the distribution of the total matter is close to spherical and the hot gas is smooth. The cluster in the second row is in a transition phase, with a clear peak of the matter distribution in the center and some infalling satellites, that are also well traced by the star component. Finally, the third cluster is in the merging process: two main structures are colliding in the center, the stars are aligned along the merger direction and the gas distribution is more clumpy.

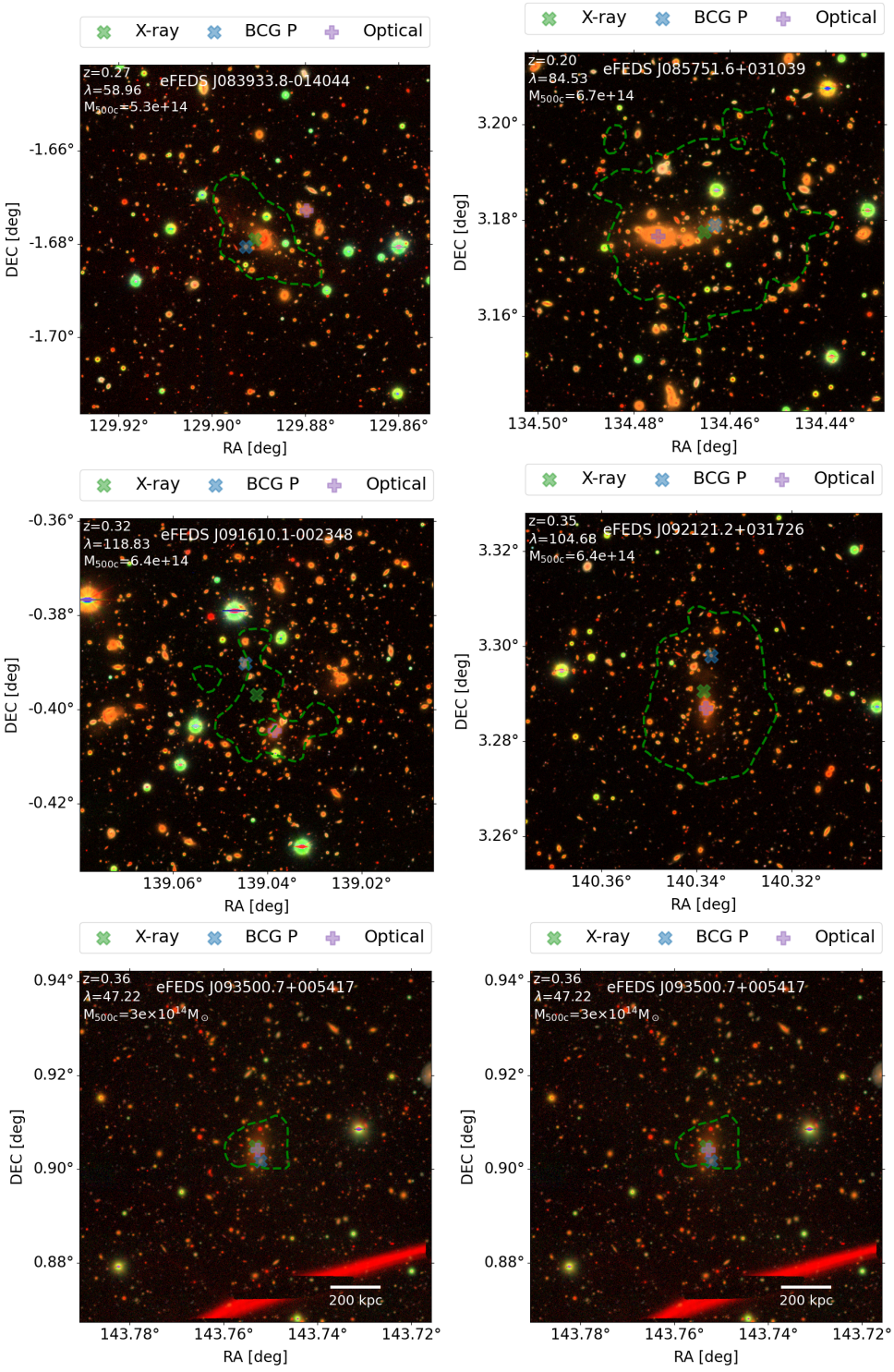

Abbildung 4.7: eFEDS clusters. Each panel shows a different cluster with the relative name, redshift (z), richness (λ), and mass (M_{500c}). The optical image is an RGB cube built from HSC data using g, r, and z bands. The green lines denote the 3σ contours of the X-ray emission. The green cross identifies the X-ray center found by eSASS, the blue one denotes the position of the galaxy with the largest membership probability, and the pink plus sign locates the optical center identified the centering algorithm of redMaPPer.

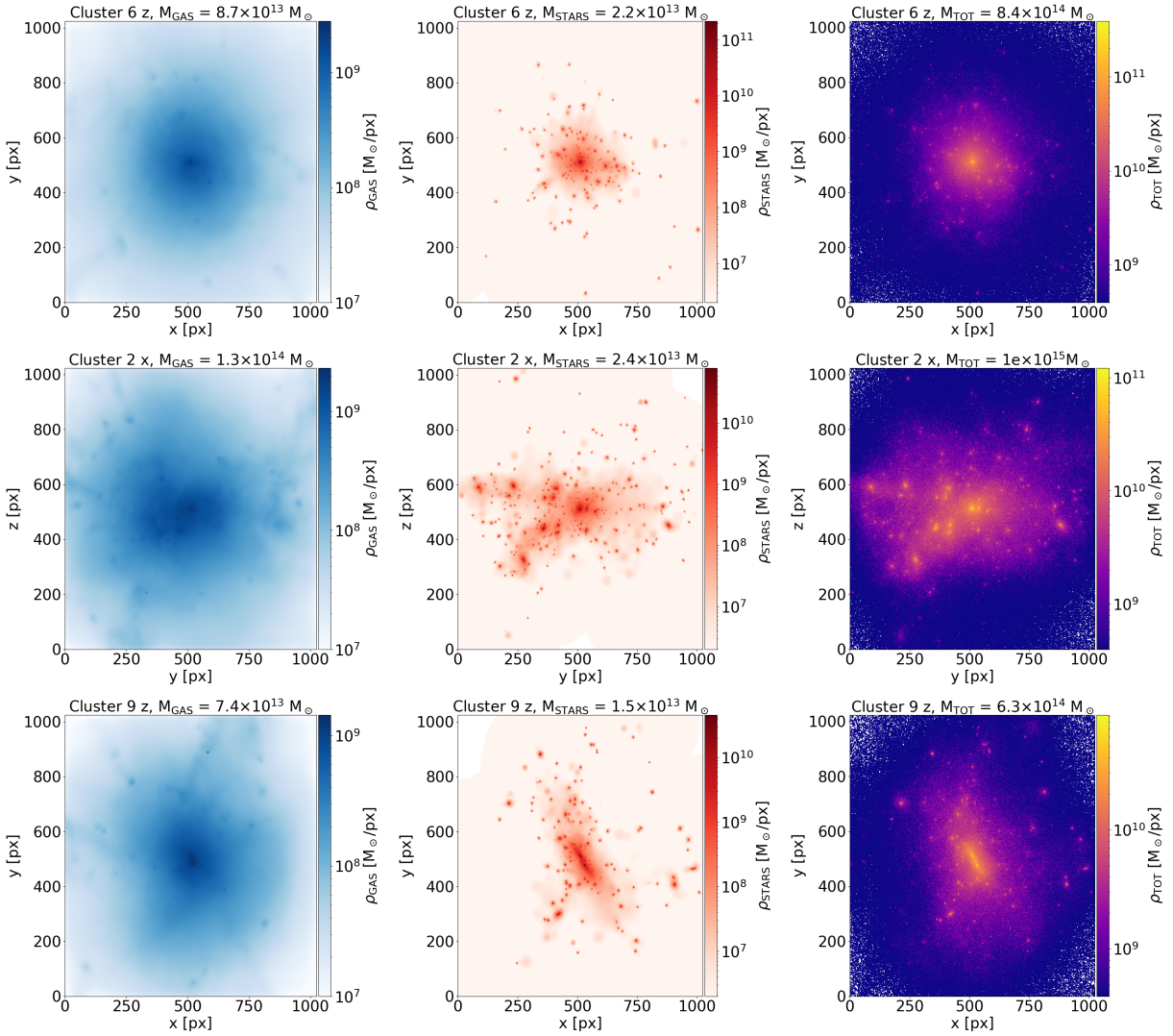

Abbildung 4.8: Clusters of galaxies in the Magneticum Box2/hr simulation at z=0.25. Each row displays one object. Each column shows a different component: the hot gas in the left-hand column, the stars in the middle one, and the total matter distribution in the right-hand one. The panels are $4\times R_{500c}$ large. The side of the maps contains 1024 pixels, so the pixel size is $\sim 3.9\times10^{-3} R_{500c}$. The figures are color coded according to the density of each component in units of solar masses per pixel. The first and third clusters are projected along the z axis, the middle one along the x axis. The mass of each component inside R_{500c} is reported in the title of each panel.

Kapitel 5

Conclusions

In this final chapter, we summarize the main results of the thesis and their role within the development of eROSITA science, including clusters of galaxies in cosmological and astrophysical contexts.

5.1 Summary and outlook

Clusters of galaxies are a great cosmological probe for studying dark matter and dark energy. The cluster abundance as a function of mass and redshift is mostly sensitive to the total amount of matter in the Universe and the amplitude of initial density perturbations at early times, but also to the dark energy equation of state. The eROSITA telescope is revolutionizing X-ray astronomy and will provide the largest X-ray-selected sample of clusters and groups, with great cosmological potential. Understanding uncertainties in the data and characterizing the samples of detected sources is essential to reaching the desired levels of precision. When comparing the abundance of clusters in observations with theoretical models, it is key to account for faint clusters that are not bright enough to be detected (completeness), random background fluctuations that are mistakenly detected as physical sources (spurious detections), and bright point-like sources, such as AGN and stars, that are classified as extended and wrongly included in the cluster sample (contamination). A cluster cosmological experiment will produce biased results if not accounting for these effects properly.
The first goal of this thesis was to study these aspects by producing and analyzing a digital twin of the first eROSITA all-sky survey. We use precise models for the populations of clusters and AGN, that reproduce their number density distributions compared to observations. We include a model of the X-ray background that is resampled from the real eRASS1 data. The combination of these models provides a very similar catalog to the real one. We find that a cut in detection likelihood reduces the fraction of spurious sources, while a cut in extension likelihood is necessary to lower contamination. We find that the fraction of detected clusters mainly depends on flux and exposure time, with secondary trends as a function of the central emissivity and the size of the cluster. We study the trade-off bet-

ween completeness and contamination as a function of extension likelihood cuts. This is key to selecting a cluster sample for cosmological purposes that is as complete and pure as possible. Different improvements on eROSITA simulations are possible in the future. The constant progress in computational power allows the production of large simulations with excellent resolution. For example, the implementation of the Uchuu simulations (Ishiyama et al., 2021) within our framework would provide an extremely detailed description of the dark matter halos hosting AGN, but also a proper characterization of galaxies, without suffering significantly from cosmic variance. The development of an optical model together with the X-ray one presented in this thesis would allow combined modeling of the optical and X-ray selection functions. In addition, multiple realizations would increase the number of objects, providing a more accurate measure of the completeness, and the assignment of detection probabilities, by checking the fraction of simulations in which a source is detected.

The very large sample of clusters observed by eROSITA will provide stringent constraints on the mass-observable scaling relation. Therefore, the contribution of the uncertainties on the models within the total error budget becomes relevant in a cosmological experiment. In this era of precision cosmology, detailed modeling of the halo mass function is required.

The second goal of this thesis was to use large N-body simulations to calibrate a halo mass function that additionally includes variables related to the dynamical state of dark matter halos. We study the concentration–mass relation of massive dark matter halos and provide a model that predicts concentration as a function of mass and redshift with a single equation. We validate a general mass function framework that includes the dynamical state of dark matter halos. It allows measuring their abundance as a function of mass, redshift, spin and offset parameter, with the possibility to integrate out the preferred combination of these variables. It allows marginalizing over selection effects related to the dynamical state of clusters directly in the measure of the halo mass function. Future improvements on such modeling may be achieved with different sets of simulations, for example spanning different cosmologies, or adding baryons. One could then better understand and quantify if these effects have an impact on the halo number density and on the mean relations using concentration, spin, and offset parameter.

The cluster dynamical state has been measured with different methods, such as central density, concentration, centroid shift, or photon asymmetry. An interesting quantity in this context is the offset between the cluster center in different wavelengths.

The third goal of this thesis was to measure the offset between the X-ray and optical center of clusters detected by eROSITA. We find that on average the offset of eRASS1 clusters is larger than the eFEDS ones because the shallower depth of the former does not allow the detection of fainter clusters with low mass and small offset. In addition, we use the Magneticum and Illustris-TNG simulations as a comparison to eROSITA data. In general, we find good agreement between the offset measured in observations, the one predicted by hydrodynamical simulations, and the offset parameter predicted by marginalizing our halo mass function model on mass and spin. We find a different behavior of the tails of the offset distributions, which we attribute to baryonic effects that are absent in dark matter-only simulations. We additionally propose to link the dynamical state in observations to

N-body models using the X-ray to optical offset. We demonstrate that the calibration of a power law relation between them in hydrodynamical simulations provides an excellent recovery of the overall distribution of the offset parameter. This is key for future cluster count cosmological experiments. Given a selection function calibrated on simulations that includes the offset parameter, this approach would allow marginalizing on the different detection probability of clusters with diverse dynamical state.

In this thesis we use observations and simulations to study astrophysics and cosmology with clusters of galaxies. The mass function model and the link between the offset parameter and the displacement between X-ray and optical centers is of particular interest for future cosmological experiments with the largest samples of clusters, such as the one provided by eROSITA. Finally, the estimates of completeness and purity using simulations, such as the one described in this thesis, are key to measuring unbiased cosmological parameters with eROSITA clusters.

5.2 Acknowledgements

This work is based on data from eROSITA, the soft X-ray instrument aboard SRG, a joint Russian-German science mission supported by the Russian Space Agency (Roskosmos), in the interests of the Russian Academy of Sciences represented by its Space Research Institute (IKI), and the Deutsches Zentrum für Luft- und Raumfahrt (DLR). The SRG spacecraft was built by Lavochkin Association (NPOL) and its subcontractors, and is operated by NPOL with support from the Max Planck Institute for Extraterrestrial Physics (MPE).

The development and construction of the eROSITA X-ray instrument was led by MPE, with contributions from the Dr. Karl Remeis Observatory Bamberg & ECAP (FAU Erlangen-Nuernberg), the University of Hamburg Observatory, the Leibniz Institute for Astrophysics Potsdam (AIP), and the Institute for Astronomy and Astrophysics of the University of Tübingen, with the support of DLR and the Max Planck Society. The Argelander Institute for Astronomy of the University of Bonn and the Ludwig Maximilians Universität Munich also participated in the science preparation for eROSITA.

The eROSITA data shown here were processed using the eSASS/NRTA software system developed by the German eROSITA consortium.

The CosmoSim database used in this work is a service by the Leibniz-Institute for Astrophysics Potsdam (AIP). The MultiDark database was developed in cooperation with the Spanish MultiDark Consolider Project CSD2009-00064.

The authors gratefully acknowledge the Gauss Centre for Supercomputing e.V. (www.gauss-centre.eu) and the Partnership for Advanced Supercomputing in Europe (PRACE, www.prace-ri.eu) for funding the MultiDark simulation project by providing computing time on the GCS Supercomputer SuperMUC at Leibniz Supercomputing Centre (LRZ, www.lrz.de).

The author gratefully acknowledges the Gauss Centre for Supercomputing e.V. (www.gauss-centre.eu) for funding this project by providing computing time on the GCS Supercomputer SuperMUC at Leibniz Supercomputing Centre (www.lrz.de).

We derive posterior probability distributions and the Bayesian evidence with the ne-

sted sampling Monte Carlo algorithm MLFriends (Buchner, 2014, 2019, 2021) using the UltraNest software.

We additionally use a χ^2 minimization algorithm with CURVE_FIT python package (Virtanen et al., 2020) (`https://scipy.org/`).

In order to perform cosmology calculation, we used the python toolkit COLOSSUS (Diemer, 2018) (`https://bdiemer.bitbucket.io/colossus/`), the ASTROPY package (Astropy Collaboration et al., 2013, 2018), and the Core Cosmology Library package (CCL, Chisari et al., 2019a).

Literaturverzeichnis

Abazajian, K., Addison, G., Adshead, P. et al. CMB-S4 Science Case, Reference Design, and Project Plan. *arXiv e-prints*, art. arXiv:1907.04473, July 2019.

Abbott, T.M.C., Aguena, M., Alarcon, A. et al. Dark Energy Survey Year 1 Results: Cosmological constraints from cluster abundances and weak lensing. *Phys. Rev. D*, 102 (2):023509, July 2020. doi: 10.1103/PhysRevD.102.023509.

Abell, G.O. The Distribution of Rich Clusters of Galaxies. *ApJS*, 3:211, May 1958. doi: 10.1086/190036.

Abell, G.O., Corwin, Harold G., J. and Olowin, R.P. A Catalog of Rich Clusters of Galaxies. *ApJS*, 70:1, May 1989. doi: 10.1086/191333.

Achitouv, I., Wagner, C., Weller, J. et al. Computation of the halo mass function using physical collapse parameters: application to non-standard cosmologies. *Journal of Cosmology and Astroparticle Physics*, 2014(10):077–077, Oct 2014. ISSN 1475-7516. doi: 10.1088/1475-7516/2014/10/077. URL http://dx.doi.org/10.1088/1475-7516/2014/10/077.

Achitouv, I.E. and Corasaniti, P.S. Non-gaussian halo mass function and non-spherical halo collapse: theory vs. simulations. *Journal of Cosmology and Astroparticle Physics*, 2012 (02):002–002, Feb 2012. ISSN 1475-7516. doi: 10.1088/1475-7516/2012/02/002. URL http://dx.doi.org/10.1088/1475-7516/2012/02/002.

Adami, C., Giles, P., Koulouridis, E. et al. The XXL Survey. XX. The 365 cluster catalogue. *A&A*, 620:A5, November 2018. doi: 10.1051/0004-6361/201731606.

Aguena, M. and Lima, M. Effects of completeness and purity on cluster dark energy constraints. *Phys. Rev. D*, 98(12):123529, December 2018. doi: 10.1103/PhysRevD.98.123529.

Aird, J., Coil, A.L., Georgakakis, A. et al. The evolution of the X-ray luminosity functions of unabsorbed and absorbed AGNs out to z$\sim$ 5. *MNRAS*, 451(2):1892–1927, August 2015. doi: 10.1093/mnras/stv1062.

Alam, S., Aubert, M., Avila, S. et al. Completed SDSS-IV extended Baryon Oscillation Spectroscopic Survey: Cosmological implications from two decades of spectroscopic surveys at the Apache Point Observatory. *Phys. Rev. D*, 103(8):083533, April 2021. doi: 10.1103/PhysRevD.103.083533.

Albrecht, A., Bernstein, G., Cahn, R. et al. Report of the Dark Energy Task Force. *arXiv e-prints*, art. astro-ph/0609591, September 2006.

Allen, S.W., Schmidt, R.W., Ebeling, H. et al. Constraints on dark energy from Chandra observations of the largest relaxed galaxy clusters. *MNRAS*, 353(2):457–467, September 2004. doi: 10.1111/j.1365-2966.2004.08080.x.

Allen, S.W., Evrard, A.E. and Mantz, A.B. Cosmological Parameters from Observations of Galaxy Clusters. *ARA&A*, 49(1):409–470, Sep 2011. doi: 10.1146/annurev-astro-081710-102514.

Anderson, M.E., Gaspari, M., White, S.D.M. et al. Unifying X-ray scaling relations from galaxies to clusters. *MNRAS*, 449(4):3806–3826, June 2015. doi: 10.1093/mnras/stv437.

Andrade-Santos, F., Jones, C., Forman, W.R. et al. The Fraction of Cool-core Clusters in X-Ray versus SZ Samples Using Chandra Observations. *ApJ*, 843(1):76, July 2017. doi: 10.3847/1538-4357/aa7461.

Angulo, R.E., Baugh, C.M. and Lacey, C.G. The assembly bias of dark matter haloes to higher orders. *MNRAS*, 387(2):921–932, June 2008. doi: 10.1111/j.1365-2966.2008.13304.x.

Angulo, R.E., Springel, V., White, S.D.M. et al. Scaling relations for galaxy clusters in the Millennium-XXL simulation. *MNRAS*, 426:2046–2062, November 2012. doi: 10.1111/j.1365-2966.2012.21830.x.

Angulo, R.E. and Hahn, O. Large-scale dark matter simulations. *Living Reviews in Computational Astrophysics*, 8(1):1, December 2022. doi: 10.1007/s41115-021-00013-z.

Arbey, A. and Mahmoudi, F. Dark matter and the early Universe: A review. *Progress in Particle and Nuclear Physics*, 119:103865, July 2021. doi: 10.1016/j.ppnp.2021.103865.

Arcodia, R., Merloni, A., Nandra, K. et al. X-ray quasi-periodic eruptions from two previously quiescent galaxies. *Nature*, 592(7856):704–707, April 2021. doi: 10.1038/s41586-021-03394-6.

Arnaud, K.A. XSPEC: The First Ten Years. In Jacoby, G.H. and Barnes, J., editors, *Astronomical Data Analysis Software and Systems V*, volume 101 of *Astronomical Society of the Pacific Conference Series*, page 17, January 1996.

Ascasibar, Y. and Markevitch, M. The Origin of Cold Fronts in the Cores of Relaxed Galaxy Clusters. *ApJ*, 650(1):102–127, October 2006. doi: 10.1086/506508.

Asgari, M., Lin, C.A., Joachimi, B. et al. KiDS-1000 cosmology: Cosmic shear constraints and comparison between two point statistics. *A&A*, 645:A104, January 2021. doi: 10.1051/0004-6361/202039070.

Astropy Collaboration, Robitaille, T.P., Tollerud, E.J. et al. Astropy: A community Python package for astronomy. *A&A*, 558:A33, October 2013. doi: 10.1051/0004-6361/201322068.

Astropy Collaboration, Price-Whelan, A.M., Sipőcz, B.M. et al. The Astropy Project: Building an Open-science Project and Status of the v2.0 Core Package. *AJ*, 156(3):123, September 2018. doi: 10.3847/1538-3881/aabc4f.

Bahar, Y.E., Bulbul, E., Clerc, N. et al. The eROSITA Final Equatorial-Depth Survey (eFEDS). X-ray properties and scaling relations of galaxy clusters and groups. *A&A*, 661:A7, May 2022. doi: 10.1051/0004-6361/202142462.

Baldi, A.S., De Petris, M., Sembolini, F. et al. Rotation in galaxy clusters from MUSIC simulations with the kinetic Sunyaev-Zeldovich effect. In *Journal of Physics Conference Series*, volume 1226 of *Journal of Physics Conference Series*, page 012003, May 2019. doi: 10.1088/1742-6596/1226/1/012003.

Baldi, A.S., De Petris, M., Sembolini, F. et al. Kinetic Sunyaev-Zel'dovich effect in rotating galaxy clusters from MUSIC simulations. *MNRAS*, 479(3):4028–4040, September 2018. doi: 10.1093/mnras/sty1722.

Balmès, I., Rasera, Y., Corasaniti, P.S. et al. Imprints of dark energy on cosmic structure formation – III. Sparsity of dark matter halo profiles. *Monthly Notices of the Royal Astronomical Society*, 437(3):2328–2339, 11 2013. ISSN 0035-8711. doi: 10.1093/mnras/stt2050. URL https://doi.org/10.1093/mnras/stt2050.

Bardeen, J.M., Bond, J.R., Kaiser, N. et al. The Statistics of Peaks of Gaussian Random Fields. *ApJ*, 304:15, May 1986. doi: 10.1086/164143.

Barnes, D.J., Vogelsberger, M., Kannan, R. et al. A census of cool-core galaxy clusters in IllustrisTNG. *MNRAS*, 481(2):1809–1831, Dec 2018. doi: 10.1093/mnras/sty2078.

Bartelmann, M. and Schneider, P. Weak gravitational lensing. *Phys. Rep.*, 340(4-5):291–472, January 2001. doi: 10.1016/S0370-1573(00)00082-X.

Bartelmann, M. Gravitational lensing. *Classical and Quantum Gravity*, 27(23):233001, Nov 2010. ISSN 1361-6382. doi: 10.1088/0264-9381/27/23/233001. URL http://dx.doi.org/10.1088/0264-9381/27/23/233001.

Baxter, E.J., Sherwin, B.D. and Raghunathan, S. Constraining the rotational kinematic Sunyaev-Zel'dovich effect in massive galaxy clusters. *J. Cosmology Astropart. Phys.*, 2019(6):001, June 2019. doi: 10.1088/1475-7516/2019/06/001.

Behroozi, P., Knebe, A., Pearce, F.R. et al. Major mergers going Notts: challenges for modern halo finders. *MNRAS*, 454(3):3020–3029, December 2015. doi: 10.1093/mnras/stv2046.

Behroozi, P., Wechsler, R. and Wu, H.Y. The ROCKSTAR Phase-space Temporal Halo Finder and the Velocity Offsets of Cluster Cores. *ApJ*, 762:109, January 2013. doi: 10.1088/0004-637X/762/2/109.

Benson, B.A., Ade, P.A.R., Ahmed, Z. et al. *SPT-3G: a next-generation cosmic microwave background polarization experiment on the South Pole telescope*, volume 9153 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, page 91531P. 2014. doi: 10.1117/12.2057305.

Betoule, M., Kessler, R., Guy, J. et al. Improved cosmological constraints from a joint analysis of the SDSS-II and SNLS supernova samples. *A&A*, 568:A22, August 2014. doi: 10.1051/0004-6361/201423413.

Bett, P., Eke, V., Frenk, C.S. et al. The spin and shape of dark matter haloes in the Millennium simulation of a Lambda cold dark matter universe. *Monthly Notices of the Royal Astronomical Society*, 376(1):215–232, 02 2007. ISSN 0035-8711. doi: 10.1111/j.1365-2966.2007.11432.x. URL https://doi.org/10.1111/j.1365-2966.2007.11432.x.

Beutler, F., Blake, C., Colless, M. et al. The 6dF Galaxy Survey: baryon acoustic oscillations and the local Hubble constant. *MNRAS*, 416(4):3017–3032, October 2011. doi: 10.1111/j.1365-2966.2011.19250.x.

Bhattacharya, S., Heitmann, K., White, M. et al. Mass Function Predictions Beyond ΛCDM. *ApJ*, 732:122, May 2011. doi: 10.1088/0004-637X/732/2/122.

Bianconi, M., Ettori, S. and Nipoti, C. Gas rotation in galaxy clusters: signatures and detectability in X-rays. *MNRAS*, 434(2):1565–1575, September 2013. doi: 10.1093/mnras/stt1112.

Biffi, V., Dolag, K. and Böhringer, H. Investigating the velocity structure and X-ray observable properties of simulated galaxy clusters with PHOX. *MNRAS*, 428(2):1395–1409, January 2013. doi: 10.1093/mnras/sts120.

Biffi, V., Borgani, S., Murante, G. et al. On the Nature of Hydrostatic Equilibrium in Galaxy Clusters. *ApJ*, 827(2):112, August 2016. doi: 10.3847/0004-637X/827/2/112.

Biffi, V., Dolag, K. and Merloni, A. AGN contamination of galaxy-cluster thermal X-ray emission: predictions for eRosita from cosmological simulations. *MNRAS*, 481(2):2213–2227, December 2018. doi: 10.1093/mnras/sty2436.

Biffi, V., Dolag, K., Reiprich, T.H. et al. The eROSITA view of the Abell 3391/95 field: Case study from the Magneticum cosmological simulation. *A&A*, 661:A17, May 2022. doi: 10.1051/0004-6361/202141107.

Bilton, L.E., Hunt, M., Pimbblet, K.A. et al. The rotational profiles of cluster galaxies. *MNRAS*, 490(4):5017–5032, December 2019. doi: 10.1093/mnras/stz2927.

Biviano, A., Murante, G., Borgani, S. et al. On the efficiency and reliability of cluster mass estimates based on member galaxies. *A&A*, 456(1):23–36, September 2006. doi: 10.1051/0004-6361:20064918.

Biviano, A. From Messier to Abell: 200 Years of Science with Galaxy Clusters. In Durret, F. and Gerbal, D., editors, *Constructing the Universe with Clusters of Galaxies*, page 1, January 2000.

Blanton, M.R., Bershady, M.A., Abolfathi, B. et al. Sloan Digital Sky Survey IV: Mapping the Milky Way, Nearby Galaxies, and the Distant Universe. *AJ*, 154(1):28, July 2017. doi: 10.3847/1538-3881/aa7567.

Bleem, L.E., Stalder, B., Brodwin, M. et al. A New Reduction of the Blanco Cosmology Survey: An Optically Selected Galaxy Cluster Catalog and a Public Release of Optical Data Products. *ApJS*, 216(1):20, January 2015a. doi: 10.1088/0067-0049/216/1/20.

Bleem, L.E., Stalder, B., de Haan, T. et al. Galaxy Clusters Discovered via the Sunyaev-Zel'dovich Effect in the 2500-Square-Degree SPT-SZ Survey. *ApJS*, 216(2):27, February 2015b. doi: 10.1088/0067-0049/216/2/27.

Blumenthal, G.R., Faber, S.M., Primack, J.R. et al. Formation of galaxies and large-scale structure with cold dark matter. *Nature*, 311:517–525, October 1984. doi: 10.1038/311517a0.

Bocquet, S., Saro, A., Dolag, K. et al. Halo mass function: baryon impact, fitting formulae, and implications for cluster cosmology. *MNRAS*, 456:2361–2373, March 2016. doi: 10.1093/mnras/stv2657.

Bocquet, S., Dietrich, J.P., Schrabback, T. et al. 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, June 2019. doi: 10.3847/1538-4357/ab1f10.

Bocquet, S., Heitmann, K., Habib, S. et al. The Mira-Titan Universe. III. Emulation of the Halo Mass Function. *ApJ*, 901(1):5, September 2020. doi: 10.3847/1538-4357/abac5c.

Boella, G., Butler, R.C., Perola, G.C. et al. BeppoSAX, the wide band mission for X-ray astronomy. *A&AS*, 122:299–307, April 1997. doi: 10.1051/aas:1997136.

Böhringer, H., Voges, W., Huchra, J.P. et al. The Northern ROSAT All-Sky (NORAS) Galaxy Cluster Survey. I. X-Ray Properties of Clusters Detected as Extended X-Ray Sources. *ApJS*, 129(2):435–474, August 2000. doi: 10.1086/313427.

Böhringer, H., Schuecker, P., Guzzo, L. et al. The ROSAT-ESO Flux Limited X-ray (REFLEX) Galaxy cluster survey. V. The cluster catalogue. *A&A*, 425:367–383, October 2004. doi: 10.1051/0004-6361:20034484.

Boller, T., Freyberg, M.J., Trümper, J. et al. Second ROSAT all-sky survey (2RXS) source catalogue. *A&A*, 588:A103, April 2016. doi: 10.1051/0004-6361/201525648.

Bonafede, A., Feretti, L., Murgia, M. et al. The Coma cluster magnetic field from Faraday rotation measures. *A&A*, 513:A30, April 2010. doi: 10.1051/0004-6361/200913696.

Bond, J.R., Cole, S., Efstathiou, G. et al. Excursion set mass functions for hierarchical Gaussian fluctuations. *ApJ*, 379:440–460, October 1991. doi: 10.1086/170520.

Bondi, H. On spherically symmetrical accretion. *MNRAS*, 112:195, January 1952. doi: 10.1093/mnras/112.2.195.

Bondi, H. and Hoyle, F. On the mechanism of accretion by stars. *MNRAS*, 104:273, January 1944. doi: 10.1093/mnras/104.5.273.

Borgani, S. *Cosmology with Clusters of Galaxies*, volume 740, page 24. 2008. doi: 10.1007/978-1-4020-6941-3_9.

Borgani, S. and Kravtsov, A. Cosmological Simulations of Galaxy Clusters. *Advanced Science Letters*, 4(2):204–227, February 2011. doi: 10.1166/asl.2011.1209.

Briel, U.G., Henry, J.P. and Boehringer, H. Observation of the Coma cluster of galaxies with ROSAT during the all-sky-survey. *A&A*, 259:L31–L34, June 1992.

Brüggen, M., Bykov, A., Ryu, D. et al. Magnetic Fields, Relativistic Particles, and Shock Waves in Cluster Outskirts. *Space Sci. Rev.*, 166(1-4):187–213, May 2012. doi: 10.1007/s11214-011-9785-9.

Brunetti, G. and Lazarian, A. Stochastic reacceleration of relativistic electrons by turbulent reconnection: a mechanism for cluster-scale radio emission? *MNRAS*, 458(3):2584–2595, May 2016. doi: 10.1093/mnras/stw496.

Brunner, H., Liu, T., Lamer, G. et al. The eROSITA Final Equatorial Depth Survey (eFEDS). X-ray catalogue. *A&A*, 661:A1, May 2022. doi: 10.1051/0004-6361/202141266.

Bryan, G.L. and Norman, M.L. Statistical Properties of X-Ray Clusters: Analytic and Numerical Comparisons. *ApJ*, 495(1):80–99, March 1998. doi: 10.1086/305262.

Buchner, J. A statistical test for Nested Sampling algorithms. *arXiv e-prints*, art. arXiv:1407.5459, July 2014.

Buchner, J. Collaborative Nested Sampling: Big Data versus Complex Physical Models. *PASP*, 131(1004):108005, October 2019. doi: 10.1088/1538-3873/aae7fc.

Buchner, J. UltraNest - a robust, general purpose Bayesian inference engine. *The Journal of Open Source Software*, 6(60):3001, April 2021. doi: 10.21105/joss.03001.

Buchner, J., Georgakakis, A., Nandra, K. et al. Obscuration-dependent Evolution of Active Galactic Nuclei. *ApJ*, 802(2):89, April 2015. doi: 10.1088/0004-637X/802/2/89.

Bulbul, E., Liu, A., Pasini, T. et al. The eROSITA Final Equatorial-Depth Survey (eFEDS). Galaxy clusters and groups in disguise. *A&A*, 661:A10, May 2022. doi: 10.1051/0004-6361/202142460.

Bulbul, E., Chiu, I.N., Mohr, J.J. et al. X-Ray Properties of SPT-selected Galaxy Clusters at $0.2 < z < 1.5$ Observed with XMM-Newton. *ApJ*, 871(1):50, January 2019. doi: 10.3847/1538-4357/aaf230.

Burenin, R.A., Vikhlinin, A., Hornstrup, A. et al. The 400 Square Degree ROSAT PSPC Galaxy Cluster Survey: Catalog and Statistical Calibration. *ApJS*, 172(2):561–582, October 2007. doi: 10.1086/519457.

Cao, K., Barnes, D.J. and Vogelsberger, M. Studying galaxy cluster morphological metrics with Mock-X. *arXiv e-prints*, art. arXiv:2006.10752, June 2020.

Capasso, R., Mohr, J.J., Saro, A. et al. Mass calibration of the CODEX cluster sample using SPIDERS spectroscopy – I. The richness–mass relation. *Monthly Notices of the Royal Astronomical Society*, 486(2):1594–1607, 03 2019. ISSN 0035-8711. doi: 10.1093/mnras/stz931. URL https://doi.org/10.1093/mnras/stz931.

Carlstrom, J.E., Holder, G.P. and Reese, E.D. Cosmology with the sunyaev-zel'dovich effect. *Annual Review of Astronomy and Astrophysics*, 40(1):643–680, 2002. doi: 10.1146/annurev.astro.40.060401.093803. URL https://doi.org/10.1146/annurev.astro.40.060401.093803.

Cash, W. Parameter estimation in astronomy through application of the likelihood ratio. *ApJ*, 228:939–947, March 1979. doi: 10.1086/156922.

Castro, T., Borgani, S., Dolag, K. et al. On the impact of baryons on the halo mass function, bias, and cluster cosmology. *MNRAS*, 500(2):2316–2335, January 2021. doi: 10.1093/mnras/staa3473.

Cavaliere, A. and Fusco-Femiano, R. Reprint of 1976A&A....49..137C. X-rays from hot plasma in clusters of galaxies. *A&A*, 500:95–102, May 1976.

Chisari, N.E., Alonso, D., Krause, E. et al. Core Cosmology Library: Precision Cosmological Predictions for LSST. *ApJS*, 242(1):2, May 2019a. doi: 10.3847/1538-4365/ab1658.

Chisari, N.E., Mead, A.J., Joudaki, S. et al. Modelling baryonic feedback for survey cosmology. *The Open Journal of Astrophysics*, 2(1):4, June 2019b. doi: 10.21105/astro.1905.06082.

Chiu, I.N., Umetsu, K., Murata, R. et al. The richness-to-mass relation of CAMIRA galaxy clusters from weak-lensing magnification in the Subaru Hyper Suprime-Cam survey. *MNRAS*, 495(1):428–450, June 2020. doi: 10.1093/mnras/staa1158.

Chiu, I.N., Ghirardini, V., Liu, A. et al. 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, May 2022. doi: 10.1051/0004-6361/202141755.

Chuang, C.H., Yepes, G., Kitaura, F.S. et al. UNIT project: Universe N-body simulations for the Investigation of Theoretical models from galaxy surveys. *MNRAS*, 487(1):48–59, July 2019. doi: 10.1093/mnras/stz1233.

Churazov, E., Forman, W., Jones, C. et al. XMM-Newton Observations of the Perseus Cluster. I. The Temperature and Surface Brightness Structure. *ApJ*, 590(1):225–237, June 2003. doi: 10.1086/374923.

Cibirka, N., Cypriano, E.S., Brimioulle, F. et al. CODEX weak lensing: concentration of galaxy clusters at z $\sim$ 0.5. *MNRAS*, 468(1):1092–1116, June 2017. doi: 10.1093/mnras/stx484.

Clerc, N., Pierre, M., Pacaud, F. et al. The cosmological analysis of X-ray cluster surveys - I. A new method for interpreting number counts. *MNRAS*, 423(4):3545–3560, July 2012. doi: 10.1111/j.1365-2966.2012.21154.x.

Clerc, N., Merloni, A., Zhang, Y.Y. et al. SPIDERS: the spectroscopic follow-up of X-ray selected clusters of galaxies in SDSS-IV. *MNRAS*, 463(4):4490–4515, December 2016. doi: 10.1093/mnras/stw2214.

Clerc, N., Ramos-Ceja, M.E., Ridl, J. et al. Synthetic simulations of the extragalactic sky seen by eROSITA. I. Pre-launch selection functions from Monte-Carlo simulations. *A&A*, 617:A92, September 2018. doi: 10.1051/0004-6361/201732119.

Clerc, N., Kirkpatrick, C.C., Finoguenov, A. et al. SPIDERS: overview of the X-ray galaxy cluster follow-up and the final spectroscopic data release. *MNRAS*, 497(3):3976–3992, July 2020. doi: 10.1093/mnras/staa2066.

Clerc, N. and Finoguenov, A. X-ray cluster cosmology. *arXiv e-prints*, art. arXiv:2203.11906, March 2022.

Clowe, D., Bradač, M., Gonzalez, A.H. et al. A Direct Empirical Proof of the Existence of Dark Matter. *ApJ*, 648(2):L109–L113, September 2006. doi: 10.1086/508162.

Cole, S., Percival, W.J., Peacock, J.A. et al. The 2dF Galaxy Redshift Survey: power-spectrum analysis of the final data set and cosmological implications. *MNRAS*, 362(2): 505–534, September 2005. doi: 10.1111/j.1365-2966.2005.09318.x.

Colless, M., Dalton, G., Maddox, S. et al. The 2dF Galaxy Redshift Survey: spectra and redshifts. *MNRAS*, 328(4):1039–1063, December 2001. doi: 10.1046/j.1365-8711.2001.04902.x.

Comparat, J., Merloni, A., Salvato, M. et al. Active galactic nuclei and their large-scale structure: an eROSITA mock catalogue. *MNRAS*, 487(2):2005–2029, August 2019. doi: 10.1093/mnras/stz1390.

Comparat, J., Prada, F., Yepes, G. et al. Accurate mass and velocity functions of dark matter haloes. *MNRAS*, 469(4):4157–4174, Aug 2017. doi: 10.1093/mnras/stx1183.

Comparat, J., Eckert, D., Finoguenov, A. et al. Full-sky photon simulation of clusters and active galactic nuclei in the soft X-rays for eROSITA. *The Open Journal of Astrophysics*, 3(1):13, December 2020. doi: 10.21105/astro.2008.08404.

Comparat, J., Truong, N., Merloni, A. et al. The eROSITA Final Equatorial Depth Survey (eFEDS): X-ray emission around star-forming and quiescent galaxies at $0.05 < z < 0.3$. *arXiv e-prints*, art. arXiv:2201.05169, January 2022.

Contreras, S., Zehavi, I., Padilla, N. et al. The evolution of assembly bias. *MNRAS*, 484 (1):1133–1148, Mar 2019. doi: 10.1093/mnras/stz018.

Cooray, A. and Chen, X. Kinetic sunyaev-zeldovich effect from halo rotation. *The Astrophysical Journal*, 573(1):43–50, jul 2002. doi: 10.1086/340582. URL https://doi.org/10.1086%2F340582.

Corasaniti, P.S., Ettori, S., Rasera, Y. et al. Probing Cosmology with Dark Matter Halo Sparsity Using X-Ray Cluster Mass Measurements. *ApJ*, 862(1):40, July 2018. doi: 10.3847/1538-4357/aaccdf.

Corasaniti, P.S., Sereno, M. and Ettori, S. Cosmological Constraints from Galaxy Cluster Sparsity, Cluster Gas Mass Fraction, and Baryon Acoustic Oscillation Data. *ApJ*, 911 (2):82, April 2021. doi: 10.3847/1538-4357/abe9a4.

Correa, C.A., Wyithe, J.S.B., Schaye, J. et al. The accretion history of dark matter haloes - III. A physical model for the concentration-mass relation. *MNRAS*, 452(2):1217–1232, September 2015. doi: 10.1093/mnras/stv1363.

Costanzi, M., Rozo, E., Rykoff, E.S. et al. Modelling projection effects in optically selected cluster catalogues. *MNRAS*, 482(1):490–505, January 2019. doi: 10.1093/mnras/sty2665.

Costanzi, M., Saro, A., Bocquet, S. et al. Cosmological constraints from DES Y1 cluster abundances and SPT multiwavelength data. *Phys. Rev. D*, 103(4):043522, February 2021. doi: 10.1103/PhysRevD.103.043522.

Crocce, M., Fosalba, P., Castand er, F.J. et al. Simulating the Universe with MICE: the abundance of massive clusters. *MNRAS*, 403(3):1353–1367, April 2010. doi: 10.1111/j.1365-2966.2009.16194.x.

Croton, D.J., Springel, V., White, S.D.M. et al. The many lives of active galactic nuclei: cooling flows, black holes and the luminosities and colours of galaxies. *MNRAS*, 365(1):11–28, January 2006. doi: 10.1111/j.1365-2966.2005.09675.x.

Croton, D.J., Gao, L. and White, S.D.M. Halo assembly bias and its effects on galaxy clustering. *MNRAS*, 374(4):1303–1309, Feb 2007. doi: 10.1111/j.1365-2966.2006.11230.x.

Cui, W., Power, C., Biffi, V. et al. How does our choice of observable influence our estimation of the centre of a galaxy cluster? Insights from cosmological simulations. *MNRAS*, 456(3):2566–2575, March 2016. doi: 10.1093/mnras/stv2839.

Dauser, T., Falkner, S., Lorenz, M. et al. SIXTE: a generic X-ray instrument simulation toolkit. *A&A*, 630:A66, September 2019. doi: 10.1051/0004-6361/201935978.

Dawson, K.S., Kneib, J.P., Percival, W.J. et al. The SDSS-IV Extended Baryon Oscillation Spectroscopic Survey: Overview and Early Data. *AJ*, 151:44, February 2016. doi: 10.3847/0004-6256/151/2/44.

Dawson, W.A., Wittman, D., Jee, M.J. et al. Discovery of a Dissociative Galaxy Cluster Merger with Large Physical Separation. *ApJ*, 747(2):L42, March 2012. doi: 10.1088/2041-8205/747/2/L42.

de Jong, R. 4MOST — 4-metre Multi-Object Spectroscopic Telescope. *The Messenger*, 145:14–16, September 2011.

de Korte, P.A.J., Bleeker, J.A.M., den Boggende, A.J.F. et al. The X-Ray Imaging Telescopes on EXOSAT. *Space Sci. Rev.*, 30(1-4):495–511, March 1981. doi: 10.1007/BF01246070.

de la Torre, S., Guzzo, L., Peacock, J.A. et al. The VIMOS Public Extragalactic Redshift Survey (VIPERS) . Galaxy clustering and redshift-space distortions at z = 0.8 in the first data release. *A&A*, 557:A54, September 2013. doi: 10.1051/0004-6361/201321463.

De Luca, F., De Petris, M., Yepes, G. et al. The Three Hundred Project: Dynamical state of galaxy clusters and morphology from multi-wavelength synthetic maps. *arXiv e-prints*, art. arXiv:2011.09002, November 2020.

De Propris, R., West, M.J., Andrade-Santos, F. et al. Brightest cluster galaxies: the centre can(not?) hold. *MNRAS*, 500(1):310–318, January 2021. doi: 10.1093/mnras/staa3286.

de Sitter, W. Einstein's theory of gravitation and its astronomical consequences. Third paper. *MNRAS*, 78:3–28, November 1917. doi: 10.1093/mnras/78.1.3.

Debackere, S.N.B., Schaye, J. and Hoekstra, H. How baryons can significantly bias cluster count cosmology. *MNRAS*, 505(1):593–609, July 2021. doi: 10.1093/mnras/stab1326.

Del Popolo, A., Pace, F. and Le Delliou, M. A high precision semi-analytic mass function. *J. Cosmology Astropart. Phys.*, 3:032, March 2017. doi: 10.1088/1475-7516/2017/03/032.

DESI Collaboration, Aghamousa, A., Aguilar, J. et al. The DESI Experiment Part I: Science,Targeting, and Survey Design. *ArXiv e-prints*, October 2016.

Despali, G., Giocoli, C., Angulo, R.E. et al. The universality of the virial halo mass function and models for non-universality of other halo definitions. *MNRAS*, 456:2486–2504, March 2016. doi: 10.1093/mnras/stv2842.

Dey, A., Schlegel, D.J., Lang, D. et al. Overview of the DESI Legacy Imaging Surveys. *AJ*, 157(5):168, May 2019. doi: 10.3847/1538-3881/ab089d.

Di Matteo, T., Colberg, J., Springel, V. et al. Direct Cosmological Simulations of the Growth of Black Holes and Galaxies. *ApJ*, 676(1):33–53, March 2008. doi: 10.1086/524921.

Diemer, B. COLOSSUS: A Python Toolkit for Cosmology, Large-scale Structure, and Dark Matter Halos. *ApJS*, 239(2):35, Dec 2018. doi: 10.3847/1538-4365/aaee8c.

Diemer, B. Universal at Last? The Splashback Mass Function of Dark Matter Halos. *ApJ*, 903(2):87, November 2020. doi: 10.3847/1538-4357/abbf52.

Diemer, B. A dynamics-based density profile for dark haloes - I. Algorithm and basic results. *MNRAS*, 513(1):573–594, June 2022. doi: 10.1093/mnras/stac878.

Diemer, B. and Joyce, M. An Accurate Physical Model for Halo Concentrations. *ApJ*, 871 (2):168, February 2019. doi: 10.3847/1538-4357/aafad6.

Diemer, B. and Kravtsov, A.V. A Universal Model for Halo Concentrations. *ApJ*, 799(1): 108, January 2015. doi: 10.1088/0004-637X/799/1/108.

Dolag, K. The Magneticum Simulations, from Galaxies to Galaxy Clusters. In *IAU General Assembly*, volume 29, page 2250156, August 2015.

Dolag, K., Mevius, E. and Remus, R.S. Distribution and Evolution of Metals in the Magneticum Simulations. *Galaxies*, 5(3):35, August 2017. doi: 10.3390/galaxies5030035.

Driver, S.P., Hill, D.T., Kelvin, L.S. et al. Galaxy and Mass Assembly (GAMA): survey diagnostics and core data release. *MNRAS*, 413:971–995, May 2011. doi: 10.1111/j.1365-2966.2010.18188.x.

Driver, S.P., Norberg, P., Baldry, I.K. et al. GAMA: towards a physical understanding of galaxy formation. *Astronomy and Geophysics*, 50(5):5.12–5.19, October 2009. doi: 10.1111/j.1468-4004.2009.50512.x.

Du, W., Fan, Z., Shan, H. et al. Mass-Concentration Relation of Clusters of Galaxies from CFHTLenS. *ApJ*, 814(2):120, December 2015. doi: 10.1088/0004-637X/814/2/120.

Duffy, A.R., Schaye, J., Kay, S.T. et al. Dark matter halo concentrations in the Wilkinson Microwave Anisotropy Probe year 5 cosmology. *MNRAS*, 390(1):L64–L68, October 2008. doi: 10.1111/j.1745-3933.2008.00537.x.

Dutton, A.A. and Macciò, A.V. Cold dark matter haloes in the Planck era: evolution of structural parameters for Einasto and NFW profiles. *MNRAS*, 441(4):3359–3374, July 2014. doi: 10.1093/mnras/stu742.

Ebeling, H., Edge, A.C., Bohringer, H. et al. The ROSAT Brightest Cluster Sample - I. The compilation of the sample and the cluster log N-log S distribution. *MNRAS*, 301 (4):881–914, December 1998. doi: 10.1046/j.1365-8711.1998.01949.x.

Eckert, D., Molendi, S. and Paltani, S. The cool-core bias in X-ray galaxy cluster samples. I. Method and application to HIFLUGCS. *A&A*, 526:A79, Feb 2011. doi: 10.1051/0004-6361/201015856.

Eckert, D., Ghirardini, V., Ettori, S. et al. Non-thermal pressure support in X-COP galaxy clusters. *A&A*, 621:A40, January 2019a. doi: 10.1051/0004-6361/201833324.

Eckert, D., Ghirardini, V., Ettori, S. et al. Non-thermal pressure support in X-COP galaxy clusters. *A&A*, 621:A40, January 2019b. doi: 10.1051/0004-6361/201833324.

Eckert, D., Gaspari, M., Gastaldello, F. et al. Feedback from Active Galactic Nuclei in Galaxy Groups. *Universe*, 7(5):142, May 2021. doi: 10.3390/universe7050142.

Einasto, M., Liivamägi, L.J., Tempel, E. et al. The Sloan Great Wall. Morphology and Galaxy Content. *ApJ*, 736(1):51, July 2011. doi: 10.1088/0004-637X/736/1/51.

Einstein, A. Die Grundlage der allgemeinen Relativitätstheorie. *Annalen der Physik*, 354 (7):769–822, January 1916. doi: 10.1002/andp.19163540702.

Eisenstein, D.J. and Hu, W. Baryonic Features in the Matter Transfer Function. *ApJ*, 496 (2):605–614, March 1998. doi: 10.1086/305424.

Eisenstein, D.J., Zehavi, I., Hogg, D.W. et al. Detection of the Baryon Acoustic Peak in the Large-Scale Correlation Function of SDSS Luminous Red Galaxies. *ApJ*, 633(2): 560–574, November 2005. doi: 10.1086/466512.

Ettori, S. and Brighenti, F. On the evolution of cooling cores in X-ray galaxy clusters. *MNRAS*, 387(2):631–638, June 2008. doi: 10.1111/j.1365-2966.2008.13054.x.

Ettori, S., Morandi, A., Tozzi, P. et al. The cluster gas mass fraction as a cosmological probe: a revised study. *A&A*, 501(1):61–73, July 2009. doi: 10.1051/0004-6361/200810878.

Euclid Collaboration, Castro, T., Fumagalli, A. et al. Euclid preparation. XXIV. Calibration of the halo mass function in $\Lambda(\nu)$CDM cosmologies. *arXiv e-prints*, art. arXiv:2208.02174, August 2022.

Everett, S., Yanny, B., Kuropatkin, N. et al. Dark Energy Survey Year 3 Results: Measuring the Survey Transfer Function with Balrog. *ApJS*, 258(1):15, January 2022. doi: 10.3847/1538-4365/ac26c1.

Fabian, A.C. Observational Evidence of Active Galactic Nuclei Feedback. *ARA&A*, 50:455–489, September 2012. doi: 10.1146/annurev-astro-081811-125521.

Fabjan, D., Borgani, S., Tornatore, L. et al. Simulating the effect of active galactic nuclei feedback on the metal enrichment of galaxy clusters. *MNRAS*, 401(3):1670–1690, January 2010. doi: 10.1111/j.1365-2966.2009.15794.x.

Fanidakis, N., Baugh, C.M., Benson, A.J. et al. Grand unification of AGN activity in the LambdaCDM cosmology. *MNRAS*, 410(1):53–74, January 2011. doi: 10.1111/j.1365-2966.2010.17427.x.

Ferrarese, L. and Merritt, D. A Fundamental Relation between Supermassive Black Holes and Their Host Galaxies. *ApJ*, 539(1):L9–L12, August 2000. doi: 10.1086/312838.

Finoguenov, A., Guzzo, L., Hasinger, G. et al. The XMM-Newton Wide-Field Survey in the COSMOS Field: Statistical Properties of Clusters of Galaxies. *ApJS*, 172(1):182–195, September 2007. doi: 10.1086/516577.

Finoguenov, A., Tanaka, M., Cooper, M. et al. Ultra-deep catalog of X-ray groups in the Extended Chandra Deep Field South. *A&A*, 576:A130, April 2015. doi: 10.1051/0004-6361/201323053.

Finoguenov, A., Merloni, A., Comparat, J. et al. 4MOST Consortium Survey 5: eROSITA Galaxy Cluster Redshift Survey. *The Messenger*, 175:39–41, March 2019. doi: 10.18727/0722-6691/5124.

Finoguenov, A., Rykoff, E., Clerc, N. et al. CODEX clusters. Survey, catalog, and cosmology of the X-ray luminosity function. *A&A*, 638:A114, June 2020. doi: 10.1051/0004-6361/201937283.

Foëx, G., Motta, V., Jullo, E. et al. SARCS strong-lensing galaxy groups. II. Mass-concentration relation and strong-lensing bias. *A&A*, 572:A19, December 2014. doi: 10.1051/0004-6361/201424706.

Foreman-Mackey, D., Hogg, D.W., Lang, D. et al. emcee: The MCMC Hammer. *PASP*, 125(925):306, March 2013. doi: 10.1086/670067.

Friedmann, A. Über die Krümmung des Raumes. *Zeitschrift fur Physik*, 10:377–386, January 1922. doi: 10.1007/BF01332580.

Gaia Collaboration, Brown, A.G.A., Vallenari, A. et al. Gaia Data Release 2. Summary of the contents and survey properties. *A&A*, 616:A1, August 2018. doi: 10.1051/0004-6361/201833051.

Gao, L., Springel, V. and White, S.D.M. The age dependence of halo clustering. *MNRAS*, 363(1):L66–L70, Oct 2005. doi: 10.1111/j.1745-3933.2005.00084.x.

Garrel, C., Pierre, M., Valageas, P. et al. The XXL survey. XLVI. Forward cosmological analysis of the C1 cluster sample. *A&A*, 663:A3, July 2022. doi: 10.1051/0004-6361/202141204.

Gaspari, M., Ruszkowski, M. and Sharma, P. Cause and Effect of Feedback: Multiphase Gas in Cluster Cores Heated by AGN Jets. *ApJ*, 746(1):94, February 2012. doi: 10.1088/0004-637X/746/1/94.

Genel, S., Bryan, G.L., Springel, V. et al. A Quantification of the Butterfly Effect in Cosmological Simulations and Implications for Galaxy Scaling Relations. *ApJ*, 871(1):21, January 2019. doi: 10.3847/1538-4357/aaf4bb.

Georgakakis, A., Nandra, K., Laird, E.S. et al. A new method for determining the sensitivity of X-ray imaging observations and the X-ray number counts. *MNRAS*, 388(3):1205–1213, August 2008. doi: 10.1111/j.1365-2966.2008.13423.x.

Georgakakis, A., Comparat, J., Merloni, A. et al. Exploring the halo occupation of AGN using dark-matter cosmological simulations. *MNRAS*, 487(1):275–295, July 2019. doi: 10.1093/mnras/sty3454.

George, M.R., Leauthaud, A., Bundy, K. et al. Galaxies in X-Ray Groups. II. A Weak Lensing Study of Halo Centering. *ApJ*, 757(1):2, September 2012. doi: 10.1088/0004-637X/757/1/2.

Ghirardini, V., Eckert, D., Ettori, S. et al. Universal thermodynamic properties of the intracluster medium over two decades in radius in the X-COP sample. *A&A*, 621:A41, January 2019. doi: 10.1051/0004-6361/201833325.

Ghirardini, V., Bahar, E., Bulbul, E. et al. The eROSITA Final Equatorial-Depth Survey (eFEDS): Characterization of Morphological Properties of Galaxy Groups and Clusters. *arXiv e-prints*, art. arXiv:2106.15086, June 2021a.

Ghirardini, V., Bulbul, E., Hoang, D.N. et al. Discovery of a supercluster in the eROSITA Final Equatorial Depth Survey: X-ray properties, radio halo, and double relics. *A&A*, 647:A4, March 2021b. doi: 10.1051/0004-6361/202039554.

Giacconi, R., Murray, S., Gursky, H. et al. The Uhuru catalog of X-ray sources. *ApJ*, 178: 281–308, December 1972. doi: 10.1086/151790.

Giacconi, R., Branduardi, G., Briel, U. et al. The Einstein (HEAO 2) X-ray Observatory. *ApJ*, 230:540–550, June 1979. doi: 10.1086/157110.

Giacconi, R., Gursky, H., Paolini, F.R. et al. Evidence for x Rays From Sources Outside the Solar System. *Phys. Rev. Lett.*, 9(11):439–443, December 1962. doi: 10.1103/PhysRevLett.9.439.

Gianfagna, G., De Petris, M., Yepes, G. et al. Exploring the hydrostatic mass bias in MUSIC clusters: application to the NIKA2 mock sample. *MNRAS*, 502(4):5115–5133, April 2021. doi: 10.1093/mnras/stab308.

Gilli, R., Comastri, A. and Hasinger, G. The synthesis of the cosmic X-ray background in the Chandra and XMM-Newton era. *A&A*, 463(1):79–96, February 2007. doi: 10.1051/0004-6361:20066334.

Giocoli, C., Tormen, G. and Sheth, R.K. Formation times, mass growth histories and concentrations of dark matter haloes. *MNRAS*, 422(1):185–198, May 2012. doi: 10.1111/j.1365-2966.2012.20594.x.

Giocoli, C., Marulli, F., Moscardini, L. et al. AMICO galaxy clusters in KiDS-DR3. Cosmological constraints from large-scale stacked weak lensing profiles. *A&A*, 653:A19, September 2021. doi: 10.1051/0004-6361/202140795.

Gitti, M., Brighenti, F. and McNamara, B.R. Evidence for AGN Feedback in Galaxy Clusters and Groups. *Advances in Astronomy*, 2012:950641, January 2012. doi: 10.1155/2012/950641.

Giulia Campitiello, M., Ettori, S., Lovisari, L. et al. CHEX-MATE: Morphological analysis of the sample. *arXiv e-prints*, art. arXiv:2205.11326, May 2022.

Gladders, M.D. and Yee, H.K.C. The Red-Sequence Cluster Survey. I. The Survey and Cluster Catalogs for Patches RCS 0926+37 and RCS 1327+29. *ApJS*, 157(1):1–29, March 2005. doi: 10.1086/427327.

Gonzalez, A.H., Sivanandam, S., Zabludoff, A.I. et al. Galaxy Cluster Baryon Fractions Revisited. *ApJ*, 778(1):14, November 2013. doi: 10.1088/0004-637X/778/1/14.

Gozaliasl, G., Finoguenov, A., Tanaka, M. et al. Chandra centres for COSMOS X-ray galaxy groups: differences in stellar properties between central dominant and offset brightest group galaxies. *MNRAS*, 483(3):3545–3565, March 2019. doi: 10.1093/mnras/sty3203.

Green, T.S., Edge, A.C., Ebeling, H. et al. Hiding in plain sight – recovering clusters of galaxies with the strongest AGN in their cores. *Monthly Notices of the Royal Astronomical Society*, 465(4):4872–4885, 11 2016. ISSN 0035-8711. doi: 10.1093/mnras/stw3059. URL https://doi.org/10.1093/mnras/stw3059.

Gunn, J.E. and Gott, J. Richard, I. On the Infall of Matter Into Clusters of Galaxies and Some Effects on Their Evolution. *ApJ*, 176:1, August 1972. doi: 10.1086/151605.

Gupta, N., Saro, A., Mohr, J.J. et al. SZE observables, pressure profiles and centre offsets in Magneticum simulation galaxy clusters. *MNRAS*, 469(3):3069–3087, August 2017. doi: 10.1093/mnras/stx715.

Heitmann, K., Frontiere, N., Sewell, C. et al. The Q Continuum Simulation: Harnessing the Power of GPU Accelerated Supercomputers. *ApJS*, 219:34, August 2015. doi: 10.1088/0067-0049/219/2/34.

Henson, M.A., Barnes, D.J., Kay, S.T. et al. The impact of baryons on massive galaxy clusters: halo structure and cluster mass estimates. *Monthly Notices of the Royal Astronomical Society*, 465(3):3361–3378, 11 2016. ISSN 0035-8711. doi: 10.1093/mnras/stw2899. URL https://doi.org/10.1093/mnras/stw2899.

Henson, M.A., Barnes, D.J., Kay, S.T. et al. The impact of baryons on massive galaxy clusters: halo structure and cluster mass estimates. *MNRAS*, 465(3):3361–3378, March 2017. doi: 10.1093/mnras/stw2899.

Herschel, W. Catalogue of 500 New Nebulae, Nebulous Stars, Planetary Nebulae, and Clusters of Stars; With Remarks on the Construction of the Heavens. *Philosophical Transactions of the Royal Society of London Series I*, 92:477–528, January 1802.

Hikage, C., Mandelbaum, R., Leauthaud, A. et al. Testing redMaPPer centring probabilities using galaxy clustering and galaxy-galaxy lensing. *MNRAS*, 480(2):2689–2697, October 2018. doi: 10.1093/mnras/sty2013.

Hikage, C., Oguri, M., Hamana, T. et al. Cosmology from cosmic shear power spectra with Subaru Hyper Suprime-Cam first-year data. *PASJ*, 71(2):43, April 2019. doi: 10.1093/pasj/psz010.

Hildebrandt, H., Viola, M., Heymans, C. et al. KiDS-450: cosmological parameter constraints from tomographic weak gravitational lensing. *MNRAS*, 465(2):1454–1498, February 2017. doi: 10.1093/mnras/stw2805.

Hilton, M., Sifón, C., Naess, S. et al. The Atacama Cosmology Telescope: A Catalog of >4000 Sunyaev-Zel'dovich Galaxy Clusters. *ApJS*, 253(1):3, March 2021. doi: 10.3847/1538-4365/abd023.

Hirschmann, M., Dolag, K., Saro, A. et al. Cosmological simulations of black hole growth: AGN luminosities and downsizing. *MNRAS*, 442(3):2304–2324, August 2014. doi: 10.1093/mnras/stu1023.

Hitomi Collaboration, Aharonian, F., Akamatsu, H. et al. Atmospheric gas dynamics in the Perseus cluster observed with Hitomi. *PASJ*, 70(2):9, March 2018. doi: 10.1093/pasj/psx138.

Hoekstra, H., Bartelmann, M., Dahle, H. et al. Masses of Galaxy Clusters from Gravitational Lensing. *Space Sci. Rev.*, 177(1-4):75–118, August 2013. doi: 10.1007/s11214-013-9978-5.

Hogg, D.W. Distance measures in cosmology. *arXiv e-prints*, art. astro-ph/9905116, May 1999.

Hogg, D.W., Baldry, I.K., Blanton, M.R. et al. The K correction. *arXiv e-prints*, art. astro-ph/0210394, October 2002.

Hollowood, D.L., Jeltema, T., Chen, X. et al. Chandra Follow-up of the SDSS DR8 Redmapper Catalog Using the MATCha Pipeline. *ApJS*, 244(2):22, October 2019. doi: 10.3847/1538-4365/ab3d27.

Hoshino, H., Leauthaud, A., Lackner, C. et al. Luminous red galaxies in clusters: central occupation, spatial distributions and miscentring. *MNRAS*, 452(1):998–1013, September 2015. doi: 10.1093/mnras/stv1271.

Hubble, E. A Relation between Distance and Radial Velocity among Extra-Galactic Nebulae. *Proceedings of the National Academy of Science*, 15(3):168–173, March 1929. doi: 10.1073/pnas.15.3.168.

Huchra, J.P., Macri, L.M., Masters, K.L. et al. The 2MASS Redshift Survey—Description and Data Release. *ApJS*, 199(2):26, April 2012. doi: 10.1088/0067-0049/199/2/26.

Hudson, D.S., Mittal, R., Reiprich, T.H. et al. What is a cool-core cluster? a detailed analysis of the cores of the X-ray flux-limited HIFLUGCS cluster sample. *A&A*, 513: A37, Apr 2010. doi: 10.1051/0004-6361/200912377.

Hurier, G. and Angulo, R.E. Measuring the hydrostatic mass bias in galaxy clusters by combining Sunyaev-Zel'dovich and CMB lensing data. *A&A*, 610:L4, February 2018. doi: 10.1051/0004-6361/201731999.

Hwang, H.S. and Lee, M.G. Searching for rotating galaxy clusters in SDSS and 2dfgrs. *The Astrophysical Journal*, 662(1):236–249, jun 2007. doi: 10.1086/514328. URL https://doi.org/10.1086%2F514328.

Ider Chitham, J., Comparat, J., Finoguenov, A. et al. Cosmological constraints from CODEX galaxy clusters spectroscopically confirmed by SDSS-IV/SPIDERS DR16. *MNRAS*, 499(4):4768–4784, October 2020. doi: 10.1093/mnras/staa3044.

Ishak, M. Testing general relativity in cosmology. *Living Reviews in Relativity*, 22(1):1, December 2019. doi: 10.1007/s41114-018-0017-4.

Ishiyama, T., Enoki, M., Kobayashi, M.A.R. et al. The ν^2GC simulations: Quantifying the dark side of the universe in the Planck cosmology. *PASJ*, 67:61, August 2015. doi: 10.1093/pasj/psv021.

Ishiyama, T., Prada, F., Klypin, A.A. et al. The Uchuu simulations: Data Release 1 and dark matter halo concentrations. *MNRAS*, 506(3):4210–4231, September 2021. doi: 10.1093/mnras/stab1755.

Jansen, F., Lumb, D., Altieri, B. et al. XMM-Newton observatory. I. The spacecraft and operations. *A&A*, 365:L1–L6, January 2001. doi: 10.1051/0004-6361:20000036.

Jauzac, M., Jullo, E., Eckert, D. et al. Hubble Frontier Fields: the geometry and dynamics of the massive galaxy cluster merger MACSJ0416.1-2403. *MNRAS*, 446(4):4132–4147, February 2015. doi: 10.1093/mnras/stu2425.

Jenkins, A., Frenk, C.S., White, S.D.M. et al. The mass function of dark matter haloes. *MNRAS*, 321:372–384, February 2001. doi: 10.1046/j.1365-8711.2001.04029.x.

Jimeno, P., Broadhurst, T., Lazkoz, R. et al. Precise clustering and density evolution of redMaPPer galaxy clusters versus MXXL simulation. *MNRAS*, 466(3):2658–2674, April 2017. doi: 10.1093/mnras/stw3253.

Joudaki, S., Blake, C., Johnson, A. et al. KiDS-450 + 2dFLenS: Cosmological parameter constraints from weak gravitational lensing tomography and overlapping redshift-space galaxy clustering. *MNRAS*, 474(4):4894–4924, March 2018. doi: 10.1093/mnras/stx2820.

Käfer, F., Finoguenov, A., Eckert, D. et al. Toward a characterization of X-ray galaxy clusters for cosmology. *A&A*, 628:A43, Aug 2019. doi: 10.1051/0004-6361/201935124.

Käfer, F., Finoguenov, A., Eckert, D. et al. Toward the low-scatter selection of X-ray clusters. Galaxy cluster detection with eROSITA through cluster outskirts. *A&A*, 634: A8, February 2020. doi: 10.1051/0004-6361/201936131.

Kaiser, N. Evolution and clustering of rich clusters. *MNRAS*, 222:323–345, September 1986. doi: 10.1093/mnras/222.2.323.

Kaiser, N. Clustering in real space and in redshift space. *MNRAS*, 227:1–21, July 1987. doi: 10.1093/mnras/227.1.1.

Kapteyn, J.C. First Attempt at a Theory of the Arrangement and Motion of the Sidereal System. *ApJ*, 55:302, May 1922. doi: 10.1086/142670.

Kauffmann, G. and Haehnelt, M. A unified model for the evolution of galaxies and quasars. *MNRAS*, 311(3):576–588, January 2000. doi: 10.1046/j.1365-8711.2000.03077.x.

Kilbinger, M. Cosmology with cosmic shear observations: a review. *Reports on Progress in Physics*, 78(8):086901, Jul 2015. ISSN 1361-6633. doi: 10.1088/0034-4885/78/8/086901. URL http://dx.doi.org/10.1088/0034-4885/78/8/086901.

Klypin, A., Yepes, G., Gottlöber, S. et al. MultiDark simulations: the story of dark matter halo concentrations and density profiles. *MNRAS*, 457:4340–4359, April 2016. doi: 10.1093/mnras/stw248.

Knebe, A., Knollmann, S.R., Muldrew, S.I. et al. Haloes gone MAD: The Halo-Finder Comparison Project. *MNRAS*, 415:2293–2318, August 2011. doi: 10.1111/j.1365-2966.2011.18858.x.

Knebe, A., Pearce, F.R., Lux, H. et al. Structure finding in cosmological simulations: the state of affairs. *MNRAS*, 435(2):1618–1658, October 2013. doi: 10.1093/mnras/stt1403.

Koens, L.A., Maughan, B.J., Jones, L.R. et al. The WARPS Survey - VIII. Evolution of the galaxy cluster X-ray Luminosity Function. *MNRAS*, 435(4):3231–3242, November 2013. doi: 10.1093/mnras/stt1519.

Komatsu, E., Afshordi, N., Bartolo, N. et al. Non-Gaussianity as a Probe of the Physics of the Primordial Universe and the Astrophysics of the Low Redshift Universe. In *astro2010: The Astronomy and Astrophysics Decadal Survey*, volume 2010, page 158, January 2009.

Komatsu, E., Smith, K.M., Dunkley, J. et al. Seven-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation. *ApJS*, 192(2):18, February 2011. doi: 10.1088/0067-0049/192/2/18.

Kong, H., Burleigh, K.J., Ross, A. et al. Removing imaging systematics from galaxy clustering measurements with Obiwan: application to the SDSS-IV extended Baryon Oscillation Spectroscopic Survey emission-line galaxy sample. *MNRAS*, 499(3):3943–3960, December 2020. doi: 10.1093/mnras/staa2742.

König, O., Wilms, J., Arcodia, R. et al. X-ray detection of a nova in the fireball phase. *Nature*, 605(7909):248–250, May 2022. doi: 10.1038/s41586-022-04635-y.

Koulouridis, E., Plionis, M., Melnyk, O. et al. X-ray AGN in the XMM-LSS galaxy clusters: no evidence of AGN suppression. *A&A*, 567:A83, July 2014. doi: 10.1051/0004-6361/201423601.

Koutoulidis, L., Plionis, M., Georgantopoulos, I. et al. Clustering, bias and the accretion mode of X-ray-selected AGN. *MNRAS*, 428(2):1382–1394, January 2013. doi: 10.1093/mnras/sts119.

Kravtsov, A.V., Klypin, A.A. and Khokhlov, A.M. Adaptive Refinement Tree: A New High-Resolution N-Body Code for Cosmological Simulations. *ApJS*, 111:73–94, July 1997. doi: 10.1086/313015.

Kravtsov, A.V. and Borgani, S. Formation of Galaxy Clusters. *ARA&A*, 50:353–409, September 2012. doi: 10.1146/annurev-astro-081811-125502.

Kuijken, K., Heymans, C., Dvornik, A. et al. The fourth data release of the Kilo-Degree Survey: ugri imaging and nine-band optical-IR photometry over 1000 square degrees. *A&A*, 625:A2, May 2019. doi: 10.1051/0004-6361/201834918.

Laganá, T.F., Martinet, N., Durret, F. et al. A comprehensive picture of baryons in groups and clusters of galaxies. *A&A*, 555:A66, July 2013. doi: 10.1051/0004-6361/201220423.

Lang, M., Holley-Bockelmann, K. and Sinha, M. Voronoi Tessellation and Non-parametric Halo Concentration. *ApJ*, 811(2):152, October 2015. doi: 10.1088/0004-637X/811/2/152.

Lange, J.U., van den Bosch, F.C., Hearin, A. et al. Brightest galaxies as halo centre tracers in SDSS DR7. *MNRAS*, 473(2):2830–2851, January 2018. doi: 10.1093/mnras/stx2434.

Laureijs, R., Amiaux, J., Arduini, S. et al. Euclid Definition Study Report. *arXiv e-prints*, art. arXiv:1110.3193, Oct 2011.

Laurent, P., Le Goff, J.M., Burtin, E. et al. A 14 h^{-3} Gpc^3 study of cosmic homogeneity using BOSS DR12 quasar sample. *J. Cosmology Astropart. Phys.*, 2016(11):060, November 2016. doi: 10.1088/1475-7516/2016/11/060.

Lazarian, A., Vlahos, L., Kowal, G. et al. Turbulence, Magnetic Reconnection in Turbulent Fluids and Energetic Particle Acceleration. *Space Sci. Rev.*, 173(1-4):557–622, November 2012. doi: 10.1007/s11214-012-9936-7.

Le Brun, A.M.C., McCarthy, I.G., Schaye, J. et al. Towards a realistic population of simulated galaxy groups and clusters. *MNRAS*, 441(2):1270–1290, June 2014. doi: 10.1093/mnras/stu608.

Lemaître, A.G. A Homogeneous Universe of Constant Mass and Increasing Radius accounting for the Radial Velocity of Extra-galactic Nebulæ. *Monthly Notices of the Royal Astronomical Society*, 91(5):483–490, 03 1931. ISSN 0035-8711. doi: 10.1093/mnras/91.5.483. URL https://doi.org/10.1093/mnras/91.5.483.

Lesci, G.F., Marulli, F., Moscardini, L. et al. AMICO galaxy clusters in KiDS-DR3: Cosmological constraints from counts and stacked weak lensing. *A&A*, 659:A88, March 2022a. doi: 10.1051/0004-6361/202040194.

Lesci, G.F., Nanni, L., Marulli, F. et al. AMICO galaxy clusters in KiDS-DR3: Constraints on cosmological parameters and on the normalisation of the mass-richness relation from clustering. *arXiv e-prints*, art. arXiv:2203.07398, March 2022b.

Li, Y., Bryan, G.L., Ruszkowski, M. et al. Cooling, AGN Feedback, and Star Formation in Simulated Cool-core Galaxy Clusters. *ApJ*, 811(2):73, October 2015. doi: 10.1088/0004-637X/811/2/73.

Liddle, A.R. and Lyth, D.H. *Cosmological Inflation and Large-Scale Structure*. Cambridge University Press, 2000. doi: 10.1017/CBO9781139175180.

Lindholm, V., Finoguenov, A., Comparat, J. et al. Clustering of CODEX clusters. *A&A*, 646:A8, February 2021. doi: 10.1051/0004-6361/202038807.

Liu, A., Bulbul, E., Ghirardini, V. et al. The eROSITA Final Equatorial-Depth Survey (eFEDS). Catalog of galaxy clusters and groups. *A&A*, 661:A2, May 2022. doi: 10.1051/0004-6361/202141120.

Liu, T., Tozzi, P., Tundo, E. et al. EXSdetect: an end-to-end software for extended source detection in X-ray images: application to Swift-XRT data. *A&A*, 549:A143, January 2013. doi: 10.1051/0004-6361/201219866.

Liu, T., Merloni, A., Comparat, J. et al. Establishing the X-ray Source Detection Strategy for eROSITA with Simulations. *arXiv e-prints*, art. arXiv:2106.14528, June 2021.

Longair, M.S. *Galaxy Formation*. Astronomy and Astrophysics Library. Springer, Heidelberg, Germany, 2008. ISBN 978-3-540-73477-2, 978-3-540-73478-9. doi: 10.1007/978-3-540-73478-9.

Lovisari, L., Reiprich, T.H. and Schellenberger, G. Scaling properties of a complete X-ray selected galaxy group sample. *A&A*, 573:A118, January 2015. doi: 10.1051/0004-6361/201423954.

Lovisari, L., Forman, W.R., Jones, C. et al. X-Ray Morphological Analysis of the Planck ESZ Clusters. *ApJ*, 846(1):51, September 2017. doi: 10.3847/1538-4357/aa855f.

Lovisari, L., Schellenberger, G., Sereno, M. et al. X-Ray Scaling Relations for a Representative Sample of Planck-selected Clusters Observed with XMM-Newton. *ApJ*, 892(2):102, April 2020. doi: 10.3847/1538-4357/ab7997.

LSST Science Collaboration, Abell, P.A., Allison, J. et al. LSST Science Book, Version 2.0. *arXiv e-prints*, art. arXiv:0912.0201, December 2009.

Ludlow, A.D., Navarro, J.F., Li, M. et al. The dynamical state and mass-concentration relation of galaxy clusters. *MNRAS*, 427(2):1322–1328, December 2012. doi: 10.1111/j.1365-2966.2012.21892.x.

Ludlow, A.D., Navarro, J.F., Boylan-Kolchin, M. et al. The mass profile and accretion history of cold dark matter haloes. *MNRAS*, 432(2):1103–1113, June 2013. doi: 10.1093/mnras/stt526.

Ludlow, A.D., Navarro, J.F., Angulo, R.E. et al. The mass-concentration-redshift relation of cold dark matter haloes. *MNRAS*, 441(1):378–388, June 2014. doi: 10.1093/mnras/stu483.

Ludlow, A.D., Bose, S., Angulo, R.E. et al. The mass-concentration-redshift relation of cold and warm dark matter haloes. *MNRAS*, 460(2):1214–1232, August 2016. doi: 10.1093/mnras/stw1046.

Macciò, A.V., Dutton, A.A. and van den Bosch, F.C. Concentration, spin and shape of dark matter haloes as a function of the cosmological model: WMAP1, WMAP3 and WMAP5 results. *MNRAS*, 391(4):1940–1954, Dec 2008. doi: 10.1111/j.1365-2966.2008.14029.x.

Malyali, A., Rau, A., Merloni, A. et al. AT 2019avd: a novel addition to the diverse population of nuclear transients. *A&A*, 647:A9, March 2021. doi: 10.1051/0004-6361/202039681.

Mamon, G.A. and Łokas, E.L. Dark matter in elliptical galaxies - II. Estimating the mass within the virial radius. *MNRAS*, 363(3):705–722, November 2005. doi: 10.1111/j.1365-2966.2005.09400.x.

Mann, A.W. and Ebeling, H. X-ray-optical classification of cluster mergers and the evolution of the cluster merger fraction. *MNRAS*, 420(3):2120–2138, March 2012. doi: 10.1111/j.1365-2966.2011.20170.x.

Manolopoulou, M. and Plionis, M. Galaxy cluster's rotation. *MNRAS*, 465(3):2616–2633, March 2017. doi: 10.1093/mnras/stw2870.

Mantz, A.B., Allen, S.W., Morris, R.G. et al. Cosmology and astrophysics from relaxed galaxy clusters - I. Sample selection. *MNRAS*, 449(1):199–219, May 2015a. doi: 10.1093/mnras/stv219.

Mantz, A.B., von der Linden, A., Allen, S.W. et al. Weighing the giants - IV. Cosmology and neutrino mass. *MNRAS*, 446(3):2205–2225, January 2015b. doi: 10.1093/mnras/stu2096.

Mantz, A.B., Allen, S.W., Morris, R.G. et al. Weighing the giants- V. Galaxy cluster scaling relations. *MNRAS*, 463(4):3582–3603, December 2016. doi: 10.1093/mnras/stw2250.

Mantz, A.B., Morris, R.G., Allen, S.W. et al. Cosmological constraints from gas mass fractions of massive, relaxed galaxy clusters. *MNRAS*, 510(1):131–145, February 2022. doi: 10.1093/mnras/stab3390.

Marinacci, F., Vogelsberger, M., Pakmor, R. et al. First results from the IllustrisTNG simulations: radio haloes and magnetic fields. *MNRAS*, 480(4):5113–5139, November 2018. doi: 10.1093/mnras/sty2206.

Markevitch, M., Gonzalez, A.H., David, L. et al. A Textbook Example of a Bow Shock in the Merging Galaxy Cluster 1E 0657-56. *ApJ*, 567(1):L27–L31, March 2002. doi: 10.1086/339619.

Markevitch, M. and Vikhlinin, A. Shocks and cold fronts in galaxy clusters. *Phys. Rep.*, 443(1):1–53, May 2007. doi: 10.1016/j.physrep.2007.01.001.

Markevitch, M., Sarazin, C.L. and Vikhlinin, A. Physics of the Merging Clusters Cygnus A, A3667, and A2065. *ApJ*, 521(2):526–530, August 1999. doi: 10.1086/307598.

Martini, P., Miller, E.D., Brodwin, M. et al. The Cluster and Field Galaxy Active Galactic Nucleus Fraction at z = 1-1.5: Evidence for a Reversal of the Local Anticorrelation between Environment and AGN Fraction. *ApJ*, 768(1):1, May 2013. doi: 10.1088/0004-637X/768/1/1.

Marulli, F., Veropalumbo, A., Sereno, M. et al. The XXL Survey. XVI. The clustering of X-ray selected galaxy clusters at z 0.3. *A&A*, 620:A1, November 2018. doi: 10.1051/0004-6361/201833238.

Marulli, F., Veropalumbo, A., Moscardini, L. et al. Redshift-space distortions of galaxies, clusters, and AGN. Testing how the accuracy of growth rate measurements depends on scales and sample selections. *A&A*, 599:A106, March 2017. doi: 10.1051/0004-6361/201526885.

Marulli, F., Veropalumbo, A., García-Farieta, J.E. et al. C^3 Cluster Clustering Cosmology I. New Constraints on the Cosmic Growth Rate at z 0.3 from Redshift-space Clustering Anisotropies. *ApJ*, 920(1):13, October 2021. doi: 10.3847/1538-4357/ac0e8c.

Mathews, W.G. and Guo, F. Cosmic Ray Diffusion Fronts in the Virgo Cluster. *ApJ*, 736 (1):6, July 2011. doi: 10.1088/0004-637X/736/1/6.

Maturi, M., Meneghetti, M., Bartelmann, M. et al. An optimal filter for the detection of galaxy clusters through weak lensing. *A&A*, 442(3):851–860, November 2005. doi: 10.1051/0004-6361:20042600.

Maughan, B.J. The L_X-Y_X Relation: Using Galaxy Cluster X-Ray Luminosity as a Robust, Low-Scatter Mass Proxy. *ApJ*, 668(2):772–780, October 2007. doi: 10.1086/520831.

Mayer, L. and Bonoli, S. The route to massive black hole formation via merger-driven direct collapse: a review. *Reports on Progress in Physics*, 82(1):016901, January 2019. doi: 10.1088/1361-6633/aad6a5.

McClintock, T., Rozo, E., Becker, M.R. et al. The Aemulus Project. II. Emulating the Halo Mass Function. *ApJ*, 872(1):53, February 2019. doi: 10.3847/1538-4357/aaf568.

McDonald, M., Bayliss, M., Benson, B.A. et al. A massive, cooling-flow-induced starburst in the core of a luminous cluster of galaxies. *Nature*, 488(7411):349–352, August 2012. doi: 10.1038/nature11379.

McDonald, M., Allen, S.W., Bayliss, M. et al. The Remarkable Similarity of Massive Galaxy Clusters from z $\sim$ 0 to z $\sim$ 1.9. *ApJ*, 843(1):28, July 2017. doi: 10.3847/1538-4357/aa7740.

McNamara, B.R. and Nulsen, P.E.J. Mechanical feedback from active galactic nuclei in galaxies, groups and clusters. *New Journal of Physics*, 14(5):055023, May 2012. doi: 10.1088/1367-2630/14/5/055023.

Menanteau, F., Hughes, J.P., Sifón, C. et al. The Atacama Cosmology Telescope: ACT-CL J0102-4915 "El Gordo," a Massive Merging Cluster at Redshift 0.87. *ApJ*, 748(1):7, March 2012. doi: 10.1088/0004-637X/748/1/7.

Meneghetti, M. and Rasia, E. Reconciling extremely different concentration-mass relations. *arXiv e-prints*, art. arXiv:1303.6158, March 2013.

Meneghetti, M., Rasia, E., Vega, J. et al. The MUSIC of CLASH: Predictions on the Concentration-Mass Relation. *ApJ*, 797(1):34, December 2014. doi: 10.1088/0004-637X/797/1/34.

Merloni, A., Predehl, P., Becker, W. et al. eROSITA Science Book: Mapping the Structure of the Energetic Universe. *arXiv e-prints*, art. arXiv:1209.3114, Sep 2012.

Merten, J., Coe, D., Dupke, R. et al. Creation of cosmic structure in the complex galaxy cluster merger Abell 2744. *MNRAS*, 417(1):333–347, October 2011. doi: 10.1111/j.1365-2966.2011.19266.x.

Messier, C. Catalogue des Nébuleuses et des Amas d'Étoiles (Catalog of Nebulae and Star Clusters). Connoissance des Temps ou des Mouvements Célestes, for 1784, p. 227-267, January 1781.

Meszaros, P. The behaviour of point masses in an expanding cosmological substratum. *A&A*, 37(2):225–228, December 1974.

Migkas, K., Schellenberger, G., Reiprich, T.H. et al. Probing cosmic isotropy with a new X-ray galaxy cluster sample through the L_X-T scaling relation. *A&A*, 636:A15, April 2020. doi: 10.1051/0004-6361/201936602.

Migkas, K., Pacaud, F., Schellenberger, G. et al. Cosmological implications of the anisotropy of ten galaxy cluster scaling relations. *A&A*, 649:A151, May 2021. doi: 10.1051/0004-6361/202140296.

Mitsuda, K., Bautz, M., Inoue, H. et al. The X-Ray Observatory Suzaku. *PASJ*, 59:S1–S7, January 2007. doi: 10.1093/pasj/59.sp1.S1.

Miyazaki, S., Oguri, M., Hamana, T. et al. A large sample of shear-selected clusters from the Hyper Suprime-Cam Subaru Strategic Program S16A Wide field mass maps. *PASJ*, 70:S27, January 2018a. doi: 10.1093/pasj/psx120.

Miyazaki, S., Oguri, M., Hamana, T. et al. A large sample of shear-selected clusters from the Hyper Suprime-Cam Subaru Strategic Program S16A Wide field mass maps. *PASJ*, 70:S27, January 2018b. doi: 10.1093/pasj/psx120.

Mo, H.J. and White, S.D.M. The abundance and clustering of dark haloes in the standard ΛCDM cosmogony. *MNRAS*, 336(1):112–118, October 2002. doi: 10.1046/j.1365-8711.2002.05723.x.

Mo, H., van den Bosch, F.C. and White, S. *Galaxy Formation and Evolution*. 2010.

Molendi, S. and Pizzolato, F. Is the Gas in Cooling Flows Multiphase? *ApJ*, 560(1): 194–200, October 2001. doi: 10.1086/322387.

Molnar, S.M., Hearn, N.C. and Stadel, J.G. Merging Galaxy Clusters: Offset between the Sunyaev-Zel'dovich Effect and X-Ray Peaks. *ApJ*, 748(1):45, March 2012. doi: 10.1088/0004-637X/748/1/45.

Monteiro-Oliveira, R., Cypriano, E.S., Machado, R.E.G. et al. The merger history of the complex cluster Abell 1758: a combined weak lensing and spectroscopic view. *MNRAS*, 466(3):2614–2632, April 2017. doi: 10.1093/mnras/stw3238.

Moresco, M., Veropalumbo, A., Marulli, F. et al. C^3: Cluster Clustering Cosmology. II. First Detection of the Baryon Acoustic Oscillations Peak in the Three-point Correlation Function of Galaxy Clusters. *ApJ*, 919(2):144, October 2021. doi: 10.3847/1538-4357/ac10c9.

Moscardini, L., Matarrese, S. and Mo, H.J. Constraining cosmological parameters with the clustering properties of galaxy clusters in optical and X-ray bands. *MNRAS*, 327(2): 422–434, October 2001. doi: 10.1046/j.1365-8711.2001.04728.x.

Moster, B.P., Naab, T. and White, S.D.M. Galactic star formation and accretion histories from matching galaxies to dark matter haloes. *MNRAS*, 428(4):3121–3138, February 2013. doi: 10.1093/mnras/sts261.

Mullis, C.R., Vikhlinin, A., Henry, J.P. et al. Evolution of the Cluster X-Ray Luminosity Function. *ApJ*, 607(1):175–189, May 2004. doi: 10.1086/383234.

Murray, S.G., Power, C. and Robotham, A.S.G. HMFcalc: An online tool for calculating dark matter halo mass functions. *Astronomy and Computing*, 3:23–34, November 2013. doi: 10.1016/j.ascom.2013.11.001.

Naiman, J.P., Pillepich, A., Springel, V. et al. First results from the IllustrisTNG simulations: a tale of two elements - chemical evolution of magnesium and europium. *MNRAS*, 477(1):1206–1224, June 2018. doi: 10.1093/mnras/sty618.

Navarro, J.F., Frenk, C.S. and White, S.D.M. The Structure of Cold Dark Matter Halos. *ApJ*, 462:563, May 1996. doi: 10.1086/177173.

Nelson, D., Pillepich, A., Springel, V. et al. First results from the IllustrisTNG simulations: the galaxy colour bimodality. *MNRAS*, 475(1):624–647, March 2018. doi: 10.1093/mnras/stx3040.

Nelson, D., Springel, V., Pillepich, A. et al. The IllustrisTNG simulations: public data release. *Computational Astrophysics and Cosmology*, 6(1):2, May 2019. doi: 10.1186/s40668-019-0028-x.

Neto, A.F., Gao, L., Bett, P. et al. The statistics of Λ CDM halo concentrations. *MNRAS*, 381(4):1450–1462, Nov 2007. doi: 10.1111/j.1365-2966.2007.12381.x.

Nishimichi, T., Takada, M., Takahashi, R. et al. Dark Quest. I. Fast and Accurate Emulation of Halo Clustering Statistics and Its Application to Galaxy Clustering. *ApJ*, 884 (1):29, October 2019. doi: 10.3847/1538-4357/ab3719.

Noordeh, E., Canning, R.E.A., King, A. et al. The environmental dependence of X-ray AGN activity at z $\sim$ 0.4. *MNRAS*, 498(3):4095–4108, November 2020. doi: 10.1093/mnras/staa2682.

Nurgaliev, D., McDonald, M., Benson, B.A. et al. Testing for X-Ray-SZ Differences and Redshift Evolution in the X-Ray Morphology of Galaxy Clusters. *ApJ*, 841(1):5, May 2017. doi: 10.3847/1538-4357/aa6db4.

Ogrean, G.A., van Weeren, R.J., Jones, C. et al. Frontier Fields Clusters: Chandra and JVLA View of the Pre-merging Cluster MACS J0416.1-2403. *ApJ*, 812(2):153, October 2015. doi: 10.1088/0004-637X/812/2/153.

Oguri, M. A cluster finding algorithm based on the multiband identification of red sequence galaxies. *MNRAS*, 444(1):147–161, October 2014. doi: 10.1093/mnras/stu1446.

Oguri, M., Lin, Y.T., Lin, S.C. et al. An optically-selected cluster catalog at redshift 0.1 $<$ z $<$ 1.1 from the Hyper Suprime-Cam Subaru Strategic Program S16A data. *PASJ*, 70:S20, January 2018. doi: 10.1093/pasj/psx042.

Oldham, L.J. and Auger, M.W. Galaxy structure from multiple tracers - II. M87 from parsec to megaparsec scales. *MNRAS*, 457(1):421–439, March 2016. doi: 10.1093/mnras/stv2982.

Ondaro-Mallea, L., Angulo, R.E., Zennaro, M. et al. Non-universality of the mass function: dependence on the growth rate and power spectrum shape. *MNRAS*, 509(4):6077–6090, February 2022. doi: 10.1093/mnras/stab3337.

Ota, N., Nguyen-Dang, N.T., Mitsuishi, I. et al. The eROSITA Final Equatorial Survey (eFEDS): X-ray properties of Subaru optically-selected clusters. *arXiv e-prints*, art. arXiv:2206.09536, June 2022.

Ota, N., Mitsuishi, I., Babazaki, Y. et al. X-ray properties of high-richness CAMIRA clusters in the Hyper Suprime-Cam Subaru Strategic Program field. *PASJ*, 72(1):1, January 2020. doi: 10.1093/pasj/psz118.

Pacaud, F., Pierre, M., Refregier, A. et al. The XMM Large-Scale Structure survey: the X-ray pipeline and survey selection function. *MNRAS*, 372(2):578–590, October 2006. doi: 10.1111/j.1365-2966.2006.10881.x.

Pacaud, F., Pierre, M., Melin, J.B. et al. The XXL Survey. XXV. Cosmological analysis of the C1 cluster number counts. *A&A*, 620:A10, November 2018. doi: 10.1051/0004-6361/201834022.

Padmanabhan, T. *Structure Formation in the Universe*. 1993.

Padovani, P., Alexander, D.M., Assef, R.J. et al. Active galactic nuclei: what's in a name? *A&A Rev.*, 25(1):2, August 2017. doi: 10.1007/s00159-017-0102-9.

Pasini, T., Gitti, M., Brighenti, F. et al. A First Chandra View of the Cool Core Cluster A1668: Offset Cooling and AGN Feedback Cycle. *ApJ*, 911(1):66, April 2021. doi: 10.3847/1538-4357/abe85f.

Pasini, T., Brüggen, M., Hoang, D.N. et al. The eROSITA Final Equatorial-Depth Survey (eFEDS). LOFAR view of brightest cluster galaxies and AGN feedback. *A&A*, 661:A13, May 2022. doi: 10.1051/0004-6361/202141211.

Peebles, P.J.E. Origin of the Angular Momentum of Galaxies. *ApJ*, 155:393, February 1969. doi: 10.1086/149876.

Peebles, P.J.E. Large-scale background temperature and mass fluctuations due to scale-invariant primeval perturbations. *ApJ*, 263:L1–L5, December 1982. doi: 10.1086/183911.

Peebles, P.J.E. and Yu, J.T. Primeval Adiabatic Perturbation in an Expanding Universe. *ApJ*, 162:815, December 1970. doi: 10.1086/150713.

Perlmutter, S., Aldering, G., Goldhaber, G. et al. Measurements of Ω and Λ from 42 High-Redshift Supernovae. *ApJ*, 517(2):565–586, June 1999. doi: 10.1086/307221.

Phriksee, A., Jullo, E., Limousin, M. et al. Weak lensing analysis of CODEX clusters using dark energy camera legacy survey: mass-richness relation. *MNRAS*, 491(2):1643–1655, January 2020. doi: 10.1093/mnras/stz3049.

Pierre, M., Pacaud, F., Adami, C. et al. The XXL Survey. I. Scientific motivations - XMM-Newton observing plan - Follow-up observations and simulation programme. *A&A*, 592: A1, June 2016. doi: 10.1051/0004-6361/201526766.

Pillepich, A., Porciani, C. and Reiprich, T.H. The X-ray cluster survey with eRosita: forecasts for cosmology, cluster physics and primordial non-Gaussianity. *MNRAS*, 422 (1):44–69, May 2012. doi: 10.1111/j.1365-2966.2012.20443.x.

Pillepich, A., Nelson, D., Hernquist, L. et al. First results from the IllustrisTNG simulations: the stellar mass content of groups and clusters of galaxies. *MNRAS*, 475(1): 648–675, March 2018a. doi: 10.1093/mnras/stx3112.

Pillepich, A., Reiprich, T.H., Porciani, C. et al. Forecasts on dark energy from the X-ray cluster survey with eROSITA: constraints from counts and clustering. *MNRAS*, 481(1): 613–626, Nov 2018b. doi: 10.1093/mnras/sty2240.

Pillepich, A., Springel, V., Nelson, D. et al. Simulating galaxy formation with the IllustrisTNG model. *MNRAS*, 473(3):4077–4106, January 2018c. doi: 10.1093/mnras/stx2656.

Pinto, C., Bambic, C.J., Sanders, J.S. et al. AGN feedback in the Phoenix cluster. *MNRAS*, 480(3):4113–4123, November 2018. doi: 10.1093/mnras/sty2185.

Planck Collaboration, Ade, P.A.R., Aghanim, N. et al. Planck 2013 results. XVI. Cosmological parameters. *A&A*, 571:A16, November 2014a. doi: 10.1051/0004-6361/201321591.

Planck Collaboration, Ade, P.A.R., Aghanim, N. et al. Planck 2013 results. XX. Cosmology from Sunyaev-Zeldovich cluster counts. *A&A*, 571:A20, November 2014b. doi: 10.1051/0004-6361/201321521.

Planck Collaboration, Ade, P.A.R., Aghanim, N. et al. Planck 2013 results. XVII. Gravitational lensing by large-scale structure. *A&A*, 571:A17, November 2014c. doi: 10.1051/0004-6361/201321543.

Planck Collaboration, Ade, P.A.R., Aghanim, N. et al. Planck 2015 results. XXVII. The second Planck catalogue of Sunyaev-Zeldovich sources. *A&A*, 594:A27, September 2016a. doi: 10.1051/0004-6361/201525823.

Planck Collaboration, Ade, P.A.R., Aghanim, N. et al. Planck 2015 results. XIII. Cosmological parameters. *A&A*, 594:A13, September 2016b. doi: 10.1051/0004-6361/201525830.

Planck Collaboration, Ade, P.A.R., Aghanim, N. et al. Planck 2015 results. XXIV. Cosmology from Sunyaev-Zeldovich cluster counts. *A&A*, 594:A24, September 2016c. doi: 10.1051/0004-6361/201525833.

Planck Collaboration, Ade, P.A.R., Aghanim, N. et al. Planck intermediate results. XL. The Sunyaev-Zeldovich signal from the Virgo cluster. *A&A*, 596:A101, December 2016d. doi: 10.1051/0004-6361/201527743.

Planck Collaboration, Aghanim, N., Akrami, Y. et al. Planck 2018 results. VI. Cosmological parameters. *A&A*, 641:A6, September 2020a. doi: 10.1051/0004-6361/201833910.

Planck Collaboration, Akrami, Y., Arroja, F. et al. Planck 2018 results. IX. Constraints on primordial non-Gaussianity. *A&A*, 641:A9, September 2020b. doi: 10.1051/0004-6361/201935891.

Poveda-Ruiz, C.N., Forero-Romero, J.E. and Muñoz-Cuartas, J.C. Quantifying and Controlling Biases in Estimates of Dark Matter Halo Concentration. *ApJ*, 832(2):169, December 2016. doi: 10.3847/0004-637X/832/2/169.

Prada, F., Klypin, A.A., Cuesta, A.J. et al. Halo concentrations in the standard Λ cold dark matter cosmology. *MNRAS*, 423:3018–3030, July 2012. doi: 10.1111/j.1365-2966.2012.21007.x.

Pratt, G.W., Croston, J.H., Arnaud, M. et al. Galaxy cluster X-ray luminosity scaling relations from a representative local sample (REXCESS). *A&A*, 498(2):361–378, May 2009. doi: 10.1051/0004-6361/200810994.

Pratt, G.W., Arnaud, M., Biviano, A. et al. The Galaxy Cluster Mass Scale and Its Impact on Cosmological Constraints from the Cluster Population. *Space Sci. Rev.*, 215(2):25, February 2019. doi: 10.1007/s11214-019-0591-0.

Predehl, P., Sunyaev, R.A., Becker, W. et al. Detection of large-scale X-ray bubbles in the Milky Way halo. *Nature*, 588(7837):227–231, December 2020. doi: 10.1038/s41586-020-2979-0.

Predehl, P., Andritschke, R., Arefiev, V. et al. The eROSITA X-ray telescope on SRG. *A&A*, 647:A1, March 2021. doi: 10.1051/0004-6361/202039313.

Press, W.H. and Schechter, P. Formation of Galaxies and Clusters of Galaxies by Self-Similar Gravitational Condensation. *ApJ*, 187:425–438, February 1974. doi: 10.1086/152650.

Ragagnin, A., Dolag, K., Biffi, V. et al. A web portal for hydrodynamical, cosmological simulations. *Astronomy and Computing*, 20:52–67, July 2017. doi: 10.1016/j.ascom.2017.05.001.

Ragagnin, A., Dolag, K., Moscardini, L. et al. Dependency of halo concentration on mass, redshift and fossilness in Magneticum hydrodynamic simulations. *MNRAS*, 486(3):4001–4012, July 2019. doi: 10.1093/mnras/stz1103.

Reiprich, T.H. and Böhringer, H. The Mass Function of an X-Ray Flux-limited Sample of Galaxy Clusters. *ApJ*, 567(2):716–740, March 2002. doi: 10.1086/338753.

Remus, R.S., Dolag, K., Naab, T. et al. The co-evolution of total density profiles and central dark matter fractions in simulated early-type galaxies. *MNRAS*, 464(3):3742–3756, January 2017. doi: 10.1093/mnras/stw2594.

Rephaeli, Y. Comptonization Of The Cosmic Microwave Background: The Sunyaev-Zeldovich Effect. *ARA&A*, 33:541–580, January 1995. doi: 10.1146/annurev.aa.33.090195.002545.

Riebe, K., Partl, A.M., Enke, H. et al. The MultiDark Database: Release of the Bolshoi and MultiDark cosmological simulations. *Astronomische Nachrichten*, 334(7):691–708, Aug 2013. doi: 10.1002/asna.201211900.

Riess, A.G., Filippenko, A.V., Challis, P. et al. Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant. *AJ*, 116(3):1009–1038, September 1998. doi: 10.1086/300499.

Riess, A.G., Yuan, W., Macri, L.M. et al. A Comprehensive Measurement of the Local Value of the Hubble Constant with 1 km s^{-1} Mpc^{-1} Uncertainty from the Hubble Space Telescope and the SH0ES Team. *ApJ*, 934(1):L7, July 2022. doi: 10.3847/2041-8213/ac5c5b.

Robertson, H.P. On the Foundations of Relativistic Cosmology. *Proceedings of the National Academy of Science*, 15(11):822–829, November 1929. doi: 10.1073/pnas.15.11.822.

Rodriguez-Puebla, A., Behroozi, P., Primack, J. et al. Halo and Subhalo Demographics with Planck Cosmological Parameters: Bolshoi-Planck and MultiDark-Planck Simulations. *ArXiv e-prints*, February 2016.

Rosati, P., Borgani, S. and Norman, C. The Evolution of X-ray Clusters of Galaxies. *ARA&A*, 40:539–577, January 2002. doi: 10.1146/annurev.astro.40.120401.150547.

Rossetti, M., Gastaldello, F., Ferioli, G. et al. Measuring the dynamical state of Planck SZ-selected clusters: X-ray peak - BCG offset. *MNRAS*, 457(4):4515–4524, April 2016. doi: 10.1093/mnras/stw265.

Rossetti, M., Gastaldello, F., Eckert, D. et al. The cool-core state of Planck SZ-selected clusters versus X-ray-selected samples: evidence for cool-core bias. *MNRAS*, 468(2):1917–1930, June 2017. doi: 10.1093/mnras/stx493.

Rozo, E. and Rykoff, E.S. redMaPPer II: X-Ray and SZ Performance Benchmarks for the SDSS Catalog. *ApJ*, 783(2):80, March 2014. doi: 10.1088/0004-637X/783/2/80.

Rubin, V.C., Ford, W. K., J. and Thonnard, N. Rotational properties of 21 SC galaxies with a large range of luminosities and radii, from NGC 4605 (R=4kpc) to UGC 2885 (R=122kpc). *ApJ*, 238:471–487, June 1980. doi: 10.1086/158003.

Rubin, V.C. and Ford, W. Kent, J. Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions. *ApJ*, 159:379, February 1970. doi: 10.1086/150317.

Rykoff, E.S., Rozo, E., Busha, M.T. et al. redMaPPer. I. Algorithm and SDSS DR8 Catalog. *ApJ*, 785(2):104, April 2014. doi: 10.1088/0004-637X/785/2/104.

Rykoff, E.S., Rozo, E., Hollowood, D. et al. The RedMaPPer Galaxy Cluster Catalog From DES Science Verification Data. *ApJS*, 224(1):1, May 2016. doi: 10.3847/0067-0049/224/1/1.

Salvati, L., Saro, A., Bocquet, S. et al. Combining Planck and SPT Cluster Catalogs: Cosmological Analysis and Impact on the Planck Scaling Relation Calibration. *ApJ*, 934(2):129, August 2022. doi: 10.3847/1538-4357/ac7ab4.

Salvati, L., Douspis, M. and Aghanim, N. Impact of systematics on cosmological parameters from future Galaxy Clusters surveys. *arXiv e-prints*, art. arXiv:2005.10204, May 2020.

Salvato, M., Wolf, J., Dwelly, T. et al. The eROSITA Final Equatorial-Depth Survey (eFEDS). Identification and characterization of the counterparts to point-like sources. *A&A*, 661:A3, May 2022. doi: 10.1051/0004-6361/202141631.

Sánchez, A.G., Scóccola, C.G., Ross, A.J. et al. The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: cosmological implications of the large-scale two-point correlation function. *MNRAS*, 425(1):415–437, September 2012. doi: 10.1111/j.1365-2966.2012.21502.x.

Sánchez, A.G., Ruiz, A.N., Jara, J.G. et al. Evolution mapping: a new approach to describe matter clustering in the non-linear regime. *MNRAS*, 514(4):5673–5685, August 2022. doi: 10.1093/mnras/stac1656.

Sanders, J.S., Fabian, A.C., Russell, H.R. et al. Hydrostatic Chandra X-ray analysis of SPT-selected galaxy clusters - I. Evolution of profiles and core properties. *MNRAS*, 474 (1):1065–1098, February 2018. doi: 10.1093/mnras/stx2796.

Sanders, J.S., Dennerl, K., Russell, H.R. et al. Measuring bulk flows of the intracluster medium in the Perseus and Coma galaxy clusters using XMM-Newton. *A&A*, 633:A42, January 2020. doi: 10.1051/0004-6361/201936468.

Sarazin, C.L. X-ray emission from clusters of galaxies. *Reviews of Modern Physics*, 58(1): 1–115, January 1986. doi: 10.1103/RevModPhys.58.1.

Sarazin, C.L. The Physics of Cluster Mergers. In Feretti, L., Gioia, I.M. and Giovannini, G., editors, *Merging Processes in Galaxy Clusters*, volume 272 of *Astrophysics and Space Science Library*, pages 1–38, June 2002. doi: 10.1007/0-306-48096-4_1.

Saro, A., Bocquet, S., Rozo, E. et al. Constraints on the richness-mass relation and the optical-SZE positional offset distribution for SZE-selected clusters. *MNRAS*, 454(3): 2305–2319, December 2015. doi: 10.1093/mnras/stv2141.

Schellenberger, G. and Reiprich, T.H. HICOSMO - cosmology with a complete sample of galaxy clusters - I. Data analysis, sample selection and luminosity-mass scaling relation. *MNRAS*, 469(3):3738–3761, August 2017a. doi: 10.1093/mnras/stx1022.

Schellenberger, G. and Reiprich, T.H. HICOSMO: cosmology with a complete sample of galaxy clusters - II. Cosmological results. *MNRAS*, 471(2):1370–1389, October 2017b. doi: 10.1093/mnras/stx1583.

Schneider, P.C., Freund, S., Czesla, S. et al. The eROSITA Final Equatorial-Depth Survey (eFEDS). The Stellar Counterparts of eROSITA sources identified by machine learning and Bayesian algorithms. *A&A*, 661:A6, May 2022. doi: 10.1051/0004-6361/202141133.

Sehgal, N., Bode, P., Das, S. et al. Simulations of the Microwave Sky. *ApJ*, 709(2):920–936, February 2010. doi: 10.1088/0004-637X/709/2/920.

Seppi, R., Comparat, J., Nandra, K. et al. The mass function dependence on the dynamical state of dark matter haloes. *A&A*, 652:A155, August 2021. doi: 10.1051/0004-6361/202039123.

Seppi, R., Comparat, J., Bulbul, E. et al. Detecting clusters of galaxies and active galactic nuclei in an eROSITA all-sky survey digital twin. *A&A*, 665:A78, September 2022. doi: 10.1051/0004-6361/202243824.

Sereno, M., Giocoli, C., Ettori, S. et al. The mass-concentration relation in lensing clusters: the role of statistical biases and selection effects. *MNRAS*, 449(2):2024–2039, May 2015. doi: 10.1093/mnras/stv416.

Sevilla-Noarbe, I., Bechtol, K., Carrasco Kind, M. et al. Dark Energy Survey Year 3 Results: Photometric Data Set for Cosmology. *ApJS*, 254(2):24, June 2021. doi: 10.3847/1538-4365/abeb66.

Seward, F.D. and Charles, P.A. *Exploring the X-Ray Universe*. 1995.

Shan, H., Kneib, J.P., Li, R. et al. The Mass-Concentration Relation and the Stellar-to-halo Mass Ratio in the CFHT Stripe 82 Survey. *ApJ*, 840(2):104, May 2017. doi: 10.3847/1538-4357/aa6c68.

Sherwin, B.D., Das, S., Hajian, A. et al. The Atacama Cosmology Telescope: Cross-correlation of cosmic microwave background lensing and quasars. *Phys. Rev. D*, 86(8):083006, October 2012. doi: 10.1103/PhysRevD.86.083006.

Sheth, R.K. and Tormen, G. Large-scale bias and the peak background split. *MNRAS*, 308:119–126, September 1999. doi: 10.1046/j.1365-8711.1999.02692.x.

Sheth, R.K. and Tormen, G. An excursion set model of hierarchical clustering: ellipsoidal collapse and the moving barrier. *MNRAS*, 329:61–75, January 2002. doi: 10.1046/j.1365-8711.2002.04950.x.

Shostak, G.S. Integral properties of late-type galaxies derived from H I observations. *A&A*, 68:321–341, August 1978.

Singh, P., Saro, A., Costanzi, M. et al. Cosmology dependence of galaxy cluster scaling relations. *MNRAS*, 494(3):3728–3740, May 2020. doi: 10.1093/mnras/staa1004.

Skibba, R.A., van den Bosch, F.C., Yang, X. et al. Are brightest halo galaxies central galaxies? *MNRAS*, 410(1):417–431, January 2011. doi: 10.1111/j.1365-2966.2010.17452.x.

Skillman, S.W., Warren, M.S., Turk, M.J. et al. Dark Sky Simulations: Early Data Release. *ArXiv e-prints*, July 2014.

Smith, G.P., Mazzotta, P., Okabe, N. et al. LoCuSS: Testing hydrostatic equilibrium in galaxy clusters. *MNRAS*, 456(1):L74–L78, February 2016. doi: 10.1093/mnrasl/slv175.

Smith, R.E., Peacock, J.A., Jenkins, A. et al. Stable clustering, the halo model and non-linear cosmological power spectra. *MNRAS*, 341(4):1311–1332, June 2003. doi: 10.1046/j.1365-8711.2003.06503.x.

Smith, R.K., Brickhouse, N.S., Liedahl, D.A. et al. Collisional Plasma Models with APEC/APED: Emission-Line Diagnostics of Hydrogen-like and Helium-like Ions. *ApJ*, 556(2):L91–L95, August 2001. doi: 10.1086/322992.

Snowden, S.L., Mushotzky, R.F., Kuntz, K.D. et al. A catalog of galaxy clusters observed by XMM-Newton. *A&A*, 478(2):615–658, February 2008. doi: 10.1051/0004-6361:20077930.

Somboonpanyakul, T., McDonald, M., Gaspari, M. et al. The Clusters Hiding in Plain Sight (CHiPS) Survey: Complete Sample of Extreme BCG Clusters. *ApJ*, 910(1):60, March 2021. doi: 10.3847/1538-4357/abe1bc.

Song, H., Hwang, H.S., Park, C. et al. A Redshift Survey of the Nearby Galaxy Cluster A2107: Global Rotation of the Cluster and Its Connection to Large-scale Structures in the Universe. *ApJ*, 869(2):124, December 2018. doi: 10.3847/1538-4357/aaed27.

Soucail, G., Fort, B., Mellier, Y. et al. A blue ring-like structure in the center of the A 370 cluster of galaxies. *A&A*, 172:L14–L16, January 1987.

Spergel, D., Gehrels, N., Baltay, C. et al. Wide-Field InfrarRed Survey Telescope-Astrophysics Focused Telescope Assets WFIRST-AFTA 2015 Report. *arXiv e-prints*, art. arXiv:1503.03757, March 2015.

Springel, V. The cosmological simulation code GADGET-2. *MNRAS*, 364:1105–1134, December 2005. doi: 10.1111/j.1365-2966.2005.09655.x.

Springel, V., White, S.D.M., Jenkins, A. et al. Simulations of the formation, evolution and clustering of galaxies and quasars. *Nature*, 435:629–636, June 2005. doi: 10.1038/nature03597.

Springel, V. and Hernquist, L. Cosmological smoothed particle hydrodynamics simulations: a hybrid multiphase model for star formation. *MNRAS*, 339(2):289–311, February 2003. doi: 10.1046/j.1365-8711.2003.06206.x.

Springel, V., Pakmor, R., Pillepich, A. et al. First results from the IllustrisTNG simulations: matter and galaxy clustering. *MNRAS*, 475(1):676–698, March 2018. doi: 10.1093/mnras/stx3304.

Staniszewski, Z., Ade, P.A.R., Aird, K.A. et al. Galaxy Clusters Discovered with a Sunyaev-Zel'dovich Effect Survey. *ApJ*, 701(1):32–41, August 2009. doi: 10.1088/0004-637X/701/1/32.

Steinborn, L.K., Dolag, K., Hirschmann, M. et al. A refined sub-grid model for black hole accretion and AGN feedback in large cosmological simulations. *MNRAS*, 448(2): 1504–1525, April 2015. doi: 10.1093/mnras/stv072.

Steinborn, L.K., Dolag, K., Comerford, J.M. et al. Origin and properties of dual and offset active galactic nuclei in a cosmological simulation at z=2. *MNRAS*, 458(1):1013–1028, May 2016. doi: 10.1093/mnras/stw316.

Suchyta, E., Huff, E.M., Aleksić, J. et al. No galaxy left behind: accurate measurements with the faintest objects in the Dark Energy Survey. *MNRAS*, 457(1):786–808, March 2016. doi: 10.1093/mnras/stv2953.

Sunyaev, R.A. and Zeldovich, Y.B. Small-Scale Fluctuations of Relic Radiation. *Ap&SS*, 7(1):3–19, April 1970. doi: 10.1007/BF00653471.

Sunyaev, R.A. and Zeldovich, Y.B. 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, November 1972.

Sunyaev, R.A., Norman, M.L. and Bryan, G.L. On the Detectability of Turbulence and Bulk Flows in X-ray Clusters. *Astronomy Letters*, 29:783–790, December 2003. doi: 10.1134/1.1631411.

Takahashi, T., Kokubun, M., Mitsuda, K. et al. The ASTRO-H (Hitomi) x-ray astronomy satellite. In den Herder, J.W.A., Takahashi, T. and Bautz, M., editors, *Space Telescopes and Instrumentation 2016: Ultraviolet to Gamma Ray*, volume 9905 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, page 99050U, July 2016. doi: 10.1117/12.2232379.

Teklu, A.F., Remus, R.S., Dolag, K. et al. Connecting Angular Momentum and Galactic Dynamics: The Complex Interplay between Spin, Mass, and Morphology. *ApJ*, 812(1): 29, October 2015. doi: 10.1088/0004-637X/812/1/29.

The Dark Energy Survey Collaboration. The Dark Energy Survey. *arXiv e-prints*, art. astro-ph/0510346, October 2005.

Thomas, P.A., Muanwong, O., Pearce, F.R. et al. A simulated τCDM cosmology cluster catalogue: the NFW profile and the temperature-mass scaling relations. *MNRAS*, 324 (2):450–462, June 2001. doi: 10.1046/j.1365-8711.2001.04330.x.

Tinker, J., Kravtsov, A.V., Klypin, A. et al. Toward a Halo Mass Function for Precision Cosmology: The Limits of Universality. *ApJ*, 688:709-728, December 2008. doi: 10.1086/591439.

Tinker, J.L., Robertson, B.E., Kravtsov, A.V. et al. The Large-scale Bias of Dark Matter Halos: Numerical Calibration and Model Tests. *ApJ*, 724(2):878–886, December 2010. doi: 10.1088/0004-637X/724/2/878.

Tinker, J.L., Sheldon, E.S., Wechsler, R.H. et al. Cosmological Constraints from Galaxy Clustering and the Mass-to-number Ratio of Galaxy Clusters. *ApJ*, 745(1):16, January 2012. doi: 10.1088/0004-637X/745/1/16.

To, C., Krause, E., Rozo, E. et al. Dark Energy Survey Year 1 Results: Cosmological Constraints from Cluster Abundances, Weak Lensing, and Galaxy Correlations. *Phys. Rev. Lett.*, 126(14):141301, April 2021. doi: 10.1103/PhysRevLett.126.141301.

Tornatore, L., Borgani, S., Dolag, K. et al. Chemical enrichment of galaxy clusters from hydrodynamical simulations. *MNRAS*, 382(3):1050–1072, December 2007. doi: 10.1111/j.1365-2966.2007.12070.x.

Tovmassian, H.M. The Rotation of Galaxy Clusters. *Astrophysics*, 58(3):328–337, September 2015. doi: 10.1007/s10511-015-9387-4.

Trudeau, A., Garrel, C., Willis, J. et al. The XXL Survey. XLII. Detection and characterisation of the galaxy population of distant galaxy clusters in the XXL-N/VIDEO field: A tale of variety. *A&A*, 642:A124, October 2020. doi: 10.1051/0004-6361/202038982.

Truemper, J. The ROSAT mission. *Advances in Space Research*, 2(4):241–249, January 1982. doi: 10.1016/0273-1177(82)90070-9.

Umetsu, K. Cluster-galaxy weak lensing, 2020.

Valageas, P. and Clerc, N. Redshift-space correlation functions in large galaxy cluster surveys. *A&A*, 547:A100, November 2012. doi: 10.1051/0004-6361/201219646.

van den Bosch, F.C., Norberg, P., Mo, H.J. et al. Probing dark matter haloes with satellite kinematics. *MNRAS*, 352(4):1302–1314, August 2004. doi: 10.1111/j.1365-2966.2004.08021.x.

van Uitert, E., Gilbank, D.G., Hoekstra, H. et al. Weak-lensing-inferred scaling relations of galaxy clusters in the RCS2: mass-richness, mass-concentration, mass-bias, and more. *A&A*, 586:A43, February 2016. doi: 10.1051/0004-6361/201526719.

van Uitert, E., Joachimi, B., Joudaki, S. et al. KiDS+GAMA: cosmology constraints from a joint analysis of cosmic shear, galaxy-galaxy lensing, and angular clustering. *MNRAS*, 476(4):4662–4689, June 2018. doi: 10.1093/mnras/sty551.

van Weeren, R.J., Röttgering, H.J.A., Brüggen, M. et al. Particle Acceleration on Megaparsec Scales in a Merging Galaxy Cluster. *Science*, 330(6002):347, October 2010. doi: 10.1126/science.1194293.

Veropalumbo, A., Marulli, F., Moscardini, L. et al. An improved measurement of baryon acoustic oscillations from the correlation function of galaxy clusters at z $\sim$ 0.3. *MNRAS*, 442(4):3275–3283, August 2014. doi: 10.1093/mnras/stu1050.

Viel, M., Becker, G.D., Bolton, J.S. et al. Warm dark matter as a solution to the small scale crisis: New constraints from high redshift Lyman-α forest data. *Phys. Rev. D*, 88 (4):043502, August 2013. doi: 10.1103/PhysRevD.88.043502.

Viitanen, A., Allevato, V., Finoguenov, A. et al. The XMM-Newton wide field survey in the COSMOS field: Clustering dependence of X-ray selected AGN on host galaxy properties. *A&A*, 629:A14, September 2019. doi: 10.1051/0004-6361/201935186.

Vikhlinin, A., McNamara, B.R., Forman, W. et al. A Catalog of 200 Galaxy Clusters Serendipitously Detected in the ROSAT PSPC Pointed Observations. *ApJ*, 502(2):558–581, August 1998. doi: 10.1086/305951.

Vikhlinin, A., Kravtsov, A., Forman, W. et al. Chandra Sample of Nearby Relaxed Galaxy Clusters: Mass, Gas Fraction, and Mass-Temperature Relation. *ApJ*, 640(2):691–709, April 2006. doi: 10.1086/500288.

Vikhlinin, A., Burenin, R.A., Ebeling, H. et al. Chandra Cluster Cosmology Project. II. Samples and X-Ray Data Reduction. *ApJ*, 692(2):1033–1059, February 2009a. doi: 10.1088/0004-637X/692/2/1033.

Vikhlinin, A., Kravtsov, A.V., Burenin, R.A. et al. Chandra Cluster Cosmology Project III: Cosmological Parameter Constraints. *ApJ*, 692(2):1060–1074, February 2009b. doi: 10.1088/0004-637X/692/2/1060.

Virtanen, P., Gommers, R., Oliphant, T.E. et al. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17:261–272, 2020. doi: 10.1038/s41592-019-0686-2.

Voges, W., Aschenbach, B., Boller, T. et al. The ROSAT all-sky survey bright source catalogue. *A&A*, 349:389–405, September 1999.

Wang, J., Bose, S., Frenk, C.S. et al. Universality in the structure of dark matter haloes over twenty orders of magnitude in halo mass. *arXiv e-prints*, art. arXiv:1911.09720, November 2019.

Weinberg, D.H., Mortonson, M.J., Eisenstein, D.J. et al. Observational probes of cosmic acceleration. *Phys. Rep.*, 530(2):87–255, September 2013. doi: 10.1016/j.physrep.2013.05.001.

Weinberger, R., Springel, V., Hernquist, L. et al. Simulating galaxy formation with black hole driven thermal and kinetic feedback. *MNRAS*, 465(3):3291–3308, March 2017. doi: 10.1093/mnras/stw2944.

Weinberger, R., Springel, V., Pakmor, R. et al. Supermassive black holes and their feedback effects in the IllustrisTNG simulation. *MNRAS*, 479(3):4056–4072, September 2018. doi: 10.1093/mnras/sty1733.

Weinberger, R., Springel, V. and Pakmor, R. The AREPO Public Code Release. *ApJS*, 248(2):32, June 2020. doi: 10.3847/1538-4365/ab908c.

Weinmann, S.M., van den Bosch, F.C., Yang, X. et al. Properties of galaxy groups in the Sloan Digital Sky Survey - I. The dependence of colour, star formation and morphology on halo mass. *MNRAS*, 366(1):2–28, February 2006. doi: 10.1111/j.1365-2966.2005.09865.x.

Weisskopf, M.C., Tananbaum, H.D., Van Speybroeck, L.P. et al. Chandra X-ray Observatory (CXO): overview. In Truemper, J.E. and Aschenbach, B., editors, *X-Ray Optics, Instruments, and Missions III*, volume 4012 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, pages 2–16, July 2000. doi: 10.1117/12.391545.

Weißmann, A., Böhringer, H., Šuhada, R. et al. Studying the properties of galaxy cluster morphology estimators. *A&A*, 549:A19, January 2013. doi: 10.1051/0004-6361/201219333.

Wen, Z.L., Han, J.L. and Liu, F.S. A Catalog of 132,684 Clusters of Galaxies Identified from Sloan Digital Sky Survey III. *ApJS*, 199(2):34, April 2012. doi: 10.1088/0067-0049/199/2/34.

White, S.D.M., Navarro, J.F., Evrard, A.E. et al. The baryon content of galaxy clusters: a challenge to cosmological orthodoxy. *Nature*, 366(6454):429–433, December 1993. doi: 10.1038/366429a0.

Wiersma, R.P.C., Schaye, J., Theuns, T. et al. Chemical enrichment in cosmological, smoothed particle hydrodynamics simulations. *MNRAS*, 399(2):574–600, October 2009. doi: 10.1111/j.1365-2966.2009.15331.x.

Will, C.M. The Confrontation between General Relativity and Experiment. *Living Reviews in Relativity*, 17(1):4, December 2014. doi: 10.12942/lrr-2014-4.

Wilms, J., Allen, A. and McCray, R. On the Absorption of X-Rays in the Interstellar Medium. *ApJ*, 542(2):914–924, October 2000. doi: 10.1086/317016.

Wojtak, R. and Łokas, E.L. Mass profiles and galaxy orbits in nearby galaxy clusters from the analysis of the projected phase space. *MNRAS*, 408(4):2442–2456, November 2010. doi: 10.1111/j.1365-2966.2010.17297.x.

Wolter, H. Spiegelsysteme streifenden Einfalls als abbildende Optiken für Röntgenstrahlen. *Annalen der Physik*, 445(1):94–114, January 1952. doi: 10.1002/andp.19524450108.

Yan, Z., Raza, N., Van Waerbeke, L. et al. An analysis of galaxy cluster mis-centring using cosmological hydrodynamic simulations. *MNRAS*, 493(1):1120–1129, March 2020. doi: 10.1093/mnras/staa295.

Yang, X., Mo, H.J., van den Bosch, F.C. et al. Galaxy Groups in the SDSS DR4. I. The Catalog and Basic Properties. *ApJ*, 671(1):153–170, December 2007. doi: 10.1086/522027.

York, D.G., Adelman, J., Anderson, John E., J. et al. The Sloan Digital Sky Survey: Technical Summary. *AJ*, 120(3):1579–1587, September 2000. doi: 10.1086/301513.

Zenteno, A., Hernández-Lang, D., Klein, M. et al. A joint SZ-X-ray-optical analysis of the dynamical state of 288 massive galaxy clusters. *MNRAS*, 495(1):705–725, June 2020. doi: 10.1093/mnras/staa1157.

Zhang, C., Yu, Q. and Lu, Y. Offsets between the X-Ray and the Sunyaev-Zel'dovich-effect Peaks in Merging Galaxy Clusters and their Cosmological Implications. *ApJ*, 796(2):138, December 2014. doi: 10.1088/0004-637X/796/2/138.

Zhang, Y., Jeltema, T., Hollowood, D.L. et al. Dark Energy Surveyed Year 1 results: calibration of cluster mis-centring in the redMaPPer catalogues. *MNRAS*, 487(2):2578–2593, August 2019. doi: 10.1093/mnras/stz1361.

Zhao, D.H., Mo, H.J., Jing, Y.P. et al. The growth and structure of dark matter haloes. *MNRAS*, 339(1):12–24, February 2003. doi: 10.1046/j.1365-8711.2003.06135.x.

Zhao, D.H., Jing, Y.P., Mo, H.J. et al. Accurate Universal Models for the Mass Accretion Histories and Concentrations of Dark Matter Halos. *ApJ*, 707(1):354–369, December 2009. doi: 10.1088/0004-637X/707/1/354.

Zhuravleva, I., Churazov, E., Kravtsov, A. et al. Quantifying properties of ICM inhomogeneities. *MNRAS*, 428(4):3274–3287, February 2013. doi: 10.1093/mnras/sts275.

Zitrin, A., Bartelmann, M., Umetsu, K. et al. Miscentring in galaxy clusters: dark matter to brightest cluster galaxy offsets in 10 000 Sloan Digital Sky Survey clusters. *MNRAS*, 426(4):2944–2956, November 2012. doi: 10.1111/j.1365-2966.2012.21886.x.

Zubeldia, Í. and Challinor, A. Cosmological constraints from Planck galaxy clusters with CMB lensing mass bias calibration. *MNRAS*, 489(1):401–419, October 2019. doi: 10.1093/mnras/stz2153.

Zwicky, F. Die Rotverschiebung von extragalaktischen Nebeln. *Helvetica Physica Acta*, 6:110–127, January 1933.

Danksagung

The past three years have been very challenging, but incredibly rewarding as well. The covid pandemic partially ruined the adventure, but it also taught us to appreciate the privilege of our own journey on a daily basis. Words cannot express my gratitude to a very long list of people, who helped me during this PhD. I'll do my best to acknowledge all of them.

Not a single line of this thesis would have been written without the support and guidance from Johan. I will forever owe you so much for everything you did for me, I wish you and your family all the very best for whatever the future holds, thank you for investing time and energy in me.

I'm extremely grateful to my formal supervisor Paul, for always providing very useful insights and suggestions about my projects, and the regular Friday Quarantinis during the lockdown :)
I would also like to thank Esra, for creating a very exciting working environment in the cluster group at MPE; Vittorio, for the informal and science discussions, and sharing the Trentino office in the beginning; Andrea and Mara, for always looking out for the students; Annette, for taking care of the IMPRS program; Birgit, for helping with all the paperwork and the HEG events; Harald, for the efficient IT support; and everyone else in the cluster group and the high energy group at MPE.

I would like to express my deepest gratitude to the new friends I met in Munich. In particular Martin, Jacob, Adam, Emre, Sophia, Soumya, Francesco, Toska, David, Zsofi, thank you for the lunch breaks, coffees, the various adventures, hikes, trips, birthdays, dinners, and creating a friendly atmosphere around me; Ricky, for also sharing Ws and Ls on the court and being the best v0 version of me; Aish, Max, Joaquin, for the fun times at the beginning of covid; you were all fundamental parts of my life in Munich and made this so much more than just a PhD! Thanks again to my office mate David for helping with my Zusammenfassung, reaching the state of the art of the German language (die Zustand der Kunstmodelle!). Thanks to all the IMPRS students I met, each one of us is essential in creating this involving international community, I was glad to be an IMPRS rep and I hope that each current and future student will have a good experience in our program. Special thanks to my DM friends from Bologna, we barely met the last three years, but staying in touch with you and sharing our experiences from different places has been helpful; in particular to Fede, for the Bamberg–Munich exchanges; and Chiara, for the support during the last few months.

Thanks to Safyie, Soumya, and Adam for joining the eRUNita team.
Thanks to coach Nik, Marius, and all my teammates of MILBaskets, playing this game that I love so much was of the utmost importance for me and doing it with you is just perfect. TEAMWORK!

Muchas gracias to Iago, for lightening the atmosphere at the house during the crazy start of the pandemic; and Solem, for the light-hearted times, the stimulating discussions, the italian-iranian-spanish meals, and for making the apartment the chill and fun place I needed daily.

Thank you Pol, for sharing the first marathon of our running career together; Gullino and JJ, our holidays were some of the funniest times I had and you allowed me to forget about work for some time (Richi +1 per questo PhD); to all my other friends back in the valley, for welcoming me back everytime like I never left.

Voglio fare un ringraziamento in particolare ai miei nonni, Marisa e Roberto, che mi controllano sempre dall'alto, Angelina e Adriano, mi avete sempre accolto a braccia aperte ad ogni visita, grazie infinite.
Infine, non avrei mai portato a termine questo percorso senza i miei genitori Rosella e Angelo, il vostro supporto non mi é mai mancato quando ne ho avuto bisogno, negli ultimi 28 anni cosí come durante questo dottorato; e mia sorella Federica, grazie per tutte le videochiamate, l'ospitalitá durante le visite a Milano e le riunioni per le festivitá, ti ho sentita vicina ogni giorno nonostante la distanza e questo é stato fondamentale.