Universidade de São Paulo
Instituto de Astronomia, Geofísica e Cincias Atmosféricas
Departamento de Astronomia

Isabel Bonomo Pederneiras

The evolution of the mass - X-ray luminosity relation in the CODEX galaxy cluster sample

São Paulo
2024

Isabel Bonomo Pederneiras

The evolution of the mass - X-ray luminosity relation in the CODEX galaxy cluster sample

Dissertação apresentada ao Departamento de Astronomia do Instituto de Astronomia, Geofísica e Ciências Atmosféricas da Universidade de São Paulo como requisito parcial para a obtenção do título de Mestre em Ciencias.

Versão Corrigida. O original encontra-se disponível na Unidade.

Área de Concentração: Astronomia
Orientador: Prof. Dr. Eduardo Serra Cypriano

São Paulo

2024

Agradecimentos

Gostaria de expressar minha gratidão a todos que contribuíram para a realização deste trabalho de mestrado.

Em primeiro lugar, meu profundo agradecimento ao meu orientador, Prof. Eduardo Cypriano, pela orientação excepcional, dedicação e por ter me ensinado tanto ao longo deste processo. Sua orientação foi fundamental para o desenvolvimento deste trabalho, e sou imensamente grata pela oportunidade de trabalhar sob sua liderança.

Ao meu colaborador, Prof. Alexis Finoguenov, expresso minha gratidão pela orientação adicional fornecida. Além das contribuições mais concretas, como o compartilhamento de dados e códigos, suas sugestões e correções também foram fundamentais para a qualidade e rigor deste estudo.

Minha sincera gratidão aos meus pais, Paulo e Paula, que sempre me inspiraram pela determinação e força com que levam a vida. Agradeço por oferecerem o apoio necessário quando optei por embarcar no mestrado e mudar para outra cidade, e pela compreensão durante todo o período de pesquisa. Agradeço também aos meus irmãos, Lucas e Maria, que me apoiaram e aturaram na infância, para atualmente se tornarem meus amigos mais próximos.

Ao meu namorado, Matheus, pela paciência e compreensão de quem presenciou de perto toda essa jornada acadêmica. Agradeço por sempre acreditar no meu potencial e me incentivar nos momentos de dificuldade. Mais importante ainda, obrigada pela sua presença nas ocasiões alegres e descontraídas, que me fazem ansear pelo futuro.

Aos meu amigos de longa data, que sempre estiveram ao meu lado e me ajudaram a superar os obstáculos em cada passo da vida. Obrigada por fazerem parte das minhas memórias mais importantes e dos meus momentos de conquista. Agradeço também às

novas amizades feitas durante este percurso, que me acolheram no Instituto e me guiaram nesta nova etapa da vida acadêmica, sempre me encorajando a seguir em frente.

Aos funcionários e docentes da Universidade de São Paulo, que formam a base essencial para um ambiente de trabalho próspero. Não posso deixar de mencionar agradecimentos especiais aos técnicos de informática, Marco e Luis, cuja expertise e prontidão foram cruciais para superar obstáculos e garantir a continuidade da pesquisa.

A todos vocês, meu mais sincero obrigado. Este trabalho não seria possível sem o apoio e contribuições de cada um.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) por meio do Programa de Excelência Acadêmica (PROEX) – Código de Financiamento 88887.638120/2021-00.

"*No darkness lasts forever. And even there, there are stars.*"

Ursula K. Le Guin

"*Não se iludam, não me iludo, tudo agora mesmo pode estar por um segundo.*"

Gilberto Gil

Resumo

Medições precisas das relações de escala de raios-X em aglomerados de galáxias são fundamentais para investigar a cosmologia e estudar processos barionicos no meio intra-aglomerado. A correlação entre a luminosidade de raios-X e a massa total dos aglomerados, conforme estabelecida no modelo auto-similar puramente gravitacional (Kaiser 1986), fornece um padrão de referência. Contudo, desvios desse modelo tornam-se evidentes quando processos não gravitacionais, como resfriamento radiativo e feedback por AGNs, exercem influência, podendo resultar em uma evolução temporal da relação $M - L_X$.

Até o desenvolvimento deste trabalho, a calibração precisa da relação $M - L_X$ e sua evolução, utilizando estimativas de massa por lentes gravitacionais e uma metodologia estatística consistente, estava ausente na literatura. Nossa abordagem utiliza o catálogo COnstrain Dark Energy with X-ray clusters (CODEX), construído meticulosamente através de seleções no óptico e em raios-X de aglomerados na área de sobreposição do Rosat All Sky Survey (RASS) e do Sloan Digital Sky Survey (SDSS). Analisamos 101 aglomerados de galáxias do CODEX para os quais temos estimativas de massa por lentes gravitacionais. Este método se destaca no projeto por ser sensível a toda a matéria do aglomerado, sem a necessidade de suposições sobre a dinâmica do sistema.

Introduzindo um modelo bayesiano hierárquico para a análise da relação de escala, consideramos erros de medição heteroscedásticos, dispersão intrínseca, funções de seleção no óptica e em raios-X, e os critérios de construção da subamostra. Nossa amostra de aglomerados CODEX se ajusta bem a uma lei de potência no espaço logarítmico, com uma inclinação de 0,3±0,05, valor significativamente menor do que a previsão auto-similar. Também relatamos um parâmetro de evolução de 1,54 ± 0,55, um valor estatisticamente significativo para inferir uma certa evolução positiva da relação de escala.

Ao investigarmos esses resultados intrigantes, destacamos nossa abordagem estatística distinta e nossa análise centrada em uma amostra de aglomerados de alta massa como possíveis fatores. Também conjecturamos sobre os efeitos de feedback por AGNs e um possível viés de evolução devido a uma dependência com o redshift da distribuição de L_X. Observamos ainda que, em comparação com estudos anteriores, nosso trabalho foi o primeiro a refinar ainda mais o parâmetro de evolução para erros abaixo do valor medido.

Abstract

Precise measurements of X-ray scaling relations in galaxy clusters are imperative to probe cosmology and to study baryonic processes in the intercluster medium. The functional form of the correlation between X-ray luminosity and the total mass of galaxy clusters is well established in the purely gravitational self-similar model (Kaiser 1986). However, deviations from this model become apparent when non-gravitational processes, such as radiative cooling and AGN feedback, exert influence, potentially leading to a temporal evolution of the $M - L_X$ relation.

Up until this work, precise calibration of the $M - L_X$ relation and its evolution, employing weak lensing mass estimates and a consistent statistical methodology, was still lacking. Our calibration utilizes the COnstrain Dark Energy with X-ray clusters (CODEX) catalog, which was thoroughly constructed by adopting an X-ray and optical selection of clusters in the superposition area of Rosat All Sky Survey (RASS) and Sloan Digital Sky Survey (SDSS). We analyze 101 CODEX galaxy clusters for which we have weak lensing mass estimates. This method is a distinction of the project since it is sensitive to the entire matter of the cluster, with no assumptions needed for the system dynamics.

By introducing a Bayesian hierarchical model for scaling relation analysis, we account for heteroscedastic measurement errors, intrinsic scatter, optical and X-ray selection functions, and the subsample's construction criteria. The CODEX weak-lensing sample is well fit by a power law in log space with a slope of 0.3 ± 0.05, which is significantly shallower than the self-similar prediction. We also report an evolution parameter of 1.54 ± 0.55, which is statistically significant to infer a positive evolution of the scaling relation.

As we investigate these intriguing findings, we point out our distinctive statistical approach and our analysis focused on a high-mass cluster sample as potential factors. We

also conjecture the effects of AGN feedback and a possible evolution bias due to a redshift dependence of the L_X distribution. We also note that, when compared with previous studies, our work was the first to further constrain the evolution parameter to errors below the measured value.

List of Figures

List of Tables

Contents

Chapter 1

Introduction

1.1 Cosmological background

During the first decade of the 20^{th} century, the widespread acceptance of the *Lorentz transformations* and the formulation of the *special theory of relativity* unified the once separate concepts of time and space, combining them into a single four-dimensional quantity. The *spacetime* coordinate system, composed of three elements for the spatial dimensions and one element for time, is usually represented as $x^\mu = \{ct, x, y, z\}$. If before the length of a certain object or the duration of a certain event was unaffected by the observer's rest frame, these properties are no longer invariant in the four-dimensional spacetime. In this new scenario, the quantity that remains constant under changes in coordinates is the so-called *spacetime interval*, defined as:

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu \ , \tag{1.1}$$

where $g_{\mu\nu}$ is the metric tensor representing the main theoretical assumptions of the system we are interested in.

A physical assumption that permeates all the different models currently accepted in cosmology is the well-known '*Cosmological Principle*', which was first clearly mentioned by Newton at the end of the 17^{th} century. Essentially it places all points in the Universe on an equal footing by stating that their properties do not change according to location or direction, making any place undistinguished to any other. This may seem off when we consider the lack of uniformity seen in the structures surrounding us, such as the abundance of astronomical objects in some regions while others are essentially empty spaces. This is to say that the Universe must be homogeneous and isotropic on *large* scales, over ~ 200

Mpc (Lima Neto, 2020; Scrimgeour et al., 2012).

The general metric system $g_{\mu\nu}$ used to describe our Universe must then respect such Cosmological Principle. From the isotropic assumption, it follows that, if the Universe expands or contracts, this phenomenon must be global (i.e. independent of spatial coordinates), leading to the definition of a *scale factor* $a = a(t)$. As for the homogeneous assumption, it states that our metric must be the same at every point in space, such that we must have $g_{\mu\nu}(x^\alpha) = g_{\mu\nu}$.

With Eintein's *theory of general relativity* published in 1915, we also introduce gravitation to our spacetime framework, where gravity is essentially a manifestation of the *curvature* of the spacetime. One of the simplest examples of a metric system is that of a static universe with no gravity (i.e. free-fall) and hence no curvature of the spacetime, which is called the *Minkowski metric*, given by $ds^2 = c^2dt^2 - dx^2 - dy^2 - dz^2$.

The more general metric representing an expanding (or contracting) homogeneous and isotropic universe with a certain curvature K, is the so-called Friedmann–Lemaître–Robertson–Walker metric (FLRW):

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

Indeed, the expansion of our Universe was conjectured in the 1920s with the works of Hubble (1929) and Lemaître (1927), who showed that galaxies are moving away with velocities proportional to their distances, expressed as $v = H_0D$. The expansion rate is time-dependent and is called the *Hubble-Lamaitre parameter* $H(t) = \dot{a}(t)/a(t)$, with its present value denoted as $H(t_0) = H_0$. In this scenario, the scale factor is linked to the redshift as

$$a(z) = \frac{1}{(1+z)} . \tag{1.3}$$

Furthermore, Einstein's field equations relate the metric system, which describes how spacetime behaves, to the total matter content of our Universe. Here the term 'matter' is used in a more broad sense, which includes baryonic matter, dark matter, radiation, and dark energy. The amount of each component affects the expansion rate differently, for example,

$$\begin{aligned} &\text{non-relativistic matter} \rightarrow \rho_m \propto a^{-3} \\ &\text{relativistic matter and radiation} \rightarrow \rho_r \propto a^{-4} \\ &\text{dark energy} \rightarrow \rho_\Lambda \propto a^{0}. \end{aligned} \tag{1.4}$$

Using the solutions of Einstein's field equations for the FLRW metric one can find the Friedmann equation:

$$\frac{\dot{a}^2}{a^2} = \frac{8\pi G}{3}\rho - \frac{Kc^2}{a^2} \ . \tag{1.5}$$

By setting $K = 0$, we can define a critical density today as

$$\rho_{c,0} = 3H_0^2/8\pi G \ . \tag{1.6}$$

We can also describe the ratio of each matter component as a dimensionless density $\Omega = \rho/\rho_{c,0}$. Using this notation, and introducing $\Omega_K = Kc^2/\dot{a}^2$ as a 'curvature component', we can rewrite equation 1.5 as

$$H^2(t) = \frac{\dot{a}^2}{a^2} = H_0^2[\Omega_\Lambda + \Omega_r a^{-4} + \Omega_m a^{-3} + \Omega_K a^{-2}] \ . \tag{1.7}$$

Various ratios of these matter components may describe distinct epochs of our universe, i.e., different behaviors in the expansion and growth of structures. A flat universe with only non-relativistic, pressureless, matter ($\Omega_m = 1$) is called an Einstein-de Sitter Universe and is a good approximation for a matter-dominated epoch of our Universe, such as during the formation of most structures we observe today. The standard Λ Cold Dark Matter (ΛCDM) Universe, on the other hand, assumes a predominance of dark energy (around 70%) and no radiation, and it is the most adopted model to describe our local Universe. Throughout this work, we assume a flat ΛCDM cosmology with $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$ and $H_0 = 70\ h_{70}\ km\ s^{-1}\ Mpc^{-1}$, unless stated otherwise.

1.1.1 Structure formation

In our current understanding of an expanding Universe, the Cosmological Principle - and the observed structures we find today - suggest the presence of primordial fluctuations in an otherwise extremely homogeneous early Universe, that inflated rapidly, intensifying the density contrasts. Over time, the gravitational instability in regions of overdensity grows until reaching the collapse stage, ultimately resulting in the formation of a structure.

The strong gravitational influence from the most massive peaks of overdensity in the early universe leads to the formation of large filamentary structures, which eventually creates what we now observe to be a 'Cosmic Web'. The collapse of dark matter (DM) halos through accretion, necessary for the formation of galaxy clusters, occurs in the 'nodes' of these filamentary structures, as illustrated in figure 1.1.

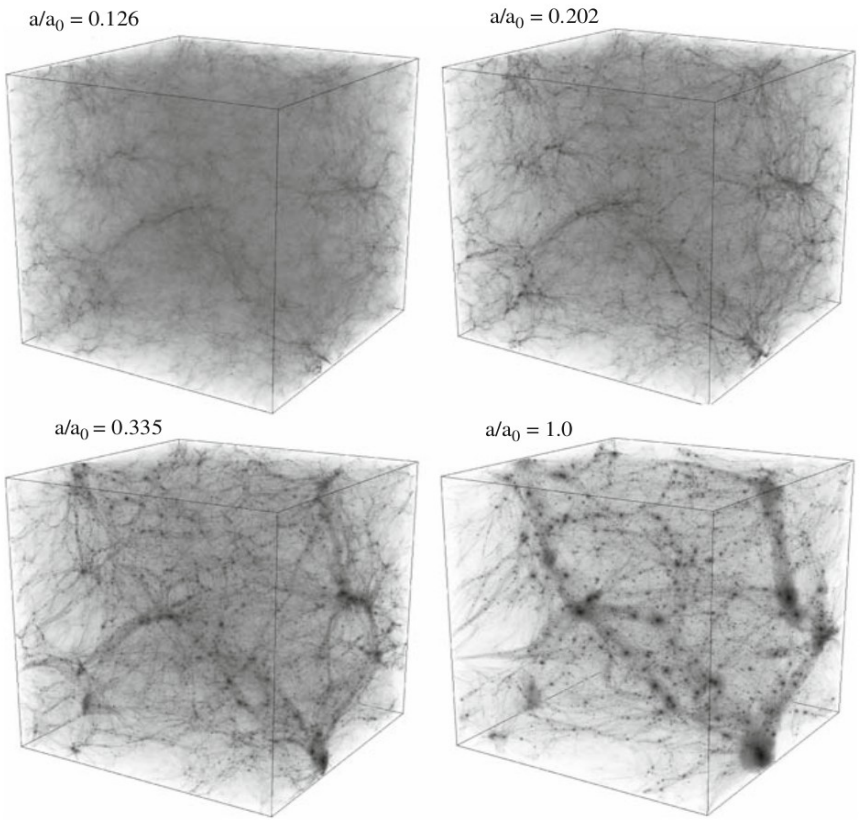


Figure 1.1: A N-body simulation for the formation of filamentary structures and of galaxy clusters. The figure taken from Rich (2010) shows four epochs, using 512^3 particles in a volume of $20h^{-1}$ Mpc.

According to the most accepted model of cold dark matter (CDM), the objects we observe today in the Universe would have been formed in a hierarchical manner (Press and Schechter, 1974; White and Rees, 1978; Blumenthal et al., 1984). In this scenario,

the first structures would have been small ($M \sim 10^6 M_\odot$) and progressively merged during the evolution of the Universe, giving rise to larger objects. Therefore, in a Universe filled with large structures and low-density regions, galaxy clusters stand out for being the most massive systems that have already reached a certain degree of dynamic maturity.

1.2 Galaxy clusters

At the end of the 18th century, astronomers Charles Messier and William Herschel realized that galaxies - at the time perceived only as diffuse objects (nebulae) - tended to cluster. The discovery that such nebulae were in fact galaxies like the Milky Way, only observed at great distances, was made by Edwin Hubble in the 1920s, and from this moment on began the most in-depth studies on clusters of galaxies (Hubble, 1925, 1926). During the following decade, measurements of velocity dispersion of galaxies in some clusters (e.g. Coma and Virgo clusters) showed that the masses of these structures were more than 200 times greater than the observed mass of all cluster stars, if one assumes the systems were in virial equilibrium (Zwicky, 1933; Smith, 1936). To explain this discrepancy between masses, astronomer Fritz Zwicky coined the term 'dark matter' when referring to a large amount of non-visible matter that supposedly permeates the clusters. Over the course of the century, even with more accurate measurements of stellar masses and the discovery that a large amount of mass is actually present in the form of hot gas throughout the cluster, the need for the presence of DM in galaxy clusters became increasingly evident. Besides being an important structure for studying DM properties, galaxy clusters are also thought to store characteristics of the Universe at the time they were formed. For example, the baryonic content in clusters is expected to approach the cosmological baryon fraction $f = \Omega_{bary}/\Omega_m$ (Angelinelli et al., 2023; Kravtsov and Borgani, 2012).

Figure 1.2: Near-infrared image of galaxy cluster SMACS 0723 from NASA's James Webb Space Telescope.

In their unique position among other astronomical objects, galaxy clusters have become ideal laboratories for testing models of structure formation, constraining parameters of cosmological models, and for unveiling dark matter properties (e.g. Voit, 2005; Allen et al., 2011). Several other astrophysical phenomena also have galaxy clusters as a focus of interest, such as the physics of the hot and ionized gas in the intercluster medium, and the effects of perturbations caused by non-gravitational energy sources such as Active Galactic Nuclei (AGN) (McNamara and Nulsen, 2007). Therefore, it is not surprising that the quest to understand the formation and evolution of these objects has become one of the central efforts of modern astrophysics.

1.2.1 The intracluster medium

According to the hierarchical model, in the formation of large structures such as galaxy clusters, the gas is heated to high temperatures by adiabatic compression and shocks during the collapse, when it reaches hydrostatic equilibrium within the potential well itself. Once sufficiently dense, the cluster begins to cool, triggering star formation and accretions into supermassive black holes, which can result in energy injection by supernovae (SN) and AGNs (Kravtsov and Borgani, 2012).

The general notion of when a cluster reaches its dynamical maturity - or 'relaxation' - is given by the *virial equilibrium*, i.e. when we apply the virial theorem to a static system of an ideal fluid with no external pressure (Mo et al., 2020); on which case we obtain

$$-2K = W \ , \tag{1.8}$$

where K and W are the kinetic and internal energies, respectively.

The collapse described in this scenario also results in the formation of the intracluster medium (ICM), consisting of a diffuse plasma composed mainly of ionized hydrogen and helium atoms, but which can be enriched with heavier atoms by feedback processes (e.g. SN and AGNs). The ICM represents most of the volume of normal baryonic matter in clusters, with a mass up to ten times greater than the stellar mass in member galaxies (Mo et al., 2020) - only $\sim 5\%$ of the mass in clusters of galaxies is due to stars, and about $\sim 15\%$ is contributed by the intergalactic gas, whereas the remaining $\sim 80\%$ consists of dark matter (Schneider, 2015).

The intracluster gas is very rarefied, with a central density of approximately $10^{-2} < n < 10^{-3}\, cm^{-3}$, and an average temperature of $T \sim 10^{7} - 10^{8}\, K$ (Lima Neto, 2020). In this configuration, it can be considered that most of the free electron population of the ICM are thermal (i.e. non-relativistic), which is consistent with the strong X-ray emission observed by these structures (section 1.4).

1.3 Self-similar model

Due to the general complexity of galaxy cluster systems, their formation is usually approximated by non-dissipative models, which are very successful in predicting the existence and functional form of correlations between cluster properties. One of the models adopted is the so-called *self-similar* model, proposed by Kaiser (1986), and is based on three main assumptions: (1) clusters form from the gravitational collapse of peaks of the initial density field in an Einstein-de Sitter Universe ($\Omega_m = 1$), guaranteeing a scale-free collapse; (2) the amplitudes of the density fluctuations are power law functions of the size of these fluctuations, such that the initial perturbations are also scale-free; and finally, (3) that other processes that influence halo properties do not introduce new scales to the model (Kravtsov and Borgani, 2012). By adopting such assumptions, we consider clusters to be scale versions of each other, and hence it is expected that the main physical properties of any cluster (e.g. L_X, T_X, velocity dispersion) can be described by power laws.

To obtain these power laws, apart from the self-similar hypothesis, we also consider a

spherically symmetric system in which the intracluster medium is in thermal equilibrium inside the potential well, and hence the hydrostatic equilibrium equation (HE) is valid:

$$\nabla P(r) = -\rho(r)\nabla\Phi(r) \; . \tag{1.9}$$

Adopting the equation of state for an ideal gas, we can write the pressure and potential gradients as

$$\frac{dP}{dr} = \frac{d}{dr}\left(\frac{k_B T \rho}{\mu m_p}\right) \; ; \;\; \frac{d\Phi}{dr} = \frac{GM}{r^2} \; , \tag{1.10}$$

where μ is the mean molecular weight of the gas, m_p is the proton mass and ρ is the mass density.

Substituting these expressions in the HE equation, it is possible to infer the mass M within a characteristic radius R as being proportional to RT.

$$\begin{gathered} M = \frac{k_B}{G\mu m_p} TR\left(-\frac{d\log\rho}{d\log r} - \frac{d\log T}{d\log r}\right) \\ \Rightarrow M \propto RT \; . \end{gathered} \tag{1.11}$$

If we consider the baryon fraction f to be constant in galaxy clusters, and define the gas mass as being simply the integral of the mass density ρ over the volume (i.e. $M_{gas} \propto \rho \, R^3$), we find

$$\begin{aligned} f &= \frac{M_{gas}}{M} = cost. \\ &\Rightarrow R \propto T^{1/2} \, \rho^{-1/2} \; , \end{aligned} \tag{1.12}$$

where we assumed the same radius for both masses.

Using the proportionalities found in 1.11 and 1.12, it is possible to define the scaling relation:

$$\Rightarrow T \propto M^{2/3} \, \rho^{1/3} \; . \tag{1.13}$$

To make the $M - L_X$ relation explicit, one must consider that the X-ray luminosity is given by the integral of the *bremsstrahlung* emissivity ($\epsilon \propto \rho^2 \, T^{-1/2} \, exp(E/kT)$) over the energy and volume, which will be discussed in detail in section 1.4. From this integration, we find that

$$L_X \propto \rho^2 \; R^3 \; T^{1/2} \; . \tag{1.14}$$

Using equation 1.13 to replace T, we find

$$\begin{aligned} L_X &\propto \rho M\ T^{1/2} \\ L_X &\propto M^{4/3}\ \rho^{7/6}\,. \end{aligned} \tag{1.15}$$

The density ρ, in turn, is directly related to the critical density defined in equation 1.6

$$\rho = \Delta\ \rho_c\ , \rho_c \propto H^2(z) \tag{1.16}$$

where Δ is a constant that represents the characteristic density contrast.

From the Friedmann equation 1.7 we can also define the parameter $E(z) = H(z)/H_0$ which dictates the dynamics of the Universe. For our flat ΛCDM model, we have $E(z) = \sqrt{\Omega_M(1+z)^3 + \Omega_\Lambda}$. With these definitions, we can state that

$$\rho \propto \Delta\ E^2(z)\ . \tag{1.17}$$

Hence, equation 1.15 can be written as

$$L_X \propto M_\Delta^{4/3}\ E(z)^{7/3}\,, \tag{1.18}$$

where we define M_Δ as the mass within a radius R_Δ in which the density is Δ times the critical density.

The introduction of $E(z)$ however, is a natural consequence of the expansion of the Universe - since ρ_c is expected to increase with redshift (equation 1.16), for a constant cluster density, the R_Δ will decrease and so will the definition of the cluster's mass M_Δ (Kravtsov and Borgani, 2012). Hence, this term is not responsible for describing the temporal evolution of a scale relation, even though it depends on the redshift.

Finally, equation 1.18 defines the scale correlation between $M_\Delta - L_X$ predicted by the self-similar model, which follows the power law as expected. Such correlation is more commonly worked in logarithmic space as a linear relation, with the slope and intersection parameters β and α, respectively, and the parameter γ for the temporal evolution parameterized by *redshift*:

$$\log[L_X E(z)^{-1}] = \alpha + \beta \log[M_\Delta E(z)] + \gamma \log(1+z) \tag{1.19}$$

Considering equation 1.18, self-similarity predicts $\beta = 4/3$ and does not account for temporal evolution, i.e. $\gamma = 0$.

1.3.1 Deviations from self-similarity

Nonetheless, the purely gravitational model developed by Kaiser applies to situations in which the temperature and density profiles as a function of dimensionless radii are entirely independent of mass, which is not rigorously observed (Navarro et al., 1997; Nagai, 2006; Ascasibar et al., 2006). The self-similar scaling for $L_X - T$, for example, has been called into question by numerous works, that found a generally steeper slope (~ 3) than the one predicted by the model (Markevitch, 1998; Pratt et al., 2009; Maughan et al., 2012).

Indeed, studies of cooling flow in cool-core clusters have given a strong indication of the presence of feedback mechanisms in galaxy clusters, where their impact on the gas dynamics of the ICM may invalidate HE assumptions and thus cause deviations in self-similar predictions (see section 1.4.2 for more details).

Furthermore, X-ray observations also indicate a possible increase in the gas mass fraction with total mass, such that the lower mass systems would present a smaller amount of gas and, consequently, a lower X-ray emission (Pratt et al., 2009; Planelles et al., 2015). This observational trend could result in the steepening of the $M - L_X$ slope, especially when the analysis includes the galaxy group regime.

Various works have calibrated the $M - L_X$ relation using hydrostatic mass estimates (i.e. through X-ray or Sunyaev–Zel'dovich observations) and the consensus seems to be a steeper value for the slope when compared with the self-similar prediction. Chen et al. (2007a) analyzed 106 local groups and clusters from ROSAT and ASCA observations, and Eckmiller et al. (2011) complemented this work by adding 26 other galaxy groups with *Chandra* observations. Both papers found larger slope values than expected by the model ($\beta_{self} = 1.3$). Similar works that focused on the high mass regime also agree with this steepening, such as the ones from Bulbul et al. (2019) and Lovisari et al. (2020). The two works use cluster samples from *XMM-Newton* observations and mass estimates through

SZ signal, and adopt a more robust statistical analysis (e.g. Eddington bias correction). They also investigated the redshift evolution of the scaling relation and found it consistent with the self-similar prediction.

In contrast, works utilizing weak gravitational lensing masses to calibrate the $M - L_X$ relation are not as common. Even so, they tend to agree with the other studies (e.g Leauthaud et al., 2010; Hoekstra et al., 2015), with perhaps the exception of the work done in Kettula et al. (2015). The latter analyzed the galaxy groups of Leauthaud et al. (2010) and clusters from Hoekstra et al. (2015), adding 12 new clusters with CFHTLS observations. They found a slope value of $\beta = 1.27^{+0.16}_{-0.15}$ which is consistent with the self-similar prediction. We note that, although these studies do not detail any selection function or adopt a statistical analysis to account for the different samples used, we compare our results with theirs in section 4. This is especially relevant as we also utilize weak lensing masses in our analysis.

Contrary to the self-similar theory, when accounting for non-gravitational contributions to the gas, deviations from equilibrium and spherical symmetry are expected to be mass-dependent. Consequently, so are the cluster's thermodynamic profiles (e.g. T, P, n_e). Braspenning et al. (2023) used hydrodynamic FLAMINGO simulations of galaxy clusters to analyze X-ray properties, revealing strong mass dependence of these profiles. While not the primary focus of the paper, they also examined the mass-luminosity relation for their simulation, depicted in figure 1.3. Notably, a mild difference in slope is observed between low and high-mass regions, with a steeper slope in the galaxy group regime. Additionally, they reported no significant redshift evolution in the $M - L_X$ relation.

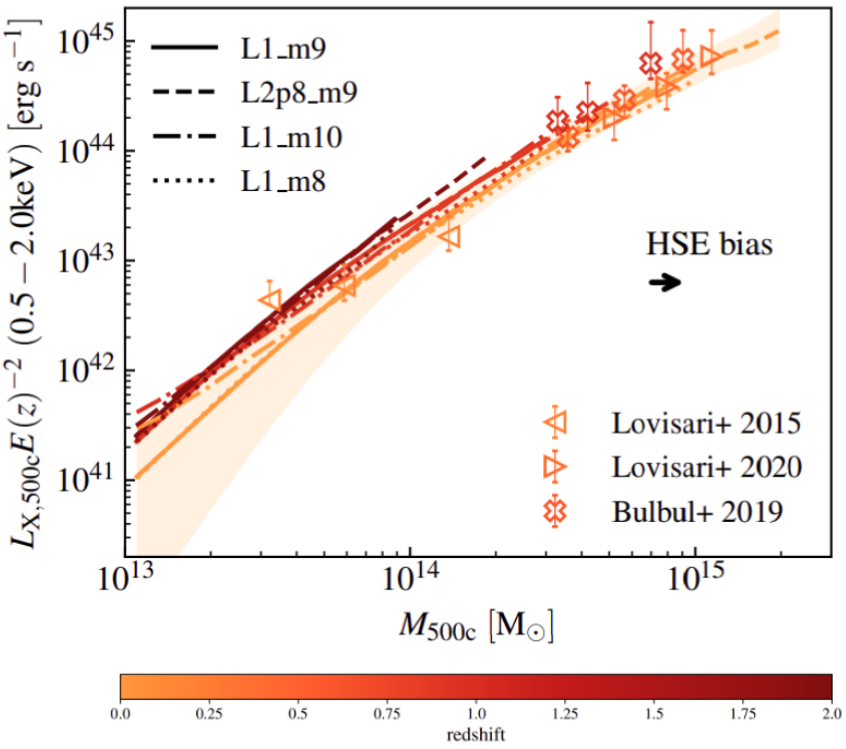


Figure 1.3: X-ray luminosity and halo mass scaling relation. The solid, dashed, dot-dashed, and dotted lines show the results for different box sizes and resolutions, and the colormap represents the redshift. HSE bias indicates the systematic shift applied to the observational data with hydrostatic equilibrium inferred masses to correct for the hydrostatic mass bias. Modified figure from Braspenning et al. (2023).

With that being said, efforts to obtain a consistent analysis of the $M - L_X$ scaling relation are still ongoing. Applying efficient cluster detection methods through combinations of different techniques, and robust procedures for estimating the cluster mass, as well as corrections for the appropriate statistical bias, is a promising goal.

In the following sections, we present a brief review of the main properties of galaxy clusters, highlighting essential features for this work, such as the physics behind the strong X-ray emission of the ICM, and the gravitational weak lensing method used to estimate the halo masses.

1.4 X-ray emission

The first detections of X-ray emission in galaxy clusters occurred in the 1970s, initially by balloon detectors, whose main targets were the Virgo, Perseus, and Coma clusters. The observations indicated a spatially extensive and non-varying X-ray source, such that it could not be associated with a particular AGN and therefore it must have occurred in

the thermal gas of the ICM. Shortly afterward, there was a boost in the studies of X-ray emission in galaxy clusters with the launch of the satellite *Uhuru*, which allowed extensive observations and complete surveys in X-rays (Sarazin, 1988).

In the 1980s, it was already possible to image clusters of galaxies in X-rays, and a decade later, spatially resolved spectroscopy was also performed. These and other advances in the area of X-ray emission from galaxy clusters became feasible with the launch of satellites such as *Einstein* (1978 - 1981), ROSAT (1990 - 1999), *Chandra* (1999 -), XMM-*Newton* (1999 -), among others (Sarazin, 1988; Lima Neto, 2020).

In the next section, we present an overview of the physics behind the X-ray emission detected in the ICM. The mathematical approach was mainly developed as in (Rybicki G., 1985).

1.4.1 Thermal bremsstrahlung

The X-ray emission observed in galaxy clusters is caused by a process known as thermal *Bremsstrahlung* or *free-free* emission, that occurs due to the acceleration of a charge in an electron-ion collision. Since both particles have essentially the same charge, and we know that the relative acceleration is inversely proportional to the mass, we can consider the electrons as primary radiators moving through the ion's electromagnetic field (Rybicki G., 1985). Generally speaking, when an electron passes through the ion, it slows down, changing its trajectory and losing energy. The energy variation caused by this encounter is then emitted as an X-ray photon.

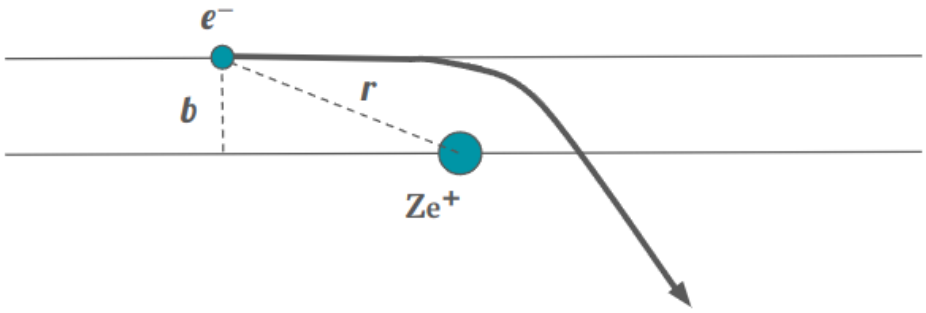


Figure 1.4: Esquematization of an electron of charge e^- passing through an ion of charge Ze^+, with impact parameter b.

If we consider a *small angle scattering* regime, we can analyze the emission by considering a single electron with charge e^- and velocity $\boldsymbol{v}$, passing an ion of charge Ze^+ with impact parameter b, as illustrated in figure 1.4.

1.4.1.1 Dipole aproximation

If we assume that the electron-ion collision is taking place at a 'length' L and within a time scale τ such that $\tau >> \Delta t = L/c$, i.e the particles are not relativistic, we can use the dipole approximation to calculate the *Bremsstrahlung* emission. In this case, the dipole moment and its second derivative is given by

$$\begin{aligned} \boldsymbol{d} &= -e\boldsymbol{r} \\ \ddot{\boldsymbol{d}} &= -e\dot{\boldsymbol{v}}\,, \end{aligned} \tag{1.20}$$

and the Fourier transform of $\ddot{\boldsymbol{d}}$ in terms of the frequency ω, is given by

$$\hat{\ddot{\boldsymbol{d}}}(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \ddot{\boldsymbol{d}}(t)e^{i\omega t}dt = -\frac{e}{2\pi}\int_{-\infty}^{\infty} \dot{\boldsymbol{v}}e^{i\omega t}dt. \tag{1.21}$$

Knowing that

$$\hat{\ddot{\boldsymbol{d}}}(\omega) = -\omega^2\hat{\boldsymbol{d}}(\omega)\,, \tag{1.22}$$

we see that the transform of the dipole moment is

$$\hat{\boldsymbol{d}}(\omega) = \frac{e}{2\pi\omega^2}\int_{-\infty}^{\infty} \dot{\boldsymbol{v}} e^{i\omega t} dt. \tag{1.23}$$

The integral above is meaningful only during the collision time τ, which is in order of $\tau \sim b/v$, and we can evaluate the integral in two limits: When $\omega\tau >> 1$ the exponent oscillates rapidly, hence the integral's contribution can be negligible. On the other hand, when $\omega\tau << 1$ the exponent is approximately zero, and the exponential term is ~ 1. With that being said, we can write

$$\hat{\boldsymbol{d}}(\omega) = \begin{cases} \frac{e}{2\pi\omega^2}\Delta\boldsymbol{v}, & \omega\tau << 1 \\ 0, & \omega\tau >> 1 \end{cases} \tag{1.24}$$

where $\Delta\boldsymbol{v}$ is the change in velocity during the collision.

We can also write the part of the electromagnetic field that describes the radiation in terms of the dipole (Rybicki G., 1985):

$$\boldsymbol{E}_{rad}(t) = \frac{\boldsymbol{n}\times(\boldsymbol{n}\times\ddot{\boldsymbol{d}})}{c^2 R} = \frac{\ddot{\boldsymbol{d}}(t)}{c^2 R}\sin\theta\ , \tag{1.25}$$

where $\boldsymbol{n}$ is a unit vector in the direction of motion, and R is the radial distance from the particle's trajectory. If we take the Fourier transform of $\ddot{\boldsymbol{d}}(t)$ and use 1.22, we can express the equation above in terms of the frequency

$$\hat{\boldsymbol{E}}_{rad}(\omega) = -\frac{\omega^2\hat{\boldsymbol{d}}(\omega)}{c^2 R}\sin\theta\ . \tag{1.26}$$

Knowing that the radiation spectrum (e.i. energy per frequency per unit area) is given by

$$\frac{dW}{d\omega dA} = c|\hat{\boldsymbol{E}}_{rad}(\omega)|^2\ , \tag{1.27}$$

we can use equation 1.26 to write

$$\frac{dW}{d\omega dA} = \frac{\omega^4}{c^3 R^2}|\hat{\boldsymbol{d}}(\omega)|^2\sin^2\theta\ . \tag{1.28}$$

We must now introduce the solid angle $d\Omega = dA/R^2$ and integrate the expression above:

$$\begin{aligned} \frac{dW}{d\omega} &= \frac{\omega^4}{c^3}|\hat{\boldsymbol{d}}(\omega)|^2 \int \sin^2\theta d\Omega \\ \Rightarrow \frac{dW}{d\omega} &= \frac{8\pi\omega^4}{3c^3}|\hat{\boldsymbol{d}}(\omega)|^2 \,. \end{aligned} \tag{1.29}$$

Using the expression for the dipole found in equation 1.24, we can then obtain the total energy per frequency range emitted by this electron:

$$\frac{dW}{d\omega} = \begin{cases} \frac{2e^2}{3\pi c^3}|\Delta\boldsymbol{v}|^2, & \omega\tau << 1 \\ 0, & \omega\tau >> 1 \end{cases} \tag{1.30}$$

The change in velocity $\Delta\boldsymbol{v}$ can be determined by integrating the acceleration normal to the electron's trajectory $a_\perp$, since the tangential counterpart of the velocity is constant (see figure 1.4 for an illustration of the system):

$$\Delta\boldsymbol{v} = \int a_\perp dt \,;$$

$$a_\perp = \frac{F}{m} = \frac{Ze^2}{r^2}\frac{1}{m_e}\frac{b}{r}; \;\; r^2 = (b^2 + v^2t^2) \tag{1.31}$$

$$\Rightarrow \Delta\boldsymbol{v} = \frac{Ze^2}{m_e}\int \frac{b\,dt}{(b^2+v^2t^2)^{3/2}} = \frac{2Ze^2}{mbv}$$

At last, combining these equations, we find that the emission of a single electron in collision with one ion can be expressed as:

$$\frac{dW(b)}{d\omega} = \begin{cases} \frac{8Z^2e^2}{3\pi c^3 m_e^2 v^2 b^2}, & b << v/\omega \\ 0, & b >> v/\omega >> 1 \,. \end{cases} \tag{1.32}$$

1.4.1.2 Emission in the ICM

As mentioned before, the ICM is a hot, rarefied, ionized plasma, and therefore, it presents an abundant amount of ions and free electrons; an ideal environment for *Bremsstrah-*

lung radiation. Hence, to gain a general comprehension of the X-ray luminosity observed in these systems, we need to find the expression for the *Bremsstrahlung* emission from a gas of n_e electrons and n_i ions per cm^3.

In this case, the flux of electrons passing through a single ion in an element of area $dA = 2\pi b\, db$, and the consequent power emitted, are:

$$\begin{aligned} \frac{dN}{dtdb} db &= v n_e dA = v n_e 2\pi b\, db\,, \\ \frac{dW}{d\omega dtdb} &= \frac{dN}{dtdb}\frac{dW}{d\omega} = n_e \frac{16 Z^2 e^6}{3 m_e^2 c^3 b v}, \end{aligned} \tag{1.33}$$

where we used the expression for the electromagnetic emission found in 1.32.

In order to calculate the emission for all electrons and ions in the gas, we must now integrate the expression above over the impact parameters and multiply it by the ion number density n_i (Rybicki G., 1985):

$$\frac{dW}{d\omega dtdV} = \frac{16\pi e^6}{3\sqrt{3} m_e^2 c^3 v} n_e n_i Z^2 g(v,\omega)\,. \tag{1.34}$$

In the above equation, the term $g(v,\omega)$ is called the Gaunt factor, which accounts for the integral over the impact parameters normalized to ~ 1 and considers quantum effects.

Due to the ICM temperature ($T \sim 10^7 - 10^8\ K$) it can be assumed that the electrons are non-relativistic and exhibit a Maxwell velocity distribution. With these considerations, and using $d\omega = 2\pi d\nu$, we have that the *Bremsstrahlung* emissivity (i.e. power per frequency, per volume) that this population of electrons radiates is given by (Rybicki G., 1985):

$$\epsilon_\nu = \frac{dW}{d\nu dV dt} = 6,842 \cdot 10^{-38}\, Z^2 n_e n_i T^{-1/2} e^{-h\nu/kT} g(E,T)\ erg\, s^{-1} cm^{-3} Hz^{-1}\,, \tag{1.35}$$

where T is the temperature of the gas, and the Gaunt factor $g(E,T)$ is responsible for the quantum corrections of *Bremsstrahlung* scattering.

Integrating over all frequencies, we obtain an emissivity of the entire X-ray range:

$$\epsilon = \frac{dW}{dV dt} = 1,426 \cdot 10^{-27}\, Z^2 n_e n_i T^{1/2} g(E,T)\ erg\, s^{-1} cm^{-3}\,. \tag{1.36}$$

Taking typical values for the ICM (i.e. $Z^2 = 1,4$ ($\sim 24\%$ of He), $n_e \sim n_i = 10^{-2}\, cm^{-3}$, $T \sim 10^7\, K$, and $g(E,T) \sim 1,2$) and assuming a cluster size of $\sim 1\, Mpc$, which corresponds to a volume of $\sim 10^{72}\, cm^3$, one can estimate the X-ray luminosity to be around

$$L_x = \frac{dW}{dV\,dt} dV \sim 10^{44}\, erg\, s^{-1}\,. \tag{1.37}$$

Indeed, galaxy clusters are extremely luminous X-ray sources, with luminosities ranging from $L_x \sim 10^{43} - 10^{45}\, erg\, s^{-1}$. In fact, these clusters are the most abundant and brightest X-ray sources besides AGN - of the extragalactic X-ray sources, around 85% are AGNs and the remaining 15% are galaxy clusters (Schneider, 2015).

X-ray emission from galaxy clusters is observed in the [0.2 - 10.0] keV band; for energies below this range, the emission is strongly absorbed by neutral hydrogen present in the Milky Way. The upper limit is determined by the exponential term in the expression for the *Bremsstrahlung* emissivity, equation 1.35, where for energies greater than kT the emissivity decays exponentially.

As we can see from equation 1.36, the emissivity strongly depends on the particle density (since we usually assume $n_e \sim n_i = n$, then $\epsilon \propto n^2$) so that when we have a superficial X-ray brightness profile (integral of the emissivity along the line-of-sight), we can also obtain information of the cluster's gas density map (Mo et al., 2020). This means that the complex and often dynamic structure of the intracluster medium can be approached through X-ray measurements, making it possible to infer the occurrence of mergers, relaxation, and feedback processes during the cluster's evolution. As an example, figure 1.5 shows an X-ray image of the much-studied Coma cluster, taken with ROSAT; even though Coma is usually considered a regular cluster, we can observe a small secondary structure in the lower right region, which indicates that the system is still dynamically evolving (Briel et al., 2001).

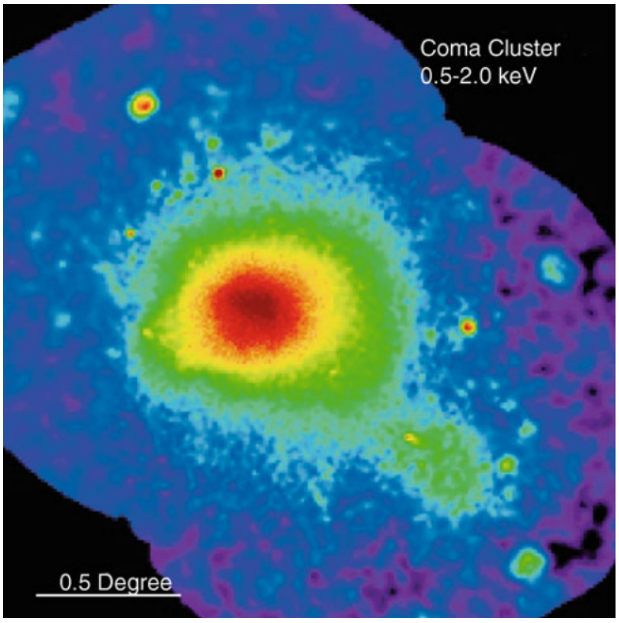


Figure 1.5: X-ray image of the Coma cluster, taken with the ROSAT-Position Sensitive Proportional Counter (PSPC) instrument. Figure taken from Schneider (2015).

1.4.2 Cooling flow and AGN feedback

With galaxy clusters being such strong X-ray emitters, it is only natural that one assumes the ICM gas will start to lose internal energy and cool down, so we can no longer expect the hydrostatic equilibrium (HE) to hold indefinitely (Cowie and Binney, 1977; Fabian and Nulsen, 1977). With that being said, the time it takes for the gas to cool can be estimated as shown in equation 1.38.

$$t_{cool} = \frac{E}{dE/dt} \sim 8.5 \times 10^{10} yr \left(\frac{n_e}{10^{-3} cm^{-3}}\right)^{-1} \left(\frac{T_g}{10^8 K}\right)^{1/2} , \tag{1.38}$$

where $E = (3/2)nkT_g$ is the internal energy for an ideal gas, and $dE/dt = \epsilon$ is the loss of energy due to the thermal *Bremsstrahlung* emission.

The cooling time is usually long for most regions in the cluster, where the HE can be used as a good approximation. However, in the most central regions, the density (n_e) can be high enough to result in a cooling time shorter than the cluster age [1]($t_{cool} < H_0^{-1}$), which would imply a rapid cooling of the gas in this region and a mass inflow to re-establish pressure equilibrium. Hence, we would expect to observe a considerable amount of gas with temperature $kT \leq 1keV$ in the center of clusters, a major increase in star

[1] The galaxy cluster age can be approximated by the age of the Universe (i.e the Hubble time H_0^{-1}) where the comparison is in order of magnitudes rather than precise values.

formation as the gas cools down, and a strong increase in X-ray emission (since $\epsilon \propto n_e^2$). Nevertheless, such high cooling rates predicted by the models are not observed. In fact, the observations indicate that the temperature of the central gas is maintained at around 1/3 of the equilibrium temperature (Peterson et al., 2003), which suggests the existence of a mechanism that injects energy back into the gas, reheating it. These are called *feedback* processes and the main candidates are the powerful energetic jets in AGNs (Lima Neto, 2020; Schneider, 2015).

1.4.2.1 Cool-core and non-cool-core clusters

Even though the dynamic between cooling flow and feedback is not yet fully understood, it is in general agreement that the so-called *cool-core* (CC) clusters were the ones that reached a sufficiently relaxed state to allow a cool and compact central region to develop, and so, they are more suitable for HE approximations in the outer regions. As for the *non-cool-core* (NCC) clusters, the absence of a cooling flow suggests that the gas is far from the equilibrium state, most likely caused by turbulent motions and/or shock fronts from recent mergers (Gianfagna et al., 2021; Ansarifard et al., 2020).

The main observational difference between CC and NCC is whether or not there is a central peak in the cluster's surface X-ray brightness profile. Moreover, CC clusters are usually the most mature and massive clusters, and studies suggested a strong evolution in the occurrence of cooling flow - the CC fraction decreases with redshift - most likely due to the higher merger rate in younger clusters (Vikhlinin et al., 2006; Chen et al., 2007b). Although the presence of a CC is usually associated with more massive clusters, this trend could lead to diversions from the self-similar scaling relations.

1.4.2.2 AGN feedback in the ICM

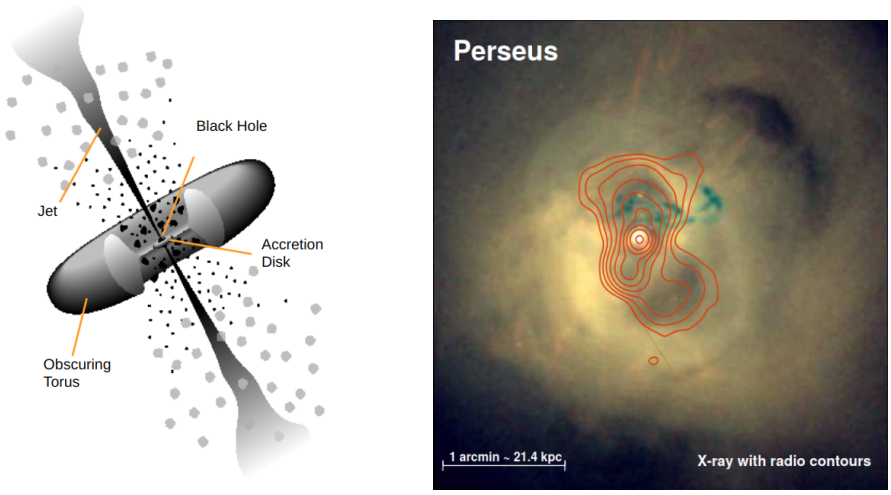
According to the unified model for Active Galactic Nuclei (AGN), these structures are essentially composed of a central supermassive black hole, an accretion disc, a surrounding dust torus, and powerful radio-emitting jets, as illustrated in figure 1.6. Although AGNs can present many facets and complexities, these objects are generally divided into two main categories; the *radiative-mode* AGN (or *radio-quiet* AGN) and the *jet-mode* AGN (or *radio-loud* AGN), which are characterized mainly by whether or not the mass accretion to the black hole is efficient and by the intensity of the jets' radio emission. While most

radiative-mode AGNs are associated with young, star-forming galaxies, the jet-mode AGN hosts are thought to be the most massive, red galaxies, usually located at the center of galaxy clusters (i.e. BCGs) (Fabian et al., 2006; Best et al., 2014). The so-called AGN feedback process is the complex interaction of the energy generated by the accretion onto the black hole and the system's surrounding gas (Fabian, 2012).

Although there is much direct evidence that radio-quiet AGNs drive wind outflows - detected by blue-shifted absorption lines in the optical, ultraviolet, and X-ray spectra, that trace the ionized gas with outflow velocities - there are still uncertainties as to its feedback effects (Veilleux et al., 2005; Fabian, 2012). The strong absorption lines measured are almost always saturated and the outflows' sizes have been largely unconstrained (Heckman and Best, 2014). We can estimate, however, that the heating source of the intercluster gas is most probably the powerful jet outflows from radio-AGNs since the presence of such activity has been detected in more than 70% of clusters, reaching nearly 100% in those with evidence of cooling flow (Burns, 1990; Heckman and Best, 2014).

Without delving into specifics, the general scenario for AGN feedback in galaxy clusters is a radio-AGN fueled by the surrounding gas that cooled out (e.g. during the cooling flow process), and the accretion energy being harnessed to generate powerful radio jets. These jets produce expanding radio lobes of relativistic plasma that inflate bubbles (i.e. cavities) in the ionized intracluster gas, as illustrated in figure 1.6. Because radio jets are highly anisotropic and the gas cooling must be balanced in all directions, it follows that it must be these bubbles, rather than direct jet interaction, that facilitate the energy transfer to the ICM gas without much disturbance. In this configuration, injecting energy in the ICM causes the gas to expend, reducing its density much more significantly than raising its temperature, and thus *reducing* the observed X-ray luminosity (Fabian, 2012).

In powerful radio galaxies, lobe energy estimates (via radio and X-ray observations) range from $\sim 10^{60}$ to 10^{62} erg; given that the thermal content of gas in a galaxy group of mass $\sim 5 \cdot 10^{13} M_{\odot}$ is roughly $\sim 10^{61}$ erg, one can conclude that AGN feedback must play a vital role in the evolution of these groups (Erlund et al., 2006; Fabian, 2012). Indeed, estimation of the energy stored at the ICM gas cavities suggests that the time-averaged heating-to-cooling ratio decreases by about one order of magnitude when going from small groups to massive clusters. Hence, the AGN heating rate must balance the cooling flow in the massive CC clusters, while in smaller systems more energy is being injected than



(a) (b)

Figure 1.6: (a) Esquematization of the main components of AGN according to the Unified Model. Modified figure from Padovani (1999). *(b)* Jet-mode AGN feedback in the Perseus cluster. The color-scaling shows the X-ray emission and the contours indicate the radio emission. One can observe the presence of cavities in the X-ray gas, where the radio emission lies. Figure from Heckman and Best (2014) and data from Fabian et al. (2006).

required (Heckman and Best, 2014).

As for the number density of AGNs in different epochs, it is usually probed by the history of the cumulative growth mass in massive black holes. In this sense, the abundance of the most luminous AGNs, associated with the most massive black holes, peaked at redshifts $z \sim 2-3$, known as the quasar era. Best et al. (2014) also showed that the radiative-mode AGN population increases with redshift, as the availability of cold gas increases, while the jet-mode AGN population decreases with redshift - since this population is associated with halos of hot gas that are cooling, an evolution similar to that of massive galaxies. This is consistent with what is observed locally; where jet-mode AGNs dominate, except for high luminosities. However, the evolution of the relation between AGN feedback and cooling flow in galaxy clusters is still very limited, especially because detecting AGN-driven cavities and identifying the cooling properties of clusters are done by X-ray observations, and therefore suffer from the strong redshift dimming of the surface profile, making measurements at high redshift extremely challenging (Ruppin et al., 2023).

These considerations indicate that the energetic feedback from AGNs is expected to bias the $M - L_X$ relation from the predicted by self-similarity. Furthermore, we have seen that these processes tend to play a more important role in lower mass systems, suggesting that deviations from the model would be more prominent for these groups (Nagai et al.,

2007; McCarthy et al., 2010; Planelles et al., 2015).

1.5 Weak gravitational lensing

The bending of light by gravity was first conjectured by Isaac Newton in 1704, but it was only later that the astronomer Johann Georg Von Soldner developed a quantified approach to the subject. In 1804, Soldner calculated the deflection of light due to its passing through the Sun, based on Newton's law of gravity, and obtained a value close to $0.9''$ (Narayan and Bartelmann, 1996). It was more than a century later that Einstein used the recently developed field equations of General Relativity (GR) to estimate the deflection angle of a photon grazing the Sun's surface, and found it to be twice the value previously predicted. According to GR, the deflection angle was $1''.7$, and the increase by a factor of two was due to the curvature of the metric, i.e. the Sun actually bends the four-dimensional space-time (Narayan and Bartelmann, 1996; Meneghetti, 2022).

The value predicted by Einstein was confirmed in an expedition led by Eddington, Dyson, and Davidson (1919) that measured the apparent angular shift of stars during an eclipse at Sobral, Brazil, and at the island of Principe, in the Gulf of Guinea. This observation marks the history of science, as it is considered the first compelling evidence in support of Einstein's GR theory (Narayan and Bartelmann, 1996; Meneghetti, 2022).

It was Zwicky (1937) that first mentioned the applications of gravitational lensing, by predicting that this phenomenon could magnify faint distant galaxies which would otherwise be impossible to detect. This means that lenses could act as 'cosmic telescopes', making it possible to determine source properties below the resolution and sensitivity limits of current observations. For many years, however, no meaningful applications of gravitational lensing were made, until the discovery of the lensed quasar QSO 0957+561A,B by Walsh, Carswell, & Weymann (1979). From there on, the field of gravitational lens only grew, and soon reached the first observations of strong lensed sources by galaxy clusters (Lynds and Petrosian, 1989), i.e. strongly distorted and elongated background galaxies perceived as luminous arcs in the foreground cluster. Later on, it was also discovered weak coherently distorted images of faint background galaxies in cluster regions, and the usage of these observations to reconstruct superficial mass profiles of the galaxy clusters (Narayan and Bartelmann, 1996). Nowadays, gravitational lensing has developed into an

essential tool for studies in astronomy.

In the following sections, we present an overview of the gravitational lensing formalism, focusing on the weak regime, which will be the case for this project. We follow the approach presented mainly by Meneghetti (2022); Narayan and Bartelmann (1996) and Wambsganss (1998).

1.5.1 Refractive index of a gravitational field

The name gravitational lensing highlights the similarities present in this phenomenon with the well-known physics area of optical lensing. In fact, considering a massive object located between the source and the observer, the gravitational field will deflect the light beams emitted by the source, as illustrated in figure 1.7. This deflection occurs so that the image perceived by the observer will be displaced, distorted and magnified, a result similar to that observed when a beam of light passes through a lens. It is possible to take this analogy further and consider the deflection of light in a curved space-time as the effect of refraction in a flat space-time, by adopting a model of a lens with a refraction index $n = c/c'$, where c' is the light speed in the medium.

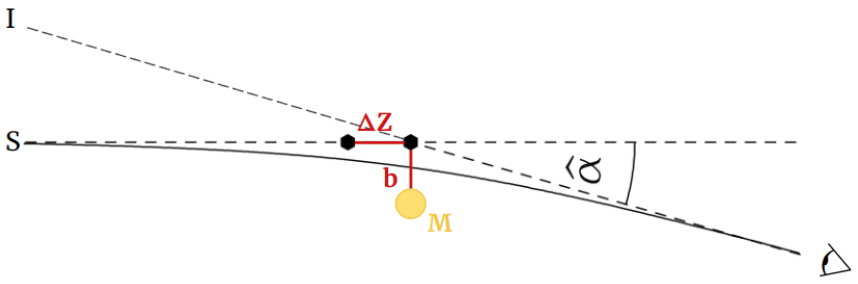


Figure 1.7: Esquematization of a light beam from a source S that suffers a deflection due to the presence of an object of mass M. The observer perceives the image I dislocated. Figure taken from Narayan and Bartelmann (1996), modified.

If we assume a spherically symmetric and static gravitational lensing system, the Schwarzschild metric can be used (Meneghetti, 2022), such that the line element in spherical coordinates is given by

$$ds^2 = \left(1 - \frac{2GM}{c^2 R}\right) c^2 dt^2 - \left(1 - \frac{2GM}{c^2 R}\right)^{-1} dR^2 - R^2(d\theta^2 + \sin^2\theta d\phi^2). \tag{1.39}$$

Introducing a new radial coordinate r in a flat space, that relates to the metric's radial coordinate R through

$$\Rightarrow R = r\left(1 + \frac{GM}{2rc^2}\right)^2, \tag{1.40}$$

and the cartesian coordinates:

$$\begin{aligned} x &= r sin\theta cos\theta \\ y &= r sin\theta sin\phi \\ z &= r cos\theta \\ dl^2 &= (dx^2 + dy^2 + dz^2) \end{aligned} \tag{1.41}$$

it is possible to rewrite the Schwarzschild metric (equation 1.39) as

$$ds^2 = \left(\frac{1 - \frac{GM}{2c^2 r}}{1 + \frac{GM}{2c^2 r}}\right) c^2 dt^2 - \left(1 + \frac{GM}{2c^2 r}\right)^4 (dx^2 + dy^2 + dz^2). \tag{1.42}$$

In the weak field limit, the gravitational potential $\Phi = -GM/r$ is much smaller than c^2, so we have $-GM/rc^2 << 1$. This is the case for most astrophysical systems, such as galaxy clusters, for which the Newtonian gravitational potential is 10^4 times smaller than c^2 (Narayan and Bartelmann, 1996). In this scenario, and using a fist order binomial approximation[2], we get

$$ds^2 = \left(1 + \frac{2\Phi}{c^2}\right) c^2 dt^2 - \left(1 - \frac{2\phi}{c^2}\right)\ dl^2. \tag{1.43}$$

We also know that a beam of light will always travel along a null geodesic, i.e. the metric must satisfy $ds^2 = 0$. In the case of a locally flat space, the path described by the beam is simply a straight line, whereas if the spacetime is deformed by a gravitational field, such as in the presence of a cluster of galaxies, the trajectory corresponding to light-like geodesics

[2] $(1+\epsilon)^n = 1 + n\epsilon,\ \epsilon << 1$

would be curved. Therefore, according to the equation 1.43, the line element on which light propagates satisfy

$$\left(1+\frac{2\Phi}{c^2}\right)c^2dt^2=\left(1-\frac{2\phi}{c^2}\right)\,dl^2. \tag{1.44}$$

The light speed in this medium is then given by

$$c'=\frac{|dl|}{dt}=c\sqrt{\frac{1+\frac{2\Phi}{c^2}}{1-\frac{2\Phi}{c^2}}}\sim c\left(1+\frac{2\Phi}{c^2}\right). \tag{1.45}$$

With the above equation, we can now infer the refractive index of the lens in a weak gravitational field:

$$n=\frac{c}{c'}=\left(1+\frac{2\Phi}{c^2}\right)^{-1}\sim 1-\frac{2\Phi}{c^2}\,. \tag{1.46}$$

Knowing the refractive index, we can find an expression for the deflection angle α of the light by applying Fermat's principle, which says that the path of light between two points is the one that takes the least time. This is equivalent to minimizing the expression for the light travel time, and it is done by taking its derivative and setting it equal to zero:

$$\begin{aligned} t&=\int n(\vec{x}(l))dl\,;\\ \Rightarrow \delta&\int n(\vec{x}(l))dl=0\,. \end{aligned} \tag{1.47}$$

After some mathematical manipulation, one can find the expression for the deflection angle:

$$\hat{\alpha}=\frac{2}{c^2}\int\nabla_{\perp}\Phi dz. \tag{1.48}$$

Consider for example a point mass lens, as illustrated by figure 1.7, where the smallest distance between the beam of light and the lens is denoted by b, and the light curve parameter is z. The lens's gravitational potential is then

$$\Phi(b,z)=-\frac{GM}{(b^2+z^2)^{1/2}}\,, \tag{1.49}$$

and the derivative perpendicular to the light path is

$$\nabla_{\perp}\Phi(b,z) = \frac{GM}{(b^2+z^2)^{3/2}}\,\vec{\boldsymbol{b}}\,, \tag{1.50}$$

where $\vec{\boldsymbol{b}}$ is the vector orthogonal to $\vec{\boldsymbol{z}}$ and points inwards to the lens.

Therefore, equation 1.48 gives the deflection angle

$$\hat{\alpha} = \frac{2}{c^2}\int_{-\infty}^{+\infty}\nabla_{\perp}\Phi dz = \frac{4GM}{c^2 b}\,. \tag{1.51}$$

Knowing the Sun's mass to be $M_{\odot} = 1.98\times10^{30}$ kg and taking the impact parameter to be the Sun's radius $b = R_{\odot} = 6.96\times10^{5}$ km, we can retrieve the result obtained by Einstein for the deflection suffered by a photon grazing the surface of our star:

$$\hat{\alpha} = \frac{4GM_{\odot}}{c^2 R_{\odot}} = 0.84\times10^{-5}\ radians = 1.7\ arcsec\,. \tag{1.52}$$

1.5.2 The thin screen approximation and the lens equation

With the expression for the deflection angle in hand, we can now analyze the case for an extended mass distribution. It is known that most of the light deflection occurs in the region $z\sim\pm b$, such that the lens can be considered very thin compared to the entire optical path, so we can consider the distribution of mass to be projected onto a plane, called the *lens plane*, as illustrated in figure 1.8.

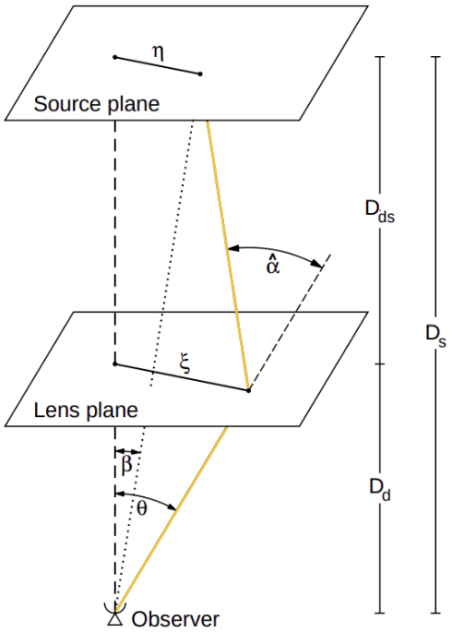


Figure 1.8: Esquematization of a gravitational lens configuration. The light ray propagates from the source S, located at a distance η from the optic axis, passing the lens at transverse distance ξ, and being deflected by an angle α. The *lens plane* is also illustrated. The angular separation of the source and the image as seen by the observer is denoted as β and θ, respectively. Figure taken from Bartelmann and Schneider (2001), modified.

In this configuration, the surface mass density of the lens, as a function of the two-dimensional vector $\vec{\xi}$ is given by

$$\Sigma(\vec{\xi}) = \int \rho(\vec{\xi}, z) dz \,. \tag{1.53}$$

In this approximation, the deflection angle at a radius $\vec{\xi}$ is that caused by the sum of mass elemets $\Sigma(\vec{\xi})d^2\xi$, so that equation 1.49 give us

$$\hat{\alpha} = \frac{4G}{c^2} \int_{-\infty}^{+\infty} \frac{(\vec{\xi} - \vec{\xi}')\Sigma(\vec{\xi})}{|\vec{\xi} - \vec{\xi}'|^2} d^2\xi \,. \tag{1.54}$$

The typical geometry of a gravitational lens system is illustrated in figure 1.8, where we can analyze the relation between the intrinsic and apparent positions of the source. In this scenario, the light ray is emitted by a source in a certain position $\vec{\beta}$ in the sky and is deflected by the lens so that the observer perceives it from a different direction $\vec{\theta}$. As we can see from figure 1.8, the relation between $\vec{\alpha}$, $\vec{\beta}$ and $\vec{\theta}$ is dependent on the angular

diameter distances of the lens D_L, the source D_S and between the lens and the source D_{LS}. We then see that the relation $\vec{\theta}D_S = \vec{\beta}D_S + \vec{\alpha}D_{LS}$ holds, and since in the majority of gravitational lens events the angles are relatively small (e.i. θ, β, $\alpha << 1$) we can simplify the relation to

$$\vec{\beta} = \vec{\theta} - \vec{\alpha}(\vec{\theta})\,, \tag{1.55}$$

where we introduced the reduce deflection angle $\vec{\alpha} = (D_{LS}/D_S)\hat{\vec{\alpha}}$. The expression 1.55 is known as the *lens equation*.

To further study the properties of gravitational lensing, it is convenient to use the Newtonian potential in three dimensions in the lens plane. The so-called *lensing potential* for an extended matter distribution is given by

$$\hat{\Psi}(\vec{\theta}) = \frac{D_{LS}}{D_S D_L}\frac{2}{c^2}\int \Phi(D_L, \theta, z)\, dz\,, \tag{1.56}$$

and can be related to the deflection angle by its gradient

$$\vec{\nabla}_\theta \hat{\Psi}(\vec{\theta}) = \vec{\alpha}(\vec{\theta})\,. \tag{1.57}$$

We also introduce the *convergence* $\boldsymbol{\kappa}$, that relates to the lensing potential by

$$\vec{\nabla}^2_\theta \hat{\Psi}(\vec{\theta}) = 2\boldsymbol{\kappa}(\vec{\theta})\,. \tag{1.58}$$

As we will see throughout this section, the convergence will represent the isotropic transformations in the source image and is defined as the dimensionless surface mass density

$$\boldsymbol{\kappa}(\vec{\theta}) = \frac{\Sigma(\vec{\theta})}{\Sigma_{crit}}\,, \tag{1.59}$$

where Σ_{crit} is the *critical surface density*:

$$\Sigma_{crit} = \frac{c^2}{4\pi G}\frac{D_S}{D_L D_{LS}}\,. \tag{1.60}$$

In principle, to determine the image's distortions caused by the gravitational lens, it would

be enough to solve the *lens equation* for each point within the extended source. Especially when the source length is much smaller than the scale in which the deflection angle varies (e.g. weak gravitational regime) the distortions in each point can be linearized.

Let us then consider two points in the lens plane with positions $\vec{\theta}$ and $\vec{\theta}' = \vec{\theta} + d\vec{\theta}$, where the deflection angles are $\vec{\alpha}$ and $\vec{\alpha}' \simeq \vec{\alpha} + d\vec{\alpha}/d\vec{\theta}\, d\vec{\theta}$, respectively. Following the *lens equation* 1.55 we can map these two points onto the source plane as $\vec{\beta} = \vec{\theta} - \vec{\alpha}$ and $\vec{\beta}' = \vec{\beta} + d\vec{\beta} = \vec{\theta}' - \vec{\alpha}'$. Figure 1.9 illustrates this situation.

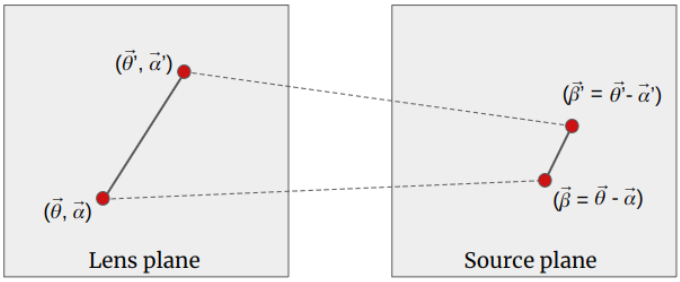


Figure 1.9: Linear mapping of the image's distortions, from the *lens plane* and the *source plane*, assuming a slow variation of the deflection angle.

Through these considerations, the vector $(\vec{\beta}' - \vec{\beta})$ can be writen as

$$(\vec{\beta}' - \vec{\beta}) = \left(I - \frac{d\vec{\alpha}}{d\vec{\theta}}\right)(\vec{\theta}' - \vec{\theta})\,. \tag{1.61}$$

Generalizing to all points in the extended source, we can describe the distortion of the image by the Jacobian matrix

$$A = \frac{\partial\vec{\beta}}{\partial\vec{\theta}} = \left(\delta_{ij} - \frac{\partial\alpha_i(\vec{\theta})}{\partial\theta_j}\right) = \left(\delta_{ij} - \hat{\Psi}_{ij}\right)\,, \tag{1.62}$$

where we used the relation between the deflection angle and the lens potential (equation 1.57) and introduced the notation

$$\frac{\partial^2\hat{\Psi}_i(\vec{\theta})}{\partial\theta_i\partial\theta_j} = \hat{\Psi}_{ij}\,. \tag{1.63}$$

When analyzing only the isotropic part of the Jacobian (the traceless part), we have

$$\left(A - \frac{1}{2} trA \cdot I\right)_{ij} = \begin{pmatrix} -\frac{1}{2}(\hat{\Psi}_{11} - \hat{\Psi}_{22}) & -\hat{\Psi}_{12} \\ -\hat{\Psi}_{12} & \frac{1}{2}(\hat{\Psi}_{11} - \hat{\Psi}_{22}) \end{pmatrix} = - \begin{pmatrix} \gamma_1 & \gamma_2 \\ \gamma_2 & -\gamma_1 \end{pmatrix} = -\Gamma, \quad (1.64)$$

where

$$\begin{aligned} \gamma_1 &= \frac{1}{2}(\hat{\Psi}_{11} - \hat{\Psi}_{22}) \\ \gamma_2 &= \hat{\Psi}_{12}\,, \end{aligned} \quad (1.65)$$

and Γ is called the *shear* tensor.

Knowing that the *shear* tensor is invariant under rotations of $\phi = 2\pi$ (spin-2), and that its engevalues are $\pm\sqrt{\gamma_1^2 + \gamma_2^2} = \pm\gamma$ (Meneghetti, 2022) we have that

$$\Gamma = \gamma \begin{pmatrix} \cos 2\phi & \sin 2\phi \\ \sin 2\phi & -\cos 2\phi \end{pmatrix}. \quad (1.66)$$

We can see from the expressions in 1.65 that the *shear* quantifies the gradient of the gravitational force (tidal field), and its physical manifestation is an asymmetric distortion in the background sources.

In a similar calculation, one finds that the other part of the Jacobian is given by $(1 - \boldsymbol{\kappa})\delta_{ij}$. So that the full Jacobian matrix is

$$A = (1 - \boldsymbol{\kappa}) \cdot I - \gamma \begin{pmatrix} \cos 2\phi & \sin 2\phi \\ \sin 2\phi & -\cos 2\phi \end{pmatrix}. \quad (1.67)$$

A quick analysis of equation 1.67 leads us to infer that the convergence is responsible for the symmetric distortions, where images are merely rescaled by a factor of $1/(1 - \boldsymbol{\kappa})$. On the other hand, the shear acts by stretching the direction in which ϕ was defined by a factor of $+\gamma$, and in the orthogonal direction by $-\gamma$. Figure 1.10 illustrates the effects of these distortions on a circular source.

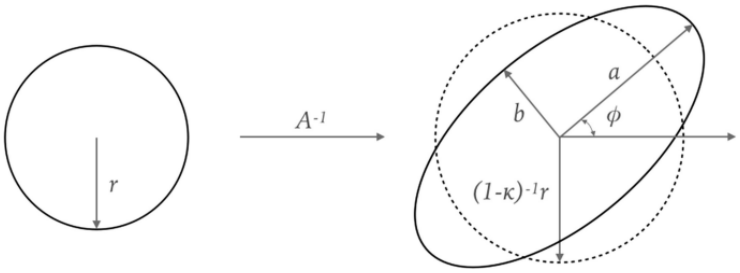


Figure 1.10: Distortion effects on a circular source. Image from Meneghetti (2022).

At this moment it is useful to define the *reduced shear* according to the semi-major and semi-minor axes:

$$g = \frac{a-b}{a+b} = \frac{\gamma}{1-\kappa}. \tag{1.68}$$

When the convergence effect is very small, which is the case for the weak lensing regime, $\kappa << 1$ and we find that $g \simeq \gamma$.

However, the representation in figure 1.10 is not exactly fiducial, since galaxies can not be considered intrinsically circular. Indeed, the ellipticity measured is a combination of the intrinsic source ellipticity (ϵ^s) and the distortion caused by the reduced shear:

$$\epsilon = \epsilon^s + g\,. \tag{1.69}$$

This introduces a complication in the estimation of the reduced shear since we have no means to evaluate the intrinsic shape of each background galaxy. Nevertheless, it is reasonable to assume that in the absence of a lens, the galaxies would be randomly orientated, such that the average ellipticity should be null $\langle \epsilon^s \rangle = 0$. In this way, we are able to overlook the intrinsic shape by simply measuring the average ellipticity of an ensemble of background galaxies:

$$\begin{aligned} \langle \epsilon \rangle &= \langle \epsilon^s \rangle + g \\ \langle \epsilon \rangle &\simeq g \simeq \gamma\,. \end{aligned} \tag{1.70}$$

1.5.3 Weak lensing mass

So far, in our ellipticity analysis, we haven't considered the position of the galaxy relative to the lens center of mass. To do so, it is useful to introduce a new set of coordinates

to analyze the shear, where we define the *tangential* and *cross* shear, $(\gamma_t, \gamma_\times)$ respectively:

$$\begin{aligned} \gamma_t &= -\gamma_1 \cos(2\phi) - \gamma_2 \sin(2\phi) \\ \gamma_\times &= \gamma_1 \sin(2\phi) - \gamma_2 \cos(2\phi)\,. \end{aligned} \tag{1.71}$$

We note that the *tangential shear* is strongly correlated with the mass distribution, such that the mean tangential shear in a certain radius from the lens is a proxy for the mean surface mass density contrast:

$$\langle \gamma_t \rangle(r) = \frac{\bar{\Sigma}(< r) - \bar{\Sigma}(r)}{\Sigma_{crit}} = \frac{\Delta\Sigma}{\Sigma_{crit}}\,, \tag{1.72}$$

where $\bar{\Sigma}(< r)$ is the mean surface mass density within the radius r.

The weak lens mass is usually estimated by fitting the radial profile of the mean tangential shear $\langle \gamma_t \rangle(r)$ measured, assuming a parametric model for the surface mass density Σ.

The mass estimation for one of the CODEX subsamples used in the work will be discussed in the following chapter, where we present an overview of the shape measurements, the surface density model, and the correlations and systematic errors considered.

Chapter 2

Samples

For a meaningful study of the $M - L_X$ scaling relation, one should seek to analyze high-quality data, with more unbiased methods to estimate the observable quantities, hence providing more accurate values for the X-ray luminosity and the cluster mass. On the other hand, to improve precision in the scaling relation analysis, it is usually preferable to work with larger samples, with parameters spanning a more extensive value range (Dawson, 2019; Witte and Witte, 2017). Unfortunately, these two aspects are difficult to conciliate in most astronomical studies.

In an effort to access these requirements, we started out by working with a cluster sample from the currently published CODEX catalog, for which the mass estimates were obtained through weak gravitational lensing from the Canada Hawaii France Telescope (CFHT) data, described in Kiiveri et al. (2021). We were granted access to their pipeline and therefore we were able to retrieve the weak lensing mass PDFs for each cluster. From this sample, we select 28 clusters (hereafter *main-subsample*) to use in this work, which will be further discussed throughout this section. Furthermore, because we have detailed knowledge of the CODEX catalog's construction, we can precisely model and account for the selection effects of the sample.

With the recent release of the extended CODEX catalog, although still not publicly available, we were able to identify and incorporate two additional samples of galaxy clusters from previously published studies, where the mass values were estimated via weak gravitational lensing (detailed in section 2.3). With the inclusion of these extra samples, we were able to analyze the $M - L_X$ relation for 101 galaxy clusters and, since all are part of the extended CODEX catalog, we can consistently apply the same statistical model and bias correction to each subsample.

In the next sections, we describe the CODEX catalog, which is further detailed in Finoguenov et al. (2020), and outline the *main-subsample* and its weak lensing analysis done by Kiiveri et al. (2021). Finally, we also present the general information and selection criteria for the extra samples used in the final analyses.

2.1 The CODEX catalog

The COnstrain Dark Energy X-ray (CODEX) catalog[1] is the first one to perform X-ray selection of galaxy clusters in the Northern hemisphere with flux limit reaching 10^{-13} $ergs$ s^{-1} cm^{-2}. The clusters were thoroughly selected in both X-ray and optical, in the overlapping area of the large-area surveys ROSAT All Sky Survey (RASS, Voges et al. (1999)) and Sloan Digital Sky Survey (SDSS, York et al. (2000)). While ROSAT is still widely used for building catalogs of galaxy clusters, its use for identifying these structures has focused mainly on the brightest subsample. The CODEX survey extends this sample to the lower flux limits accessible with RASS, and reaches high redshifts up to 0.65.

For the catalog's construction, the excesses of 4σ photons within a *wavelet* aperture in the RASS third Data Release (RASS DR3) photon images were considered (Vikhlinin et al., 1998), which means that *background* fluctuations of up to 4σ were rejected in order to avoid false detections (Cibirka, 2017). We also note that the flux measurements for the catalog were based on a few counts, down to 4 photon counts.

Initially, the *redMaPPer* algorithm (Rykoff et al., 2014) was used for every source candidate of the 8^{th} Sloan Digital Sky Survey Data Release (SDSS DR8). This algorithm looks for the presence of *early type* galaxies of similar colors around each source, as clusters typically exhibit a consistent and well-defined population of such galaxies. The identification process involves recognizing a red sequence in the color-magnitude space of the possible cluster. This optical counterpart for cluster detection enables us to differentiate and remove point-like sources (e.g. AGNs), allowing a more robust identification when compared with a purely X-ray selection (Cibirka, 2017; Leauthaud et al., 2010). The recently extended CODEX catalog, which we adopt in this work, uses this same approach with the 9_{th} and 10_{th} Sloan Digital Sky Survey Legacy Data Release (SDSS Legacy DR9, DR10),

[1] The extended CODEX catalog used in this work is not yet publicly available, but its current version can be accessed at `ftp://ftp.mpe.mpg.de/people/alexis/CODEX_aanda_full_flag.fits`. The cluster's spectroscopic properties are released as a part of SDSS-IV DR16 under SPIDERS catalog.

which provides deeper images.

The *redMaPPer* algorithm also yields an estimate for the photometric redshift from the colors of the member galaxies, for the optical center, within 400 kpc from the X-ray center, and for the cluster richness, accounting for the sum of the probabilities of each galaxy belonging to the cluster. This probability of being a cluster member is estimated using a filter function and a background density profile, with the former defined by three main components: a multicolor red sequence model, a projected NFW radial profile, and a Schechter luminosity function (Rykoff et al., 2014).

The cluster's redshift estimation spans from 0.05 to 0.65, but because the *redMaPPer* calibration was only performed for redshifts over 0.1, we limit our discussion to this range (Finoguenov et al., 2020). With the redshift information, it was then possible to obtain rest-frame properties of the sources, such as X-ray luminosity in the $0.1 - 2.4$ keV band, within a radius R_{500}.

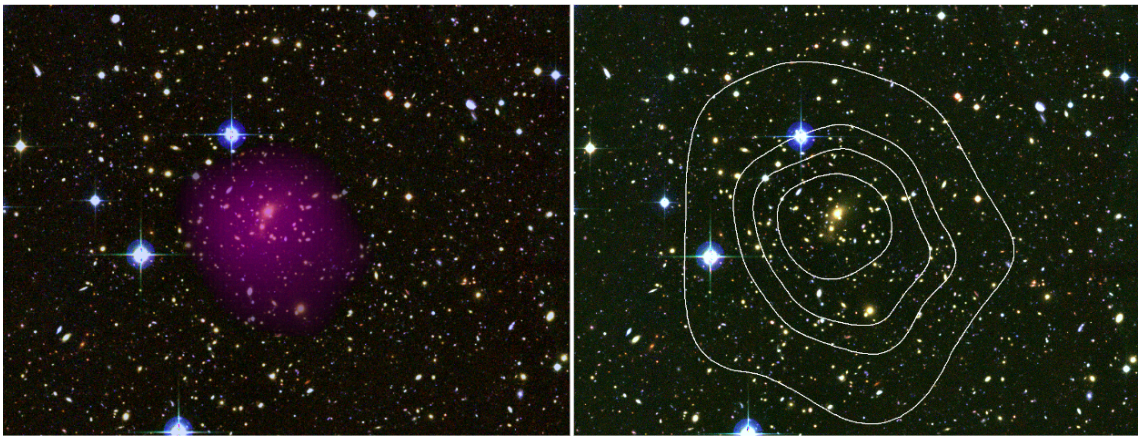


Figure 2.1: Composed optical image of bands *g, r* and *i* from CFHT observations for the CODEX cluster 24981. The X-ray data are from *XMM-Newton* observations in the $0.5 - 2$ keV band, and it is represented in shades of purple (left) and contours (right).

As a way to clean the CODEX catalog, avoiding the low completeness regions, we apply a selection depending on the cluster's richness and redshift, which is further discussed in section 3.2.1. Essentially, this function is responsible for a redshift cut of $z > 0.1$ and ensures a richness uncertainty of less than 10%.

The main distinction of utilizing the CODEX catalog is the availability of the survey's detailed selection function. The CODEX selection accounts for the probability of detecting a source given the observed source counts and the shape parameters of a surface brightness distribution. The function also considers that the cluster shape is covariant to the scatter

in the $M - L_X$ relation and that there is a correlation between the X-ray luminosity and the cluster richness. More on the formalisms of these selection functions can be found in section 3.2.1.

2.2 The main weak lensing subsample

For the study of a mass-observable relation, the use of reliable mass proxies is imperative. The vast majority of $M - L_X$ calibrations were performed using X-ray luminosity or Sunyaev–Zel'dovich signal as mass proxies, which can be affected by hydrostatic mass biases. In this scenario, the weak gravitational lensing method comes as a fine prospect for cluster mass estimates, as it is sensible to the entire mass of the structure and does not rely on assumptions about the cluster's dynamical state, making it a less biased approach. This method, however, requires deep and high-resolution images for the background galaxies' shape measurements and good photometric redshifts for the determination of the lensing critical density (1.60).

With that in mind, according to Cibirka (2017) and Kiiveri et al. (2021), there was a selection comprising the richest and high redshifted clusters in the CODEX catalog ($\lambda_{SDSS} \geq 60$ and $0.35 < z_{SDSS} < 0.65$) for follow-up observation with CFHT Legacy Survey, conducted between 2012 and 2015 with the Wide Field Optical Imaging MegaCam. We note that additional CODEX clusters identified in the CFHTLS-observed fields, which may not strictly meet the initial richness and redshift criteria, were also examined. Among all these clusters, 35 of them present redshift and richness information from the new extended CODEX catalog, as well as weak lensing mass estimations computed by the same approach adopted in Kiiveri et al. (2021). In their paper, however, they considered only the 25 clusters that presented additional CFHT richness information.

As previously mentioned, we focus our analysis on the redshift range $z > 0.1$. This redshift cut is also supported by Damsted et al. (2023), which finds a significant increase of the scatter in the X-ray luminosity for redshifts below this threshold. We also apply selections to clean the catalog and avoid contaminations, all of which will be detailed in section 3.2.1. Finally, we end up with a CODEX-CFHT weak lensing subsample of 28 clusters which we refer to as the *main-subsample*. Basic data of these galaxy clusters, including the ones that did not pass the selections, are available in table A.1.

Because we have access to the details regarding the computation of the mass estimates for this subsample, we can use the complete modeling of their mass probability distribution function (PDF). Throughout the rest of this section, we describe their weak lensing analysis done by Kiiveri et al. (2021) for the *main-subsample* and the considerations made to obtain its mass PDFs.

2.2.1 Shape measurements and redshift estimation

As seen in section 1.5.3, the observable property for estimating the cluster's weak lensing mass is the mean tangential shear $g_t \simeq \gamma_t$, which relates to the density profile $\Delta\Sigma$ by equation 1.72.

For the galaxy shape measurements, the LENSFIT algorithm (Miller et al., 2013) was used over the i-band images, where the FWHM of point-like images are usually smaller (since the atmospheric interference is smaller) when compared with the other bluer bands available (Kiiveri et al., 2021). The algorithm returns the measured ellipticities γ_1 and γ_2, as well as the intrinsic variance σ^2_{intr} and the variance due to observational uncertainties σ^2_{obs}. With these values, we find that the mean tangential shear and $\Delta\Sigma$ for a certain cluster are given by

$$g_t(r) = \sum_i w_i \gamma_{t,i} \; ; \;\; \Delta\Sigma(r) = \sum_i W_i \gamma_{t,i} / \langle \Sigma^{-1}_{crit} \rangle \; , \tag{2.1}$$

where the sum runs over all sources around the cluster, and the wights w_i and W_i are normalized to unit and satisfy

$$w_i \propto \frac{\beta_i}{\sigma^2_{intr} + \sigma^2_{obs}} \; ; \;\; W_i \propto \frac{\beta_i^2}{\sigma^2_{intr} + \sigma^2_{obs}} \; . \tag{2.2}$$

In the above equation, the term β was defined as the ratio between lens-source distance D_{LS}, and the source distance D_S:

$$\beta = \frac{D_{LS}}{D_S} \, , \tag{2.3}$$

such that the critical surface density in equation 1.60 can be writen as

$$\Sigma_{crit} = \frac{c^2}{4\pi G D_S} \beta^{-1} \, . \tag{2.4}$$

To accurately estimate β, it is important to have precise information about the cluster's redshift. We note that spectroscopic redshift information was initially unavailable for 20 clusters from the *main-subsample*, for which red-sequence galaxies were then targeted by several Nordic Optical Telescope (NOT) programs and observed in multi-object spectroscopy mode (Kiiveri et al., 2021).

The definition in equation 2.4 emphasizes the importance of knowing the source redshift to convert observables into physical mass proxies. For the *main-subsample* case, because the lensing signal of each cluster is measured as the mean tangential shear over a large number of galaxies, it is sufficient to estimate the general redshift distribution of the sources (Kiiveri et al., 2021). This is done by assigning each galaxy to a subregion in the CFHT color-magnitude space and comparing it with a reference sample of a high-quality photometric redshift catalog, also distributed in these same subregions. The use of Gruen and Brimioulle (2017) reference catalog with 9-band photometry (CFHTD2), obtained from pointings of CFHTLS, was preferable since it was created in the same way as the CODEX photometric catalog. From these considerations, the value of β of a source was defined as the mean $\langle\beta\rangle$ of all galaxies in the same color-magnitude subregion.

The COSMOS2015 sample by Laigle et al. (2016) was used as an additional reference catalog to cross-match to CFHTLS objects and validate the $\langle\beta\rangle$ estimates. The galaxies in subregions where the estimated $\langle\beta\rangle$ from COSMOS2015 redshifts was below 0.2 were removed, as well as the ones where the estimated $\langle\beta\rangle$ from both COSMOS2015 and CFHTD2 redshifts deviates more than 10% from the median ratio over all subregions. These cuts are exemplified for the cluster 50514 in figure 2.2. Further details in estimating the photometric redshift distribution of galaxy sources can be found in Kiiveri et al. (2021); Cibirka et al. (2017) and Gruen and Brimioulle (2017).

2.2.2 Surface density profile

Once we have the measured mean tangential shear and β values to estimate Σ_{crit}, we now must assume a parametric model for the surface mass density Σ. For the *main-subsample* (as well as for all the other samples used in this work) it was assumed a Navarro, Frank & White (NFW, Navarro et al. (1996, 1997)) mass profile:

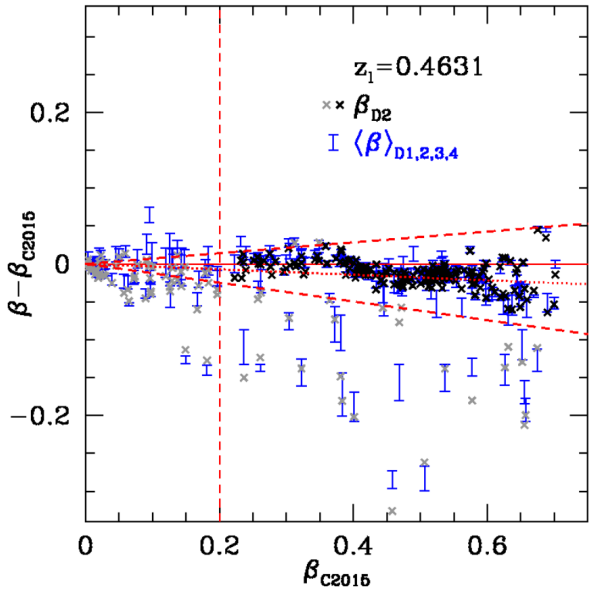


Figure 2.2: Mean value of β in subregions in 5-band color-magnitude spaces estimated from COSMOS2015 photo-z (β_{C2015}), matched galaxies with CFHT Deep + WIRDS photometry (β_{D2}, black/grey points), and CFHT Deep + WIRDS galaxies in all four Deep pointings (blue error bars). The dotted red line indicates the mean ratio $\langle \beta_{D2}/\beta_{C2015} \rangle$. Dashed red lines indicate exclusion criteria of $\beta_{C2015} < 0.2$ or 10% deviation from this median ratio. Results are shown for a lens redshift of $z_l = 0.4631$ (CODEX50514). Figure taken from Cibirka et al. (2017).

$$\rho(r) = \frac{\delta_c \rho_c(z)}{(r/r_s)(1 + r/r_s)^2} \ , \tag{2.5}$$

where ρ_c is given by equation 1.6, and r_s is the scale radius where the logarithmic profile slope changes from -1 to -3. The characteristic over-density of the halo, δ_c, is related to the concentration $c_\Delta = r_\Delta / r_s$ through

$$\delta_c = \frac{200}{3} \frac{c^3}{\ln(1+c) - c/(1+c)} \ , \tag{2.6}$$

where we chose $\Delta = 200$ for this work (Kiiveri et al., 2021).

From the above equations, we see that $\rho(r)$ depends on both the mass M and the concentration c, so in order to reduce one parameter, the M-c relation from Dutton and Macciò (2014) was used:

$$\log_{10} c = a + b \log_{10}(M/[10^{12} h^{-1} M_\odot]) \ , \tag{2.7}$$

where

$$b = -0.101 + 0.026z \ ;$$

$$a = 0.520 + (0.905 - 0.520) \ \exp(-0.617 z^{1.21}) \ .$$

The estimation of the mean $\Delta\Sigma$ was done in 12 logarithmically binned annuli, i.e. calculated within the region between concentric circles around the cluster, with a radial range of [500 - 2500 h^{-1} kpc]. The lower radial limit was best chosen to minimize the mass bias caused by off-centering since this effect is predominant in the smaller scales. Similarly, the upper limit was determined to maximize the statistical power while also excluding the regions where higher-order effects (e.g. 2-halo term) are important.

2.2.3 Covariance matrix

Several physical effects cause the measured $\Delta\Sigma_{obs}$ profile of a cluster to differ from the true $\Delta\Sigma(M)$. The definition of a covariance matrix element C_{ij} that accounts for these effects is detailed in Kiiveri et al. (2021) and we present a brief overview in this section.

The *shape noise* - C_{ii}^{shape} - takes into account the intrinsic ellipticity of source galaxies and the observational uncertainties when measuring their shape, using the LENSFIT values for σ^2_{intr} and σ^2_{obs}, as discussed in section 2.2.1.

The *uncorrelated large scale structure* - C_{ij}^{LSS} - was also considered in the covariance matrix, and it is due to the presence of randomly distributed structures in the line of sight. Since on smaller scales the Universe can be considered inhomogeneous, there are density fluctuations that cause an additional bending of the light rays, hence adding noise in the weak lensing mass estimate. Although the additional shear signal is zero on average, the overall variance due to this effect can be calculated by integrating over the convergence power spectrum defined by Limber (1954) and depends on the source sample, redshift weighting, and angular size of the anulli (Kiiveri et al., 2021).

The last component of the covariance matrix accounts for the *intrinsic variations of cluster profiles* - C_{ij}^{intr} - i.e. the noises that would be present even in ideal observational conditions. This covariance element contemplates several physical effects, such as (i) C_{ij}^{conc} for the misleading assumption that the halo density is homogeneous, where in fact the halo concentration may vary, (ii) C_{ij}^{ell} for assuming a spherically symmetric halo, when in truth there is an intrinsic ellipticity and orientation, (iii) C_{ij}^{corr} for the presence of correlated secondary halos, due to recent mergers and galaxy accretion, and (iv) C_{ij}^{off} for variations due to off-centering. All these components were described using the semi-analytical model

from Gruen et al. (2015)[2].

Altogether, we define the $\Delta\Sigma(M)$ covariance matrix as

$$C_{ij} = C_{ij}^{shape} + C_{ij}^{LSS} + C_{ij}^{intr} \ . \tag{2.8}$$

2.2.4 Mass likelihood

The weak lensing mass likelihood is proportional to the probability of measuring a $\Delta\Sigma$ for a cluster of true mass $M = M_{200c}$. According to Kiiveri et al. (2021), assuming multivariate Gaussian errors, we can write the mass probability density function (PDF) as

$$P(\Delta\Sigma|M) \propto \frac{1}{\sqrt{\det C}} \exp\left(-\frac{1}{2}E(M)^T C^{-1} E(M)\right) \ , \tag{2.9}$$

where

$$E(M) = \Delta\Sigma_{obs} - \Delta\Sigma_{model} \ , \tag{2.10}$$

and C is the covariance matrix described in the previous section.

Throughout this work, we use equation 2.9 in the logarithmic space for the *main-subsample*. We define $\mu = \log_{10} M_{200c}$, and adopt the notation $P(\tilde{\mu}|\mu)$, where the $\tilde{\mu}$ denotes the observed variable. Figure 2.3 exemplifies the mass PDF for cluster 16566 from the *main-subsample*.

In Kiiveri et al. (2021) they also introduce a parameter for the systematic uncertainties l_{sys}, that is used to compute how different the noiseless lensing masses are from the true masses due to miscalibration of lensing shapes, redshifts, and the cluster density profiles. The systematic uncertainty is used as a multiplication factor for the density profile, and the l_{sys} parameter is responsible for changing the amplitude of such factor to assimilate the errors better. As it will be further discussed in section 4, we found that this parameter does not correlate with any of the others.

2.2.5 Cosmology correction

So far, we presented a brief review of the work done by Kiiveri et al. (2021) and Gruen et al. (2015) to obtain the weak lensing mass PDF. The cosmology adopted by them for calculating the covariance matrices, described in 2.2.3, used a mass density value of

[2] The code used in Gruen et al. (2015) is available at `https://github.com/danielgruen/ccv`.

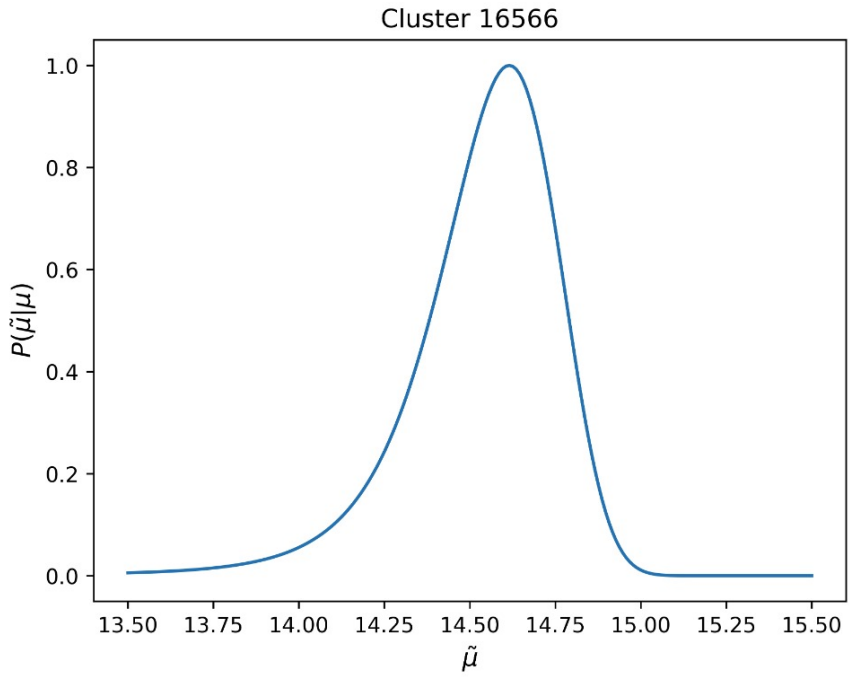


Figure 2.3: Mass PDF in logarithmic space for cluster 16566 from the WL sample.

$\Omega_m = 0.27$. Since in this work, we adopt a $\Omega_m = 0.3$, this discrepancy in cosmology can affect the weak lensing mass estimates and therefore we must perform a correction.

This correction for the weak lensing masses seems quite straightforward when considering only the general weak lensing formalism, where the Ω_m dependence relies on the estimated critical surface density Σ_{crit}. We saw from equation 2.4 that Σ_{crit} is inversely proportional to the diameter angular distance, which in turn is $\propto \Omega_m$, such that in the end, we have $\Delta\Sigma \propto \Sigma_{crit} \propto \Omega_m^{-1}$. This means that a change in the value of Ω_m will result in a difference in the mass estimate of $\Omega_m^{old}/\Omega_m^{new}$.

However, this correction might turn out to be much more complicated if we account for the covariance matrix within the mass likelihood (equation 2.9). Fortunately, although the CODEX covariance matrix includes many terms, in the end, the error is dominated by the number of source galaxies used (Finoguenov et al., 2020). Therefore, after some consideration, we have decided that it would be enough to rescale the systematic errors by the difference in $\Delta\Sigma$ ($\Omega_m^{old}/\Omega_m^{new} = 0.27/0.3 = 0.9$), which results in an increase of 10%.

2.3 Other samples

The addition of other galaxy cluster samples is interesting for increasing the precision and the mass and X-ray luminosity range of our analysis. With that in mind, we searched for previously published catalogs that computed weak lensing mass estimates for galaxy clusters, and that are also part of the extended CODEX catalog. We favored those that shared some (but not necessarily all) similar features regarding the weak lensing analysis, such as the $M - c$ relation used, the density profile, and the cosmology adopted.

Herbonnet et al. (2020) computed the weak lensing mass of 100 X-ray selected clusters from the Multi Epoch Nearby Cluster Survey (MENeaCS) and the Canadian Cluster Comparison Project (Hoekstra et al., 2015, CCCP) using CFHT images. From those, 83 clusters are present in the extended CODEX catalog. The cosmological parameters used in their work are the same as ours, and they also adopt the $M - c$ relation from Dutton and Macciò (2014). However, the redshift information available in Herbonnet et al. (2020) was obtained by searching the NASA Extragalactic Database (NED) where not all clusters had *spectroscopic* redshift estimates. We then decided to adopt the z_{spec} values available in the extended CODEX catalog from which only 7 clusters were missing information (clusters 55927, 60040, 64520, 64521, 66237, 68049, and 70574). Fortunately, after searching the literature for each cluster, either by the cluster name or by sky coordinates - adopting a 2 arcmin match and a maximum redshift difference of $\Delta z = 0.05$ between the available z_{phot} - we were able to retrieve the spectroscopic information for all these galaxy clusters.

After we applied the same cuts used in cleaning the catalog and additional selections in richness, both detailed in section 3.2.1, a total of 46 galaxy clusters were considered. This subsample (hereafter *Herbonnet20*) comprises massive clusters, with masses ranging from $\sim 3.2 \cdot 10^{14} M_{\odot}$ to $\sim 2.8 \cdot 10^{15} M_{\odot}$, and reaching a redshift value of $z \sim 0.55$. This subsample, including the clusters that did not pass the selection, is displayed in table A.2.

The hundreds of shear-selected clusters from Oguri et al. (2021) are also a great addition to our work, for which they used the Hyper Suprime-Cam Subaru Strategic Program (S19A) to compute the clusters' mass via weak gravitational lensing. Their work also adopts the same cosmological parameters used in this study, although, unlike this work, they do not use a $M - c$ relation and opted to fit the NFW profile with both parameters. From the Oguri et al. (2021) catalog, a subsample of 84 clusters is also part of the CODEX

catalog.

The redshift information available is also photometric, obtained by matching the shear-selected clusters with the optically-selected clusters from several other catalogs. We then adopt the same approach used with the Herbonnet et al. (2020) subsample for obtaining *spectroscopic* redshift information. Of the 84 CODEX clusters, further literature research was needed for 8 of them (clusters 57134, 57165, 57175, 57260, 57524, 57528, 58031, and 58158). Only the galaxy cluster 57134, also identified as RXC J0213.9-0253, did not have any spectroscopic information, for which we decided to adopt the photometric redshift value. We do not expect that this single exception will have a meaningful impact on the overall study. After analyzing the relation $z_{spec} - z_{phot}$ for the other cluster samples, as shown in figure 2.4, we can also suppose that the z_{phot} value provides an accurate representation of the cluster's true redshift.

From the sample presented above, 27 galaxy clusters pass the selections (hereafter *Oguri21*). These clusters present a mass range of $\sim [2.1 \cdot 10^{14} M_{\odot} - 1.5 \cdot 10^{15} M_{\odot}]$ and a redshift range of $[0.13 - 0.58]$. The *Oguri21* subsample and also the clusters that did not pass the selection are displayed in table A.3.

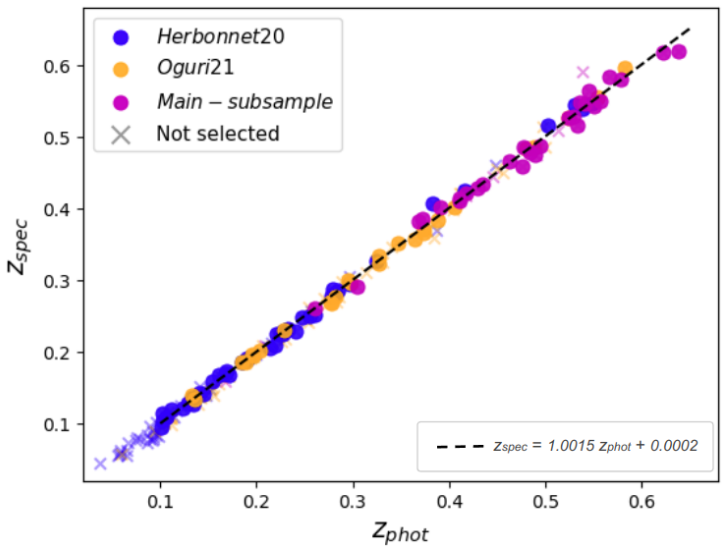


Figure 2.4:

The matches between these two samples and the extended CODEX catalog were made considering a 3-arcminute distance from the clusters' center and a redshift interval of $\Delta z = 0.05$. To ensure that there were no repeated galaxy clusters among all subsamples (i.e. *main-subsample*, *Herbonnet20*, and *Oguri21*) we applied this same matching and

found no duplicates. In conclusion, our final sample consists of 101 CODEX galaxy clusters with weak lensing mass estimates, which we will further refer to as the *WL sample*. Figure 2.5 displays our final sample and the cleaned CODEX catalog.

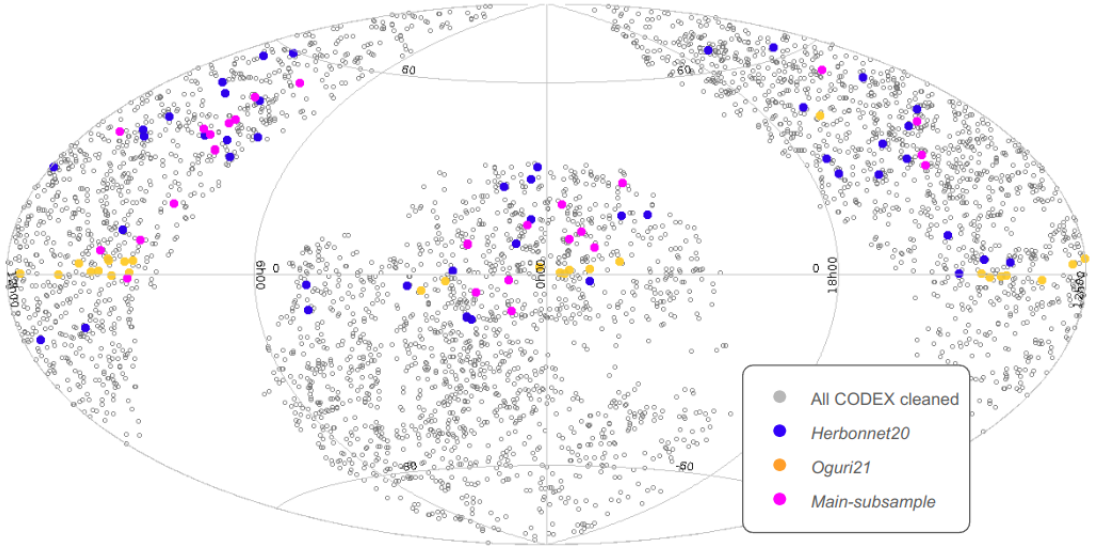


Figure 2.5: Aitoff projection of our final *WL sample* - dots in blue, orange and pink represents the *Herbonnet20*, *Oguri21*, and *main-subsample*, respectively. The grey dots are the cleaned CODEX catalog.

We also show in figure 2.6 the clusters from the *WL sample* distributed in the mass - X-ray luminosity logarithmic space, and the redshift distribution of each subsample. From the former, we can confirm that all subsamples agree with a general trend of increasing luminosity with mass. As for the letter plot, we observe that the *main-subsample* introduced by this work also increases considerably the number of high-redshifted galaxy clusters, which is particularly interesting for our goal to analyze the temporal evolution of the $M - L_X$ scaling relation.

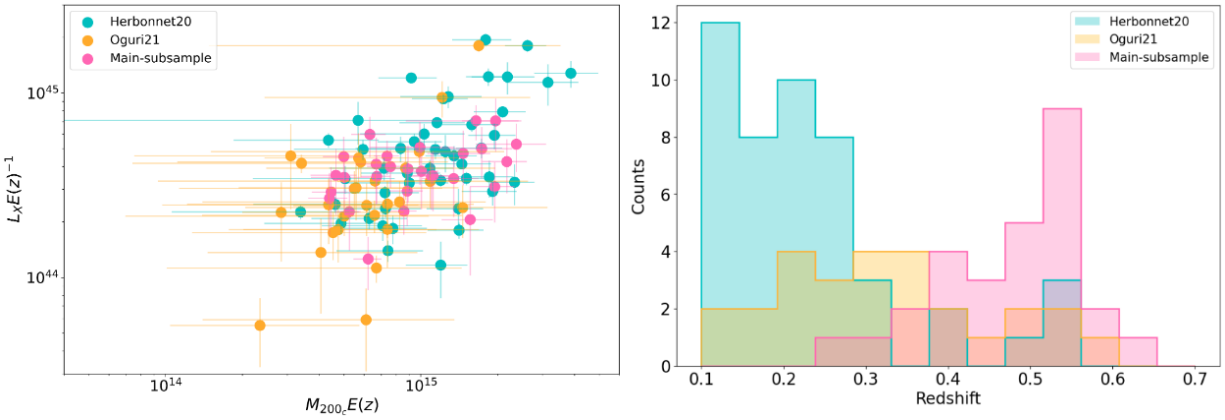


Figure 2.6: (*Right*) Galaxy clusters from the *WL sample* in the mass - X-ray luminosity logarithmic space. (*Left*) Redshift distribution of the *WL sample*.

Chapter 3

Statistical analysis

Throughout the centuries, there has been a large debate about the meaning and goals of statistics. If centuries ago it was seen as merely the compilation of data, nowadays statistics is broadly referred to as techniques derived from mathematics based on inductive reasoning, and useful in the presence of uncertainty (Feigelson and Babu, 2012). Because in any scientific endeavor, the available information is incomplete, our knowledge is necessarily probabilistic, and astronomers seek reasonable statistical approaches to obtain quantitative interpretations of the data (Feigelson and Babu, 2012; Gregory, 2005).

Scientists have long agreed on the important role of deductive reasoning in the scientific method, particularly when it comes to testing theoretical models. In this scenario, two different approaches stand out; the *frequentist* interpretation of probability, and the *Bayesian probability theory*. While the first is usually thought of as the 'conventional' statistics, currently there has been a shift in the general community's view to acknowledge the Bayesian statistical approach. In the next section, we discuss briefly the main features of these definitions, with emphasis on the Bayesian statistics, which will be adopted in this work.

3.1 Bayesian statistics

The term Bayesian statistics was first introduced in 1763, with a posthumous publication of a manuscript written by the mathematician Reverend Thomas Bayes. However, the formalism and popularization of the Bayesian methodology were developed only around a decade later, by Pierre-Simon Laplace in 1774, and even then, classical frequentist statistics remained the preferable approach (Ivezić et al., 2020; Gregory, 2005). Scientists

disregarded the use of Bayes' principle until the second half of the 20^{th} century, mainly because computers had only then reached enough speed to allow a meaningful Bayesian analysis, and because of the works adopting this approach done by renowned statisticians such as Finetti de (1937), Cox (1961) and Jaynes (1957).

The probability as seen through a frequentist view is the relative frequency of an event to occur throughout many repeated experiments and therefore relies on random variables. As for a Bayesian Inference's definition, the probability $P(A|B)$ is the measured value of the plausibility of A given that the proposition B is true. This definition allows us to directly compute the probability of a theory or model parameter based on our current state of knowledge (Gregory, 2005).

The sum and product rules, following the Bayesian formalism, can be written as

$$\begin{aligned} &\text{Sum: } P(A|B) + P(\bar{A}|B) = 1 \\ &\text{Product: } P(A, B|C) = P(A|C)P(B|A, C) = P(B|C)P(A|B, C) \ , \end{aligned} \tag{3.1}$$

where two symbols separated by a comma means that both propositions are true, i.e. the *joint probability*, and where any proposition on the right side of the bar | is assumed to be true.

If we define H_i as a proposition for our hypothesis, K as the proposition of our prior knowledge, and D as the proposition representing our data, it follows from the product rule that

$$P(H_i|D, K) = \frac{P(H_i|K)P(D|H_i, K)}{P(D|K)} \ . \tag{3.2}$$

The above equation is called the *Bayes Theorem*, where

- $P(H_i|D, K) \longrightarrow$ is the *posterior* probability of H_i given D and K.
- $P(H_i|K) \longrightarrow$ is the *prior* probability of H_i given K.
- $P(D|H_i, K) \longrightarrow$ is the probability of obtaining the data given that H_i and K are true; also known as the *likelihood function*.
- $P(D|K) \longrightarrow$ is the normalization factor; also known as the *evidence*.

Because many parameter estimation problems seek only to maximize the posterior distribution, for which we would obtain the best parameters, it is common to express Bayes theorem omitting the evidence:

$$P(H_i|D,K) \propto \; P(H_i|K)\; P(D|H_i,K)\;. \tag{3.3}$$

It is worth noting that a Bayesian probability distribution function (PDF) relates to our state of information. For example, if we only have vague prior information about our problem, $P(H_i|K)$ must be a 'broad' function, spanning a wide range of parameter values (Gregory, 2005).

Furthermore, a model is termed hierarchical if the Bayesian formulation is written in multiple levels of observational quantities. A key feature of such models is acknowledging that the data might have intrinsic dependencies, such that we are not limited to perceiving parameters as fixed values, but rather explore the notion that they might present their own distribution. To illustrate this concept, consider a scenario involving a basic linear fit, represented as $y = mx$, where we do not have access to the true variable x, but whether a measured value with error $\tilde{x} = x + \epsilon$. In conventional Bayesian notation, we would have

$$P(y|x) \propto P(x|y)\; P(y)\;. \tag{3.4}$$

However, we do not have information of $P(x|y)$, and hence we must write it in a hierarchical manner:

$$P(y|\tilde{x}) \propto \int dx\; P(\tilde{x}|x)\; P(x|y)\; P(y)\;, \tag{3.5}$$

where $P(\tilde{x}|x)$ is a function that describes the measured error in x (e.g. normal distribution with deviation ϵ).

Essentially, the hierarchical method is when you describe the entire model as a series of submodels. It is an ideal approach for describing a complete model of the data, accounting for systematic and measured errors, as well as selection effects.

3.1.1 Marginalization

In addition to bringing information into the model, the Bayesian theory of probability also allows us to 'eliminate' parameters that we are not interested in (Ivezić et al., 2020; Gregory, 2005). Let us assume a parameter estimation problem, where we have some data D and our model is defined by two distinct parameters $H = \{a, b\}$, such that our joint probability is given by $P(a, b|D, K)$. If we want to compute the implications of our data only for a, it is possible to eliminate the parameter b through what is called *marginalization*:

$$P(a|D, K) = \int dbP(a, b|D, K) \ . \tag{3.6}$$

If we introduce Bayes' theorem (equation 3.2) to this situation,

$$P(a, b|D, K) = \frac{P(a, b|K)P(D|a, b, K)}{P(D|K)} \ , \tag{3.7}$$

and assume that the priors for a and b are independent, we find that

$$P(a|D, K) \propto P(a|K) \int dbP(b|K)P(D|a, b, K) \ . \tag{3.8}$$

The expression above gives us the *marginal posterior* as a function of the average of the likelihood function $P(D|a, b, K)$.

Having presented the basics of Bayesian inference, we must now create our own hierarchical model, which will be described in the following section.

3.2 Hierarchical model

From the discussion presented in the previous section, one can state that the main reason for adopting the Bayesian view, over the frequentist tradition, is to include prior information about the model in addition to what the data can tell us. Furthermore, it also provides the means to eliminate unwanted parameters and account for the selection functions of the data sample, which is why we adopt this approach.

At first, we considered using two established Bayesian linear regression tools, the LINMIX[1] in Python, and the LInear Regression in Astronomy (LIRA) in R language (Sereno,

[1] Information available at `https://github.com/jmeyers314/linmix`.

2016). With both approaches, we were able to account for measurement errors in the dependent and independent variables, while also considering the presence of intrinsic scatter. The main difference was in the handling of the evolution parameter γ, which was not possible with LINMIX. Hence for this method, the redshift dependence had to be addressed separately.

Moreover, because in both approaches the errors had to be symmetric, we couldn't fully incorporate the information of our entire mass PDF (e.g. equation 2.9), and we were compelled to simplify it by fitting a Gaussian function and using its mean and deviation as mass values and errors, respectively.

In an attempt to remediate this issue, we tried fitting the mass PDFs using a mixture of Gaussian functions, as proposed in Sereno (2016), but found that it wasn't worth introducing a correlation between our data points. Furthermore, precise modeling of the selection functions was impractical in both approaches. After careful consideration, we have decided to create our own Bayesian hierarchical model, using Kiiveri et al. (2021) as a reference.

This work seeks to adjust the scale relation in logarithmic space, given by equation 1.19. In this scenario, where we have a relation between our observable $l_X = \log L_X$ and the true mass $\mu = \log M$, given a model with parameters $\boldsymbol{\theta} = \{\alpha, \beta, \gamma, \sigma_{intr}\}$, the joint probability distribution that there is a cluster of total mass μ, at a specific redshift z, can be described as (Kelly, 2007):

$$P(l_X, \mu|\boldsymbol{\theta}, z) = P(l_X|\mu, \boldsymbol{\theta})P(\mu|z)P(z). \tag{3.9}$$

In the above equation, the term $P(\mu|z)$ refers to the halo mass function (HMF), providing the number density of dark matter halos (e.g. of galaxy clusters) as a function of mass. We adopt the HMF defined by Tinker et al. (2008), which used 22 simulations to calibrate the function parameters:

$$\frac{dn}{dM} = f(\sigma)\frac{\bar{\rho}}{M}\frac{d\ln\sigma^{-1}}{d\ln M}\ , \tag{3.10}$$

where σ is the rms variance, $\bar{\rho}$ is the mean matter density, and $f(\sigma)$ is called halo multiplicity function. Their work fitted the local ($z = 0$) halo number with a 5% statistical

precision, considering a ΛCDM Universe (Tinker et al., 2008; Allen et al., 2011). Figure 3.1 illustrates the HMF general behavior in the logarithmic space, where we observe the prevalence of smaller structures compared to the rarer, more massive structures.

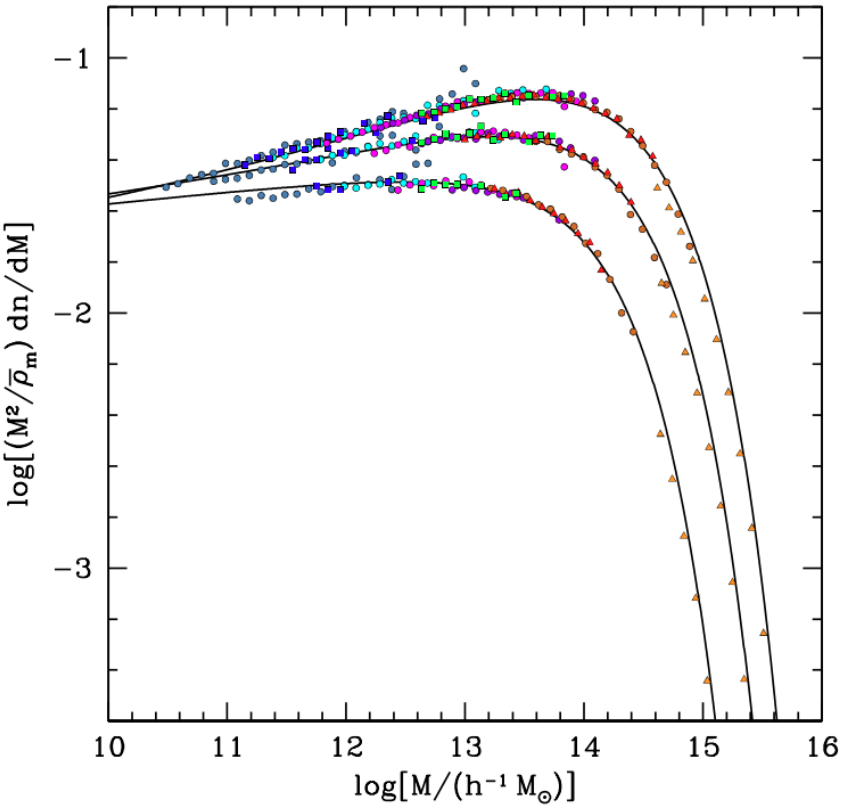


Figure 3.1: Logarithmic measured mass functions for the WMAP1 simulations and three best fit curves considering $\Delta = 200$, 800, and 3200 (*from top to bottom*). Modified figure from Tinker et al. (2008).

The term $P(z)$ in equation 3.9 is the differential comoving volume, defined as

$$\frac{dV(z)}{dz} = \left(\frac{c}{H_0}\right)^3 \frac{D^3(z)}{E(z)}\,; \qquad \text{where } D(z) = \int_0^z \frac{dz}{E(z)} \text{ is the growth function.} \tag{3.11}$$

Both terms combined account for the conditional probability of the existence of a galaxy cluster with a given mass and redshift.

As for the term $P(l_X|\mu, \boldsymbol{\theta})$, it provides the probability of having a scattered X-ray luminosity l_X given an underlying true luminosity - obtained by a true mass value and true model parameters, $\langle l_X|\mu, z, \boldsymbol{\theta}\rangle$. This function is illustrated in figure 3.2. Because it is reasonable to assume that the data is normally scattered around its theoretical expected value (Sereno, 2016; Allen et al., 2011), our function is modeled as a normal distribution

centered in l_X and with deviation given by the intrinsic scatter σ_{intr}:

$$P(l_X|\mu, z, \boldsymbol{\theta}) = \frac{1}{\sqrt{2\pi}\sigma_{intr}} exp\left[-\frac{1}{2}\frac{(\langle l_X|\mu, z, \boldsymbol{\theta}\rangle - l_X)^2}{\sigma_{intr}^2}\right] ; \tag{3.12}$$

where $\langle l_X|\mu, z, \boldsymbol{\theta}\rangle = \alpha + \beta\mu + \gamma \log(1+z)$.

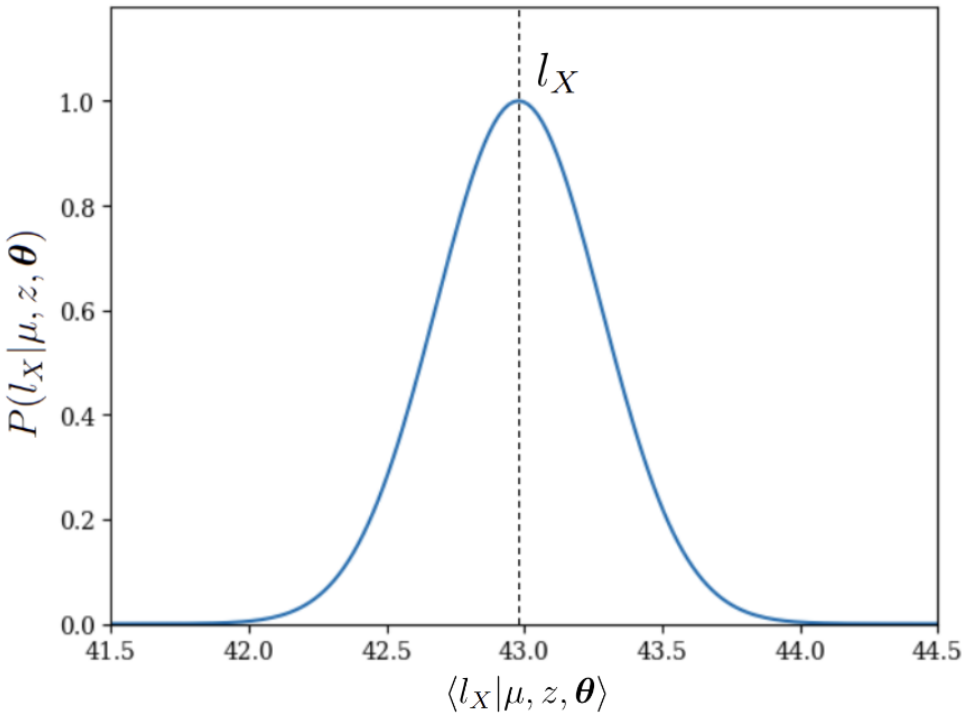


Figure 3.2: A lognormal probability distribution to account for the presence of intrinsic scatter in the X-ray luminosity, given a true mass μ, a redshift, and our model parameters $\boldsymbol{\theta}$.

When effectively incorporating the terms in equation 3.9, we are also addressing significant selection biases (i.e. Malmquist and Eddington bias).

Because our sample of clusters is X-ray selected, one must consider that this observable might suffer deviations from its intrinsic value due to effects unrelated to observational techniques (e.g. extra sources within the line-of-sight, triaxiality, and mis-centering of the radial scale). Although these effects typically manifest at levels below 10%, it is important to note that the bias is dependent on the detection method and the masses of the clusters, and thus, should be considered in the statistical analysis (Allen et al., 2011).

Figure 3.3 illustrates the situation of a cluster population selected by an X-ray flux threshold, also known as the Malmquist bias, and that presents an intrinsic scatter around

the theoretical scaling law, also known as the Eddington bias. The right-hand figure also adds the effect of the underlying mass function (i.e. HMF illustrated in figure 3.1) in the observed data, where the number density of low-mass clusters is much greater. One can observe that, for a given mass, the high X-ray luminosity clusters are more likely to be detected, and the overall effect of these biases is the resulting sample not being representative of the full cluster population, especially for the low mass and low luminosity systems. Other sample selections will be further discussed in section 3.2.1.

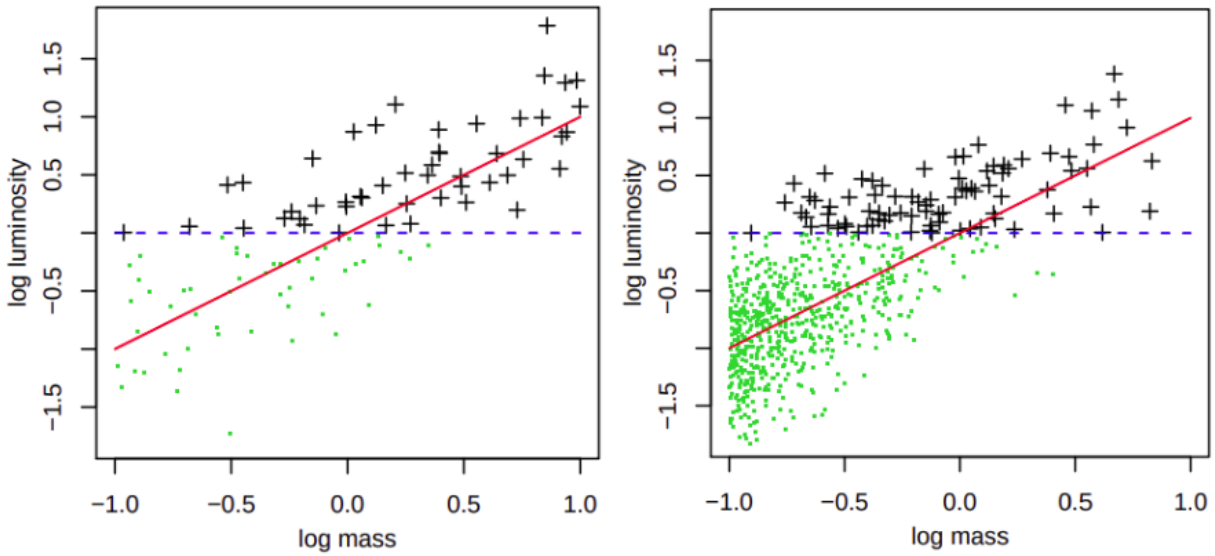


Figure 3.3: Illustration of the Malmiquist and Eddington selection biases. The *blue deshed* line represents the X-ray flux threshold and the *red solid* line represents the theoretical scaling relation. The *black crosses* and *green dots* represent the detected and undetected sources, respectively. The *right* figure represents the more realistic case of halo mass distribution. One can observe that, because of the selection effects, the detected sample is not representative of the entire population and does not reflect the underlying scaling relation. Figure from (Allen et al., 2011).

We also want to consider the effects caused by measurement errors, since we cannot directly access the observable quantity. Therefore, using Kiiveri et al. (2021) notation, we must model the probability distribution of measured observables (denoted by $\tilde{l_X}, \tilde{\mu}, \tilde{z}$) given true scattered observables (l_X, μ, z). For our case, it is reasonable to consider the spectroscopic redshift equal to the true redshift, so that $P(\tilde{z}|z)$ is modeled as a Dirac delta function, and therefore will not be explicit in the following equations (Kelly, 2007).

The probability of measuring a cluster mass $\tilde{\mu}$ through gravitational lensing is given by the mass PDF $P(\tilde{\mu}|\mu)$. For the *main-subsample*, it is simply the equation 2.9, discussed in section 2.2.4. For the other two subsamples, *Herbonnet20* and *Oguri21*, we use the information available in the respective catalogs and model the mass PDFs as Gaussian functions

in the linear space, $P(\tilde{M}_c|M_c)$. These normal distributions are centered at the weak lensing mass estimate value with deviation given by the measurement errors. Specifically for *Oguri21* subsample, because the errors are not symmetric, we create a combination of two Gaussians, both centered in the estimated mass value, with deviations given by the errors δ_M^- and δ_M^+. Figure 3.4 illustrates the normal functions described above for clusters of the *Herbonnet20* and *Oguri21* subsamples, as well as the respectively mass PDFs $P(\tilde{\mu}|\mu)$ in log space.

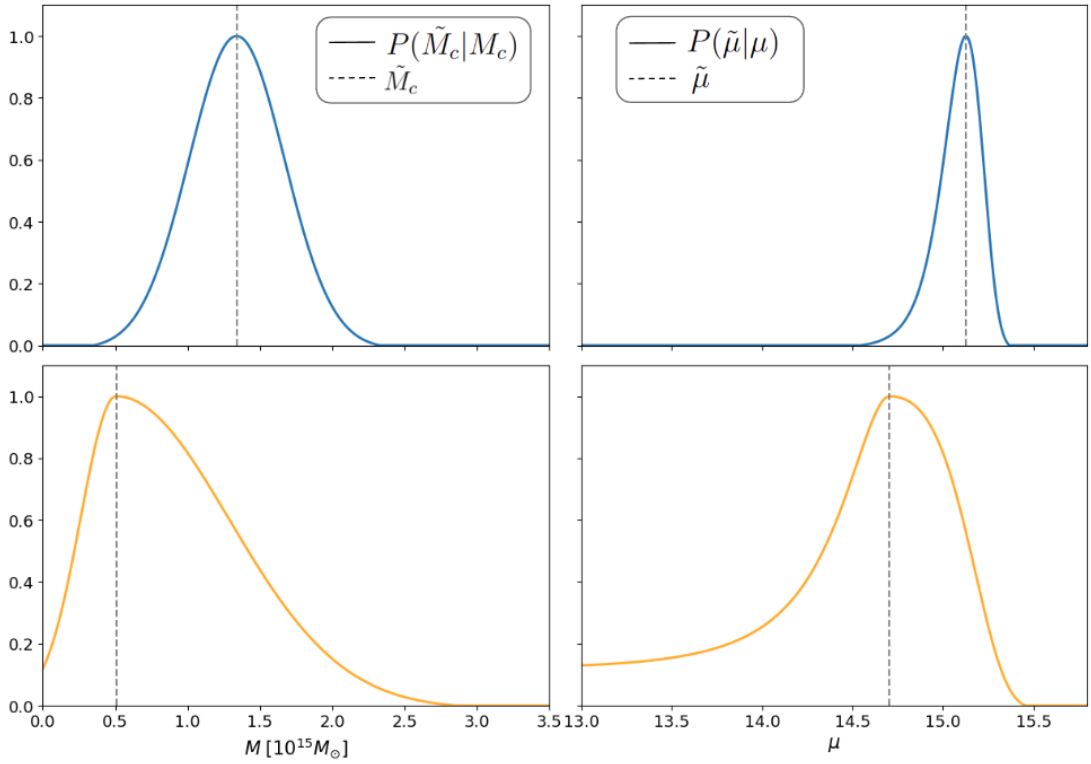


Figure 3.4: The mass probability distribution function for a cluster from the Herbonnet20 subsample (*upper panels*) and from the Oguri21 subsample (*lower panels*). The *left hand* panels are in the linear space, while the *right hand* panels are in the logarithmic space.

As for the measured observable $\tilde{l_X}$, we have $P(\tilde{l_X}|l_X)$ also modeled as a normal distribution in the linear space, that accounts for the observational uncertainties δ_{L_X}, as illustrated by figure 3.5. We compute a Gaussian distribution in the linear space (with L_X), as shown in equation 3.13, and interpolate it with the luminosity in the logarithmic space (l_X).

$$P(\tilde{L_X}|L_X) = \frac{1}{\sqrt{2\pi}\delta_{L_X}} exp\left[-\frac{1}{2}\frac{(L_X - \tilde{L_X})^2}{\delta_{L_X}^2}\right] ; \tag{3.13}$$

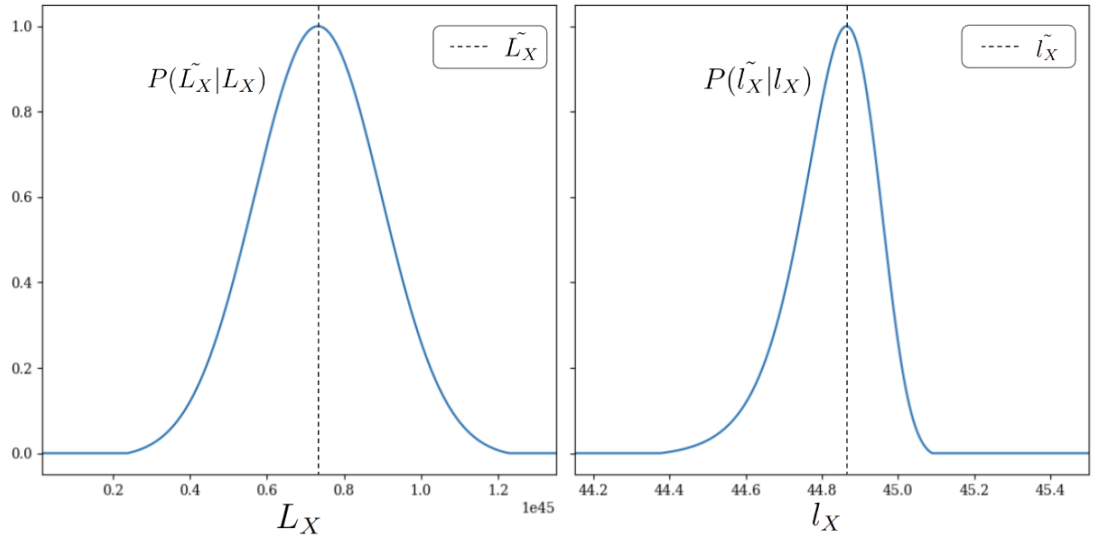


Figure 3.5: Illustration of the probability distribution of measuring a value of X-ray luminosity $\tilde{l_X}$ given that we have a scattered value of l_X. In the linear space (*left*) it is simply a Gaussian distribution.

With these considerations, we can modify equation 3.9 and, since we have access only to the measured quantities, we can marginalize it over all the true quantities, such that

$$P(\tilde{l_X}, \tilde{\mu}, \tilde{z}|\boldsymbol{\theta}) = \int dl_X \, d\mu \underbrace{P(\tilde{l_X}|l_X)\, P(\tilde{\mu}|\mu)}_{\substack{\text{Meas.-Scatt.} \\ \text{values}}} \underbrace{P(l_X|\mu, \boldsymbol{\theta})}_{\substack{\text{Scatt.-True} \\ \text{values}}} \underbrace{P(\mu|\tilde{z})}_{\text{HMF}} \underbrace{P(\tilde{z})}_{\frac{dV}{dz}} . \tag{3.14}$$

Apart from the considerations above, we must also contemplate the effects of the selections made during the construction of our subsamples. In a general context, we can introduce the variable I as a condition to whether or not the cluster passed the selection, such that the probability of having the measured observables $\tilde{\boldsymbol{o}} = \{\tilde{l_X}, \tilde{\mu}, \tilde{z}\}$ given that the cluster passed the selection is

$$P(\tilde{\boldsymbol{o}}|I, \boldsymbol{\theta}) = \frac{P(I|\tilde{\boldsymbol{o}}, \boldsymbol{\theta})\, P(\tilde{\boldsymbol{o}}|\boldsymbol{\theta})}{P(I|\boldsymbol{\theta})}. \tag{3.15}$$

Since our primary objective is to obtain the probability distribution of parameters given the observables $P(\boldsymbol{\theta}|\tilde{\boldsymbol{o}})$, according to Bayes theorem, equation 3.15 is actually the likelihood function $P(\tilde{\boldsymbol{o}}|I, \boldsymbol{\theta}) = L(\tilde{\boldsymbol{o}}|I, \boldsymbol{\theta})$ we need to model:

$$P(\boldsymbol{\theta}|\tilde{\boldsymbol{o}}) \propto \pi(\boldsymbol{\theta})\, L(\tilde{\boldsymbol{o}}|I, \boldsymbol{\theta}), \tag{3.16}$$

where $\pi(\boldsymbol{\theta})$ are the priors for the parameters, which are assumed to be independent.

We can see that the term $P(\tilde{\boldsymbol{o}}|\boldsymbol{\theta})$ in equation 3.15 represents the probability of obtaining the observables $\tilde{\boldsymbol{o}}$ given our model, and it is actually the expression we already defined in 3.14. Also, the general probability of all clusters passing the selection - denoted by $P(I|\boldsymbol{\theta})$ - can be described by simply marginalizing the numerator over all observables:

$$P(I|\boldsymbol{\theta}) = \int d\tilde{\boldsymbol{o}}\; P(I|\tilde{\boldsymbol{o}}, \boldsymbol{\theta})\, P(\tilde{\boldsymbol{o}}|\boldsymbol{\theta}). \tag{3.17}$$

Finally, we can write our full likelihood function as:

$$\begin{aligned} L(\tilde{l}_X, \tilde{\mu}, \tilde{z}|I, \boldsymbol{\theta}) \propto \int & dl_X\, d\mu\;\; P(I|\tilde{\boldsymbol{o}}, \boldsymbol{\theta}) \\ & \cdot P(\tilde{l}_X|l_X)\, P(\tilde{\mu}|\mu) \\ & \cdot P(l_X|\mu, \boldsymbol{\theta}) P(\mu|\tilde{z}) P(\tilde{z}). \end{aligned} \tag{3.18}$$

3.2.1 Selection functions

The likelihood function described previously (equation 3.18) depends on the sample's selections represented by the term $P(I|\tilde{\boldsymbol{o}}, \boldsymbol{\theta})$. In this section, we give an overview of the selection functions used in the construction of the *WL sample*, using the work done in Kiiveri et al. (2021) and Finoguenov et al. (2020) as reference.

Two optical selections were applied to the CODEX catalog sample. Finoguenov et al. (2020) simulated the completeness limits for the CODEX catalog using RASS and SDSS data in each 0.1 width bin of redshift, as illustrated in figure 3.6. It was found that the 10% completeness curve for RASS matches the Klein et al. (2019) definition of low contamination sample - a 5% contamination when identifying RASS sources using the Dark Energy Survey (DES).

As a means to account for the 10% completeness, one must then remove the sources for which the richness λ is below the curve, by applying the following selection (Finoguenov et al., 2020):

$$\lambda > 22 \left(\frac{z}{0.15}\right)^{0.8} . \tag{3.19}$$

We consider this selection in our work as the function:

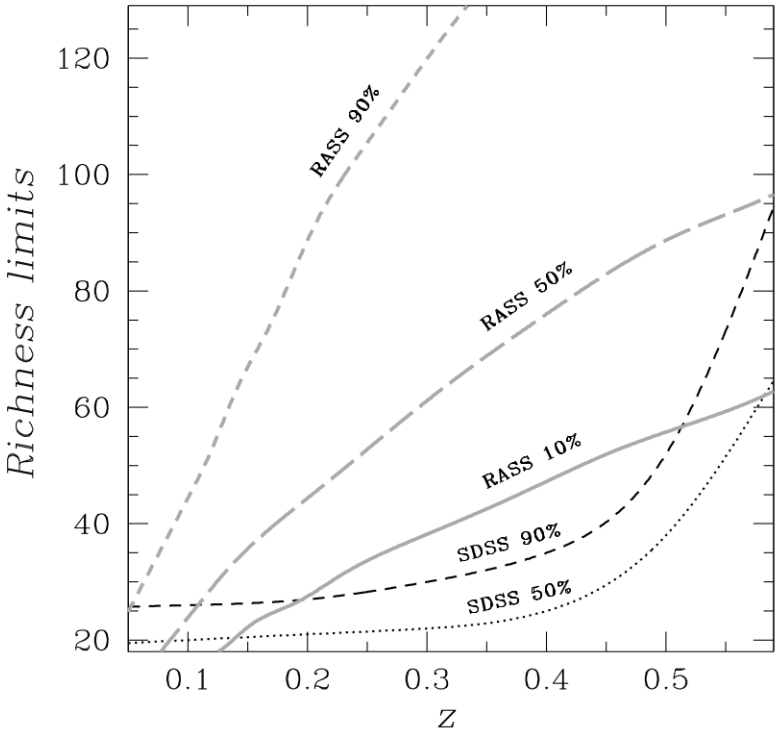


Figure 3.6: Richness completeness limits of the CODEX survey. Black curves show the 90% (dashed) and 50% (dotted) completeness limits of redMaPPer using SDSS data. Grey curves are the 10% (solid), 50% (long-dashed), and 90% (short-dashed) completeness limits of RASS data. Figure taken from Finoguenov et al. (2020).

$$P(I_{RASS}|\tilde{\lambda}, z) = \begin{cases} 1, & \text{if } \tilde{\lambda} > 22\left(\frac{z}{0.15}\right)^{0.8} \\ 0, & \text{otherwise} \end{cases} . \tag{3.20}$$

In addition, it was found that for $z < 0.2$ and $z > 0.5$, it was important to account for the optical completeness of the SDSS data. To model this selection function we follow the analytical form adopted by Finoguenov et al. (2020), obtained based on the tabulation of Rykoff et al. (2014). However, we suppress the redshift dependence of their function since we are now using deeper optical images (i.e. DR9 and DR10 instead of DR8). Therefore, the probability of detecting a cluster in the SDSS data, considering a 50% completeness, is

$$P(I_{opt}|\ln\,\lambda) = 1 - \frac{1}{2}\,erfc\left(\frac{\ln\lambda - \ln\lambda_{50\%}}{\sqrt{2}\,\sigma}\right), \tag{3.21}$$

where $\ln\lambda_{50\%} = \ln(17.2)$ and we adopt $\sigma = 0.2$ for the CODEX catalog (Finoguenov et al., 2020).

For further refining of the CODEX catalog, we also implement a redshift cut of $z > 0.1$,

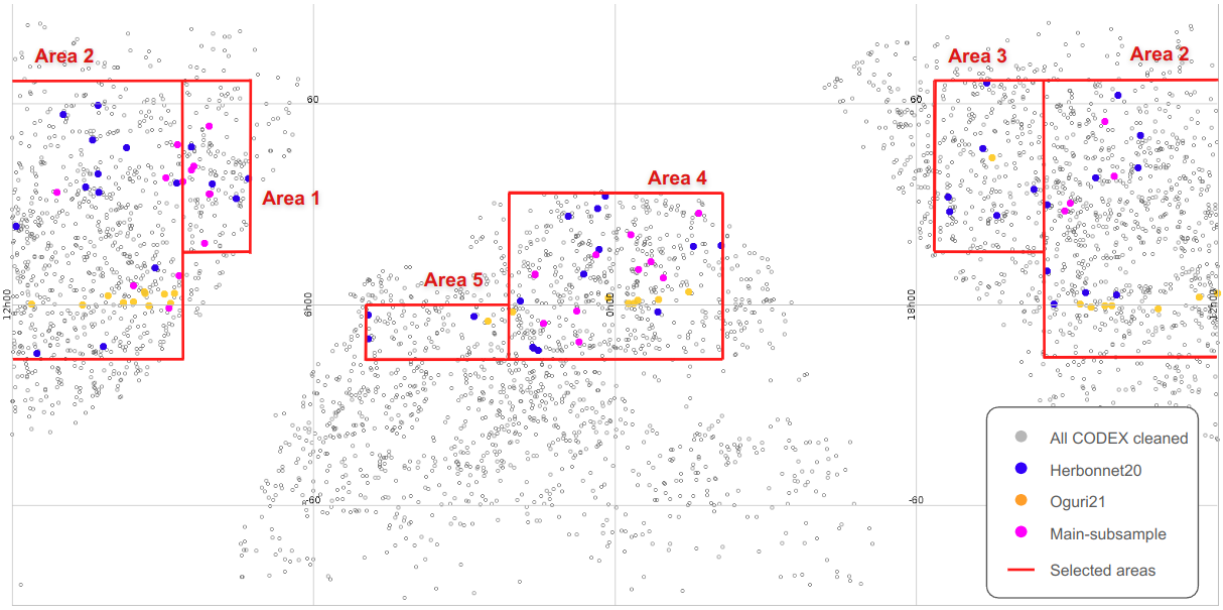


Figure 3.7: Projection of CODEX samples' coordinates onto a Cartesian plane, along with delineation of the five specified areas.

to avoid the redshift range where the *redMaPPer* algorithm is inefficient, and a richness cut of

$$\frac{\delta_\lambda}{\lambda} < 0.1 \,, \tag{3.22}$$

where δ_λ is the measurement error in the richness. This cut ensures that the richness error is less than 10%.

In addition, for our sample to be considered representative of the cluster population, the ideal scenario would be for the ratio between the number of clusters in the *WL sample* and the total number of clusters in the cleaned CODEX catalog to be ~ 1. However, as one can observe from figure 2.5, this is far from reality. A way to mitigate this issue is to define a more restricted survey area to work with, focusing on the sky regions where our subsamples lie. Figure 3.7 shows the cartesian projection of the samples' coordinates and the five areas defined for this purpose; the final survey area used in this work is essentially the combination of each of them.

With that being said, we can now implement a subsample selection function, in which all clusters with richness below 60 were excluded as a way to ensure a more pure and representative sample. However, because not all clusters with $\tilde{\lambda} \geq 60$ are actually observed, this selection differs from a step function, and was constructed following the approach adopted in Kiiveri et al. (2021). We consider the initial sample from the CODEX cleaned

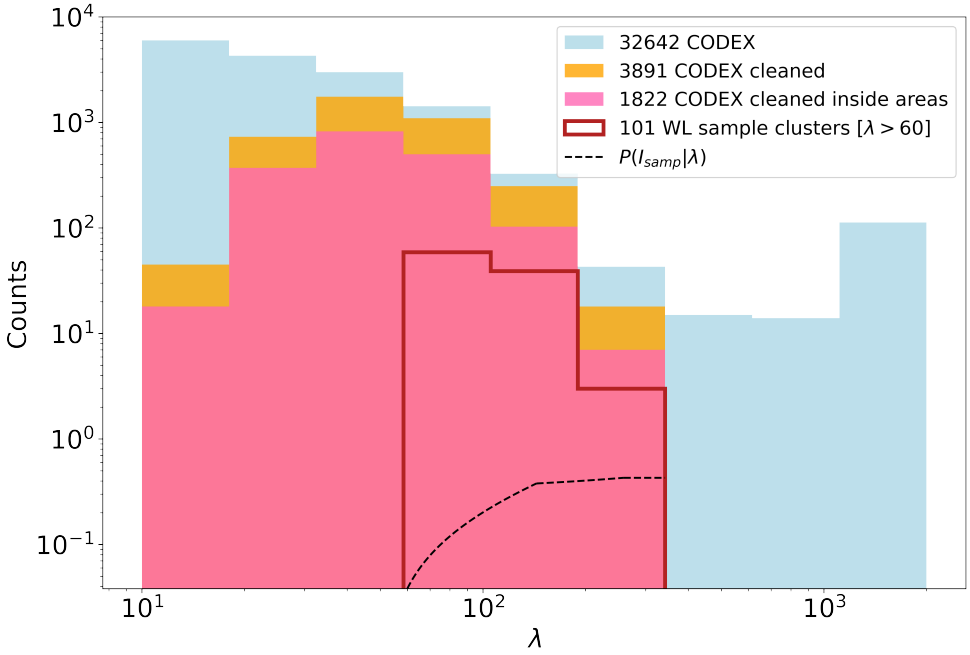


Figure 3.8: Richness distribution of the CODEX sample in equally width bins. The subsample selection function is represented by the dashed black line.

catalog (inside the defined survey area) and our final weak lensing sample of 101 clusters. We bin the samples by richness into equal logarithmic bin widths and compute the ratio of the height of the bins. After, we fit a linear piecewise function between the mean of each bin, obtaining:

$$P(I_{samp}|\tilde{\lambda}) = \begin{cases} 0, & \tilde{\lambda} < 60 \\ \frac{4}{1000}\tilde{\lambda} - \frac{202}{1000}, & 60 \leq \tilde{\lambda} < 143.5 \\ \frac{0.4}{1000}\tilde{\lambda} + \frac{316}{1000}, & 143.5 \leq \tilde{\lambda} < 258.6 \\ \frac{429}{1000}, & 258.6 \leq \tilde{\lambda} < 342 \end{cases} \tag{3.23}$$

Figure 3.8 illustrates this selection.

Finally, there is also the CODEX selection function for RASS's X-ray images, for which we give a brief overview. Further details are described in Finoguenov et al. (2020).

The CODEX selection function returns the effective survey area for the catalog and is used for correcting the probability of detecting a galaxy cluster in a comoving volume, given certain observables. This selection considers the probability of detecting a source

given the observed source counts - η^{ob} - and the shape parameters of a β-profile[2] surface brightness distribution - $\beta(\mu)$ and $r_c = \ln R_c$ - denoted by $P(I|\eta^{ob}, \beta(\mu), r_c)$.

The function also takes into account that the cluster shape is covariant to the scatter in the $M - L_X$ relation, although this effect is expected to be reduced for redshifts over 0.3 for which the RASS clusters are point-like. In addition, the function considers the presence of a correlation between the X-ray luminosity and the cluster's richness. This joint probability is denoted as $P(r_c, l_X|\mu, \nu, z)$, where the deviation from mean richness ν is defined as

$$\nu = \frac{\ln \lambda \; - \; \langle \ln \lambda | \mu \rangle}{\sigma_{\ln \lambda}}, \tag{3.24}$$

with $\langle \ln \lambda | \mu \rangle$ being the expected richness given a true mass.

The probability of observing source counts given the expected number count - $P(\eta^{ob}|\eta^{true})$ - is considered as well, and is modeled as a Poisson distribution:

$$P(\eta^{ob}|\eta^{true}) = \frac{(\eta^{true})^{\eta^{ob}} e^{-\eta^{true}}}{\eta^{ob}!}, \tag{3.25}$$

where

$$\eta^{true} = \frac{L_X \cdot S}{4\pi D_L^2 K(\langle T|L_X \rangle, z)}.$$

In the expression above, S denotes the sensitivity (counts per flux), D_L is the luminosity distance, and $K(\langle T|L_X \rangle, z)$ is the K correction.

In effect, this CODEX selection function is available as a grid, resulting from simulations that accounted for the probabilities described above, and returns the effective survey area and the values for ν, μ, l_x, σ_{intr} and z, used to calculate it. In the end, interpolating this simulated grid can give us a function for the probability of detecting a galaxy cluster with luminosity l_X, redshift z, and intrinsic scatter σ_{intr} in the CODEX catalog, denoted by $P(I_X|\tilde{z}, l_X, \sigma_{intr})$. Figure 3.9 illustrates the behavior of this selection as a function of each of the three parameters.

At last, we are able to write the full selection functions as follows:

[2] A β-profile surface brightness distribution with radius r is given by $S_B = \left[1 + \left(\frac{r}{R_c}\right)^2\right]^{-3\beta+0.5}$

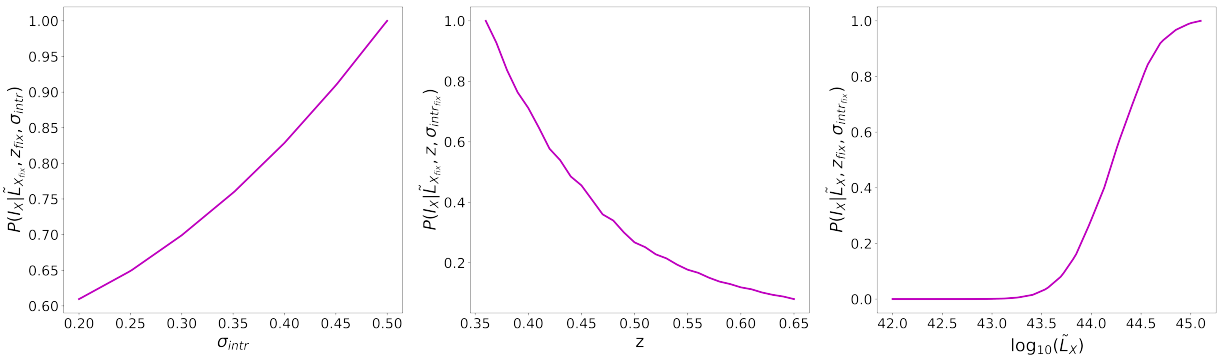


Figure 3.9: Behavior of the CODEX selection function for fixed values of intrinsic scatter, *redshift*, and X-ray luminosity, respectively.

$$
\begin{aligned}
P(I|\tilde{\lambda}, \tilde{z}, \sigma_{intr}) = \int & \, d\mu \, dl_X \\
& \cdot P(I_{RASS}|\tilde{\lambda}, \tilde{z}) \xleftarrow{\text{10\% completness of RASS}} \\
& \cdot P(I_{opt}|\ln \tilde{\lambda}) \xleftarrow{\text{50\% completness of SDSS}} \\
& \cdot P(I_{samp}|\tilde{\lambda}) \xleftarrow{\text{subsample selection}} \\
& \cdot P(I_X|\tilde{z}, l_X, \sigma_{intr}). \xleftarrow{\text{CODEX simulated matrix}}
\end{aligned}
\tag{3.26}
$$

3.2.2 Priors

As we already discussed in section 3.1, one of the distinctions in adopting the Bayesian Inference view, is the possibility of introducing extra prior information about our problem. We can combine the information from the data with our prior knowledge, to obtain the posterior probability distribution (Gregory, 2005). Depending on our choice for the functional form of the distribution, we might have a posterior that is almost entirely dominated by the prior, or determined essentially by the data information (i.e. the likelihood), as illustrated in figure 3.10.

Because we are interested in what our data reveals about the scaling relation, we might use the so-called *uninformative priors*, which incorporates weak but objective information about a parameter. It is worth noting that despite the name, these priors still affect the parameter estimates, and the results are usually different than the ones obtained by the frequentist and maximum likelihood approaches (Ivezić et al., 2020). The prior for a location parameter, for example, should not change with translations of the coordinate

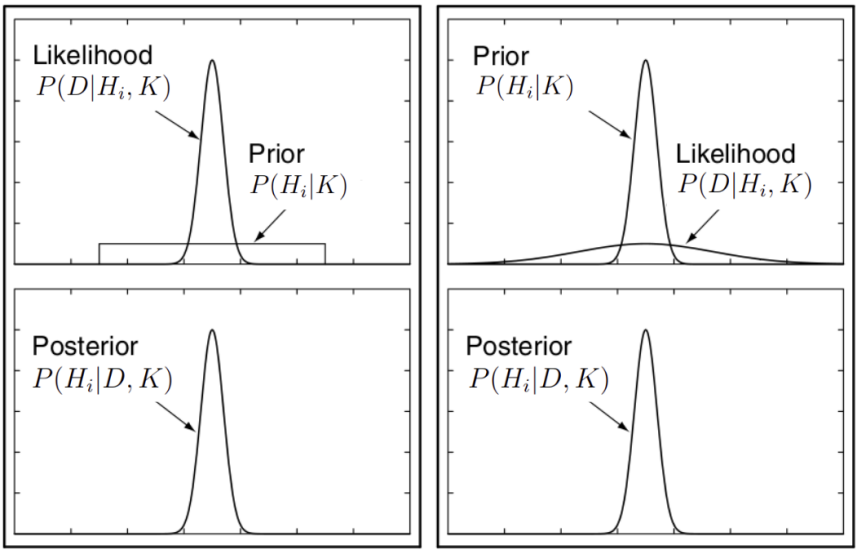


Figure 3.10: Illustration of two extreme combinations of prior and likelihood distributions (*upper panels*); When the posterior distribution is entirely dominated by the likelihood function (*lower left*) and when the posterior is essentially the prior distribution (*lower right*). Modified figure from Gregory (2005).

system, while priors for the intrinsic scatter are limited to positive values.

To our analysis, we adopt *flat uninformative priors* for the intercept, and the mass and redshift slopes:

$$\pi(\boldsymbol{\theta} = \{\alpha,\ \beta,\ \gamma\}) = \begin{cases} const.\ , & \boldsymbol{\theta}_{i_{min}} < \boldsymbol{\theta}_i < \boldsymbol{\theta}_{i_{max}} \\ 0\ , & Otherwise. \end{cases} \tag{3.27}$$

The minimum and maximum values for the α parameter were set to be $[10;\ 45]$, for β were $[0;\ 2.5]$, and for the γ parameter we chose $[-5;\ 5]$.

The intrinsic scatter parameter σ_{intr} is what one calls the *scale parameter*, i.e. it should not depend on the choice of units, such that if we rescale our measurement units by a positive factor a, we get a constraint

$$\pi(\sigma)d\sigma = \pi(a\sigma)d(a\sigma)\ . \tag{3.28}$$

The solution to the above is simply

$$\pi(\sigma) \propto \sigma^{-1}\ , \tag{3.29}$$

which is called a *scale invariant prior* or the *Jeffreys prior* (Ivezić et al., 2020; Gregory, 2005). In general literature, there is no overall agreement about the choices from weakly informative priors for the scale parameter. The Jeffreys prior described above is one of the most common approaches, although the inverse χ^2, gamma, and inverse-gamma distributions are sometimes suggested. However, one common aspect that seems to arise from all these different discussions is that the use of a uniform prior for the scale parameter is not recommended, as it results in a miscalibration towards higher values of σ (Walter G., 2010; Gelman, 2006).

It is clear from equation 3.29 that the distribution diverges at $\sigma = 0$, so if we intend to extend our lower limit to zero, we must consider using the *modified Jeffreys prior* (Gregory, 2005):

$$\pi(\sigma) = \frac{1}{(\sigma + a)\ln[(\sigma_{max} + a)/a]} , \tag{3.30}$$

where a is a constant that eliminates the singularity.

Although we can expect our values of the intrinsic scatter to be around 0.33 - that corresponds to a scatter in the mass of ~ 0.25 (Leauthaud et al., 2010) in our $M - L_X$ relation - we decide to adopt a modified Jeffreys prior with $a = 0.03$, for which the function's mean value is at $\sigma_{intr} = 0.33$. We also tested using the gamma and inverse gamma distributions but found no significant differences.

Finally, for the lensing systematic uncertainty parameter l_{sys}, we followed the definition used by (Kiiveri et al., 2021), which adopts a Gaussian distribution centered in zero with dispersion 1.

In conclusion, we model three likelihood functions (L_i) for each subsample we analyze (i.e. *main-subsample*, *Herbonnet20* and *Oguri21*). These are given by the general equation 3.18 with the discussed selections in equation 3.26, such that the posterior distribution we need to maximize in order to estimate our best parameters is, in log space,

$$\log P(\boldsymbol{\theta}|\tilde{\boldsymbol{o}}) \propto \log \pi(\boldsymbol{\theta}) + \log L_{\rm main} + \log L_{\rm Herbonnet20} + \log L_{\rm Oguri21}. \tag{3.31}$$

Chapter 4

Results and discussion

4.1 Main results

In order to sample our posterior distribution for the parameters, we used the Python package EMCEE (Goodman and Weare, 2010; Foreman-Mackey et al., 2013), which consists of a Monte Carlo Markov Chain algorithm. We found that it is enough to execute 24 chains with 2,000 steps each and burn the first 350 steps that it takes for the chains to converge. In this configuration, we identified the optimal iteration time by parallelizing the code across fourteen CPUs. As a result, the entire program can be executed in approximately four days. Our best values and their errors were computed as the median and standard deviation of the distributions for each parameter, respectively.

First, we applied our general Bayesian analysis of the 101 clusters in the *WL sample* with the uninformative priors discussed in section 3.2.2. Our results are displayed in table 4.1 and figure 4.1 shows the best parameters and their correlation in a triangle plot.

Table 4.1 - Parameters, their initial values, priors, and posteriors from the MCMC fitting with all samples.

Parameter	Initial	Prior	Posterior
Intercept α	21	flat(10, 45)	40.1 ± 0.8
Slope β	1.56	flat(0, 3)	0.30 ± 0.05
Evolution γ	0	flat(-5, 5)	1.54 ± 0.55
Scatter σ_{intr}	0.33	Jeff. prior	0.21 ± 0.01
Systematic error l_{sys}	0	N(0,1)	4.25 ± 0.37

We found a significant shallower slope than expected by self-similarity, with a value of $\beta = 0.30 \pm 0.05$. This small value might be due to our focus on analyzing only high-mass systems, which is discussed throughout the next sections.

The intrinsic scatter value was also smaller than the ones reported by most previous works - over 0.25 (e.g. Kettula et al., 2015; Leauthaud et al., 2010; Eckmiller et al., 2011). However, it is expected a smaller scatter when analyzing a high redshift sample (section 2.2, Damsted et al., 2023). Indeed, Eckmiller et al. (2011) reported a change in scatter from 0.287 for a local group sample ($z < 0.1$) to 0.153 for a cluster sample with $0.1 < z < 0.3$.

As for the temporal evolution parameter, we found a value of $\gamma = 1.54 \pm 0.55$, which is in disagreement with the self-similar prediction of zero in a 2.8σ level. We also used the MCMC chains to estimate the probability of γ being positive, done by calculating the ratio between the number of iterations with positive values and the total length of the chains, which resulted in a probability of $P(\gamma > 0) = 99.62\%$.

In essence, despite the modest slope value that might suggest the influence of unknown systematic factors, our findings render it statistically significant to propose a positive evolution in the $M - L_X$ relation. At the end of the section, we delve deeper into the interpretation and implications of this evolution parameter value, along with all the other γ results presented throughout this work.

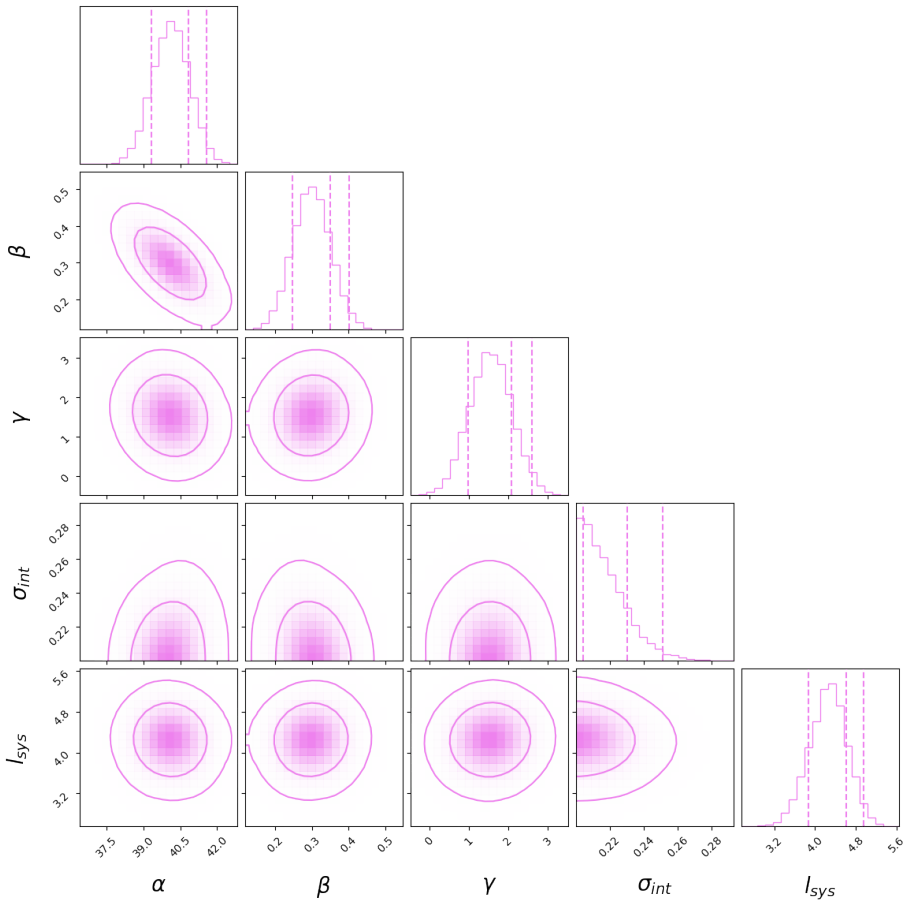


Figure 4.1: Best parameters' values from the MCMC fitting with the one and two-dimensional projections of the posterior distributions. The contours represent the 1σ (68%) and 2σ (95%) confidence regions.

We also show our best fit for the *WL sample*, which will be hereafter referred to as the main-fit, projected on the logarithmic mass - X-ray luminosity plane in figure 4.2. We also show a simple linear fit using the Python function LMFIT.MINIMIZE[1] with slope fixed in 1.3, which illustrates the self-similar prediction for comparison. The lower panels highlight the fact that we are using probability distribution functions in our analysis, by representing the mass PDFs (*lower-left*) and both the mass and X-ray luminosity PDFs (*lower-right*) for each cluster. The PDF representation might provide a clearer visualization of the relationship between the best fit and the data.

[1] Documentation available at `https://lmfit.github.io/lmfit-py/fitting.html`.

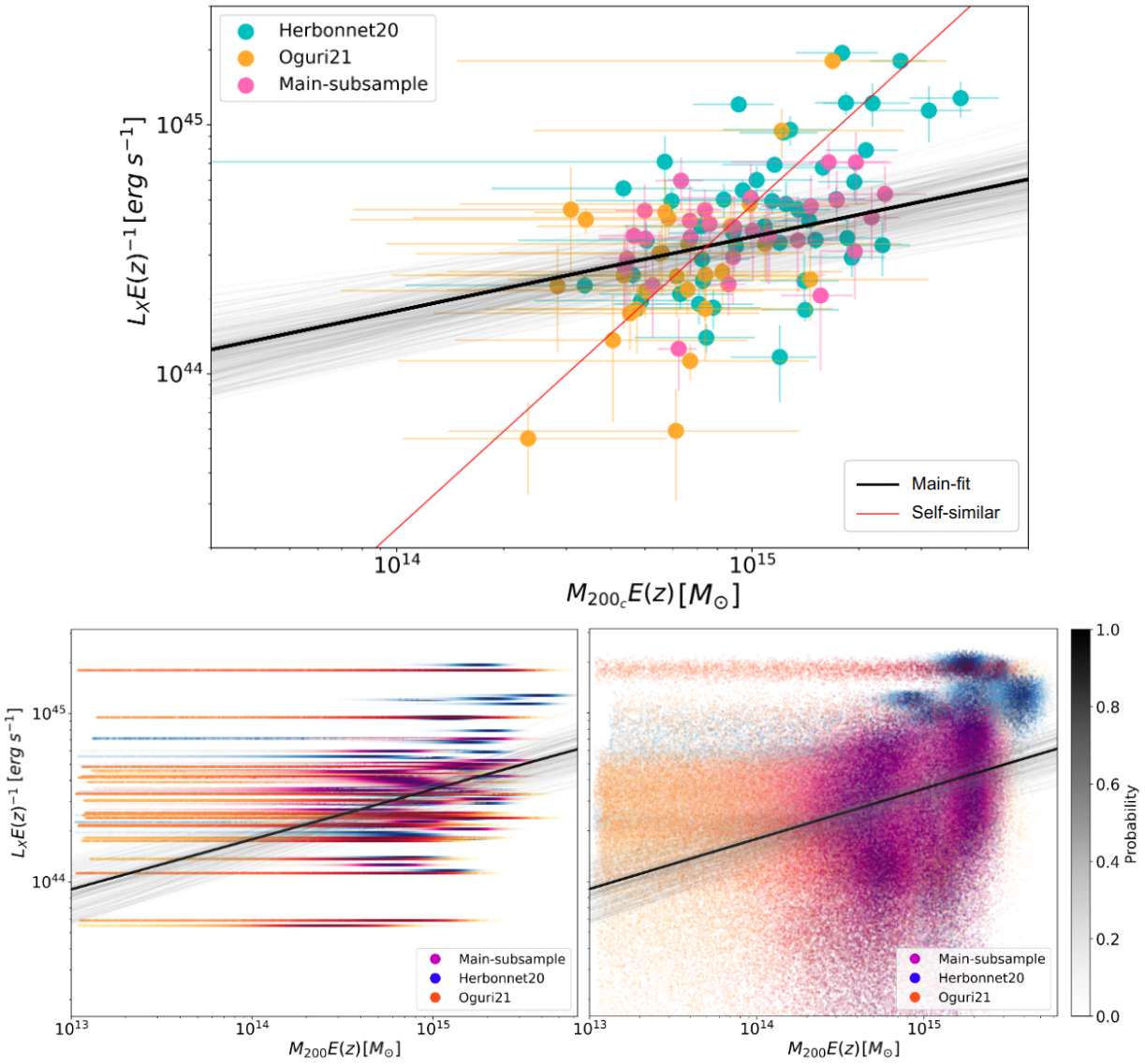


Figure 4.2: The $M - L_X$ relation in log space. The black line represents the projected best fit obtained from the MCMC. The results from a few iterations are depicted by the faded gray lines, illustrating the error associated with the best fit. In the *lower-left* panel, only the mass probability distribution functions for each cluster are displayed. In the *lower-right* panel, both the mass and X-ray luminosity PDFs are presented.

4.1.1 *Features of elements in the statistical model*

We also wanted to test the impact of our general Bayesian statistical approach and, more specifically, the impact of the selection functions on the calibration of the scaling relation. In order to achieve this, we applied three other fits to our *WL sample*:

(i) A simple linear fit using the Python function LMFIT.MINIMIZE, where we disregard

the redshift term in equation 1.19 and the evolution evaluation is done separately;

(ii) Applying our Bayesian statistical approach but neglecting all the selection functions and the Halo Mass Function (HMF) described in section 3.2;

(iii) Applying our Bayesian statistical analysis without the selection functions but considering the HMF.

For the first case, using the LMFIT.MINIMIZE function, which disregards measurement errors and intrinsic scatter, we found a slope value of 0.50 ± 0.09, which is steeper when compared with our previous results, but still in disagreement with self-similarity.

Furthermore, the temporal evolution of $M - L_X$ was also analyzed by obtaining a linear fit of the residuals and the redshift. This was compared with a second fit that fixed the slope value at zero, as depicted in figure 4.3. The latter fit has only one free parameter (instead of two) and corresponds to the model where there is no evidence of evolution.

To compare the two fits, the values of the Bayesian Information Criterion (BIC) for each of them were used to calculate $\Delta BIC = BIC_{model} - BIC_{best}$, which is the difference between a model BIC_{model} and the best model BIC_{best} (i.e. the one with the lowest BIC value). We opted to use the BIC parameter as it seeks to maximize the posterior of the fit while introducing penalty terms for the number of parameters in the model, effectively mitigating the risk of overfitting. The value of ΔBIC can be considered a test of the evidence against the new model (Arevalo et al., 2017; Fabozzi, 2014), such that:

- For $\Delta BIC < 2$, it is considered worth mentioning the evidence against the new model.
- For $2 < \Delta BIC < 6$, it is considered that the evidence against the new model is positive.
- For $6 < \Delta BIC < 10$, it is considered that the evidence against the new model is strong.
- For $\Delta BIC > 10$, it is considered that the evidence against the new model is very strong.

From this analysis, the value of $BIC_{best} = -279.8$ was obtained for the first fit — with two free parameters and corresponding to a model with some evolution — and $BIC_{model} = -278.9$ for the second fit — with only one parameter and corresponding to the model where there is no evidence of evolution. Thus, $\Delta BIC = 0.879$, and therefore, it is only worth mentioning evidence of evolution.

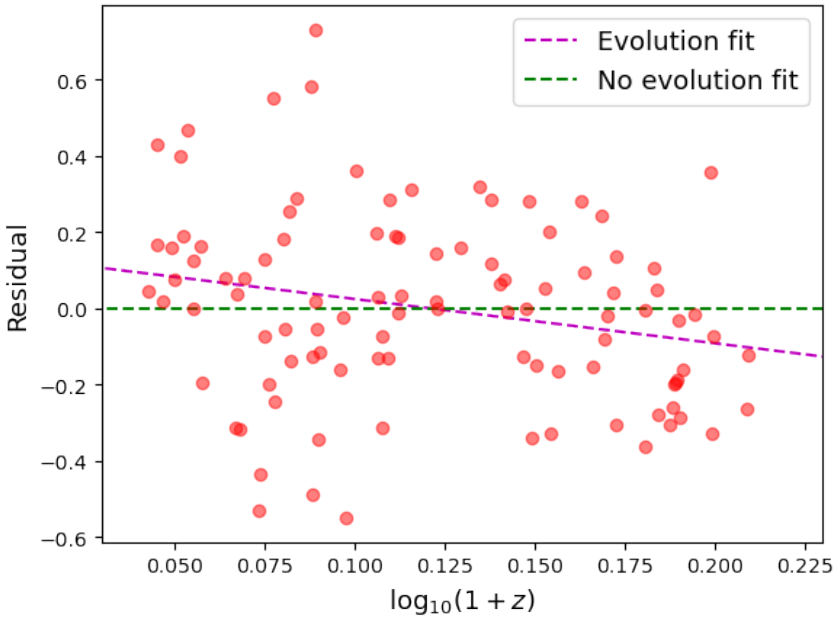


Figure 4.3: Residuals of the linear fit as a function of redshift. The two dashed lines represent the best fit for a single-parameter model (i.e. fixed slope to zero) of no temporal evolution and a two-parameter model of redshift evolution.

As for the second case, where we apply our Bayesian statistical approach neglecting all the selection functions and the HMF in the likelihood, we sampled our posterior distribution for the parameters using EMCEE with the same specifications as the main-fit analysis. We also used the uninformative priors discussed in section 3.2.2. The best parameter values are consistent with the full likelihood analysis at a 3σ level, with a small steepening of the slope when compared with the main-fit and with an even smaller value for γ.

For the last approach, still neglecting the selection functions but now including the HMF in the likelihood, we again find similar results as the ones from the main-fit, within a 3σ interval. The slope is steeper - although shallower when compared with fits (i) and (ii) - and again there is a moderately significant temporal evolution.

All the different fits - i.e. the main-fit using the full likelihood, (i) the oversimplified fit using the LMFIT.MINIMIZE function, (ii) the fit ignoring the selection functions and the

HMF in the likelihood, and (iii) the fit ignoring only the selection functions - are shown in figure 4.4. The best parameter values are displayed in table 4.2.

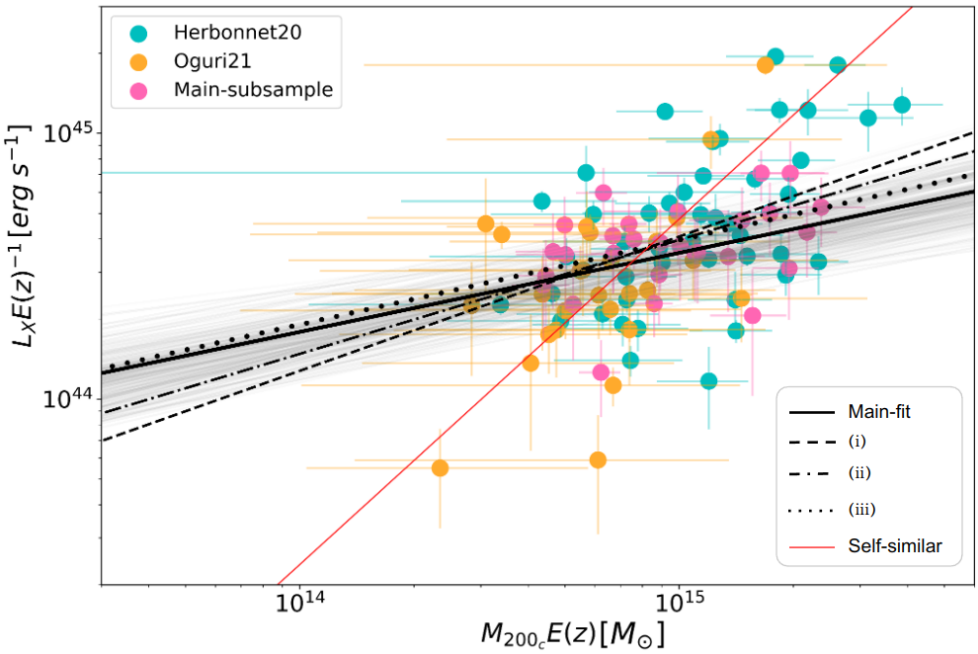


Figure 4.4: The *WL sample*, the main-fit, and the three fits discussed in this section to investigate the terms in the statistical model. The red line illustrates the slope predicted by self-similarity for comparison.

From this analysis, we can conclude that, with respect to the slope value, the inclusion of the selection functions and the HMF in our analysis have a small effect in flattening the $M - L_X$ scaling relation, with the latter being the most significant. This can be observed when comparing the slope values, which change from 0.43 to 0.32 when including the HMF, and then to 0.30 when including also the selection functions (full likelihood).

However, the statistical analysis developed throughout this work does not seem to be the main cause of the strong deviation from the self-similar model. This is evident in the results of the simple linear fit (i), which still deviates from the expected value of $\beta_{self} = 1.3$ by approximately 8.8σ. Such features suggest that our result is most affected by the sample choice, where we focus only on the most massive galaxy clusters, which will be further discussed in section 4.2.

As for the evolution parameter γ, although all fits were in agreement concerning the positive value for redshift evolution, we can observe that the inclusion of the HMF and the selection functions contribute to constraining the parameter. A comparison between (ii) and (iii) reveals that in the former (excluding the HMF and selections), the error is almost twice the measured value, whereas in the latter (considering only the HMF), the

error is constrained to be below the measured value. We can also compare the results from (iii) with the main-fit, which demonstrates that incorporating the selection functions in our analysis reduces the error by almost ten percentage points – from 44% to 34%.

4.1.2 *Contribution of each subsample*

We also aimed to verify the contribution of each subsample in the calibration of the $M - L_X$ relation. We would expect a significant impact of the new *main-subsample* in constraining the temporal evolution parameter γ since its redshift range is complementary to the other subsamples. On the other hand, the larger measurement errors in the mass estimates of *Oguri21* subsample could affect the constraints in the slope and intersection of the scaling relation.

To examine these questions - and perhaps unveil other unknown features - we applied our Bayesian statistical model with uninformative priors to 3 different sets of cluster samples: All *WL sample* but the *main-subsample*; all but the *Herbonnet20* subsample; and all but the *Oguri21* subsample;

The results for each of the three sets are listed in table 4.2 and figure 4.5 presents all the best fits in the $M - L_X$ log space, including the main-fit adopting the full *WL sample*.

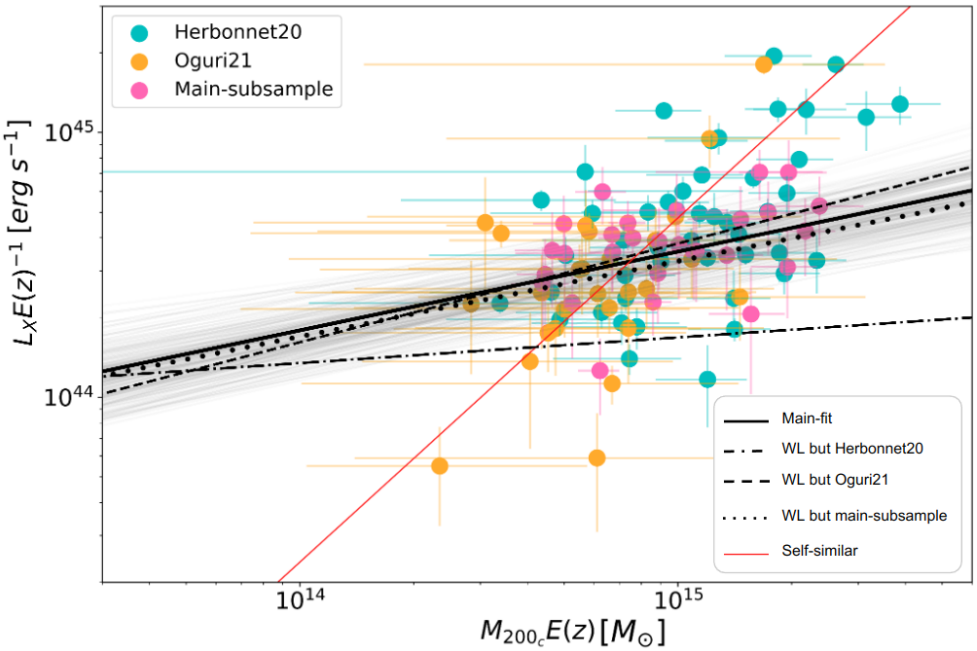


Figure 4.5: The *WL sample*, the main-fit, and the three fits discussed in this section to investigate the impact of each subsample in our analysis. The red line illustrates the slope predicted by self-similarity for comparison.

Contrary to what was expected, the inclusion of the *main-subsample* has no major impact in constraining the evolution parameter γ - when compared with the main-fit, it reduces the error by a single percentage point. Additionally, even though excluding the *Oguri21* subsample leads to an increase of the slope parameter as predicted, the difference to the results from the main-fit is under 1σ.

Upon analysis, it becomes evident that the *Herbonnet20* subsample has the most impact on our results. The slope parameter undergoes an almost 2σ change, and the error in the γ parameter differs by six percentage points. This is most likely due to the larger number of clusters in this subsample - 46 clusters, while the *main-subsample* and the *Oguri21* consist of 28 and 27 clusters, respectively.

4.1.3 Comparison with the literature

It is important to note that our result for the slope parameter is not consistent with previous works in the literature, as shown in figure 4.6. In the figure, we compare our values for slope β and the evolution parameter γ with the ones obtained by Leauthaud et al. (2010) and Kettula et al. (2015), who relied on weak lensing mass estimates, as well as with the outcomes of Lovisari et al. (2020) and Bulbul et al. (2019), who employed hydrostatic mass values. For the sake of coherence, our subsequent discussion focuses primarily on the first two studies.

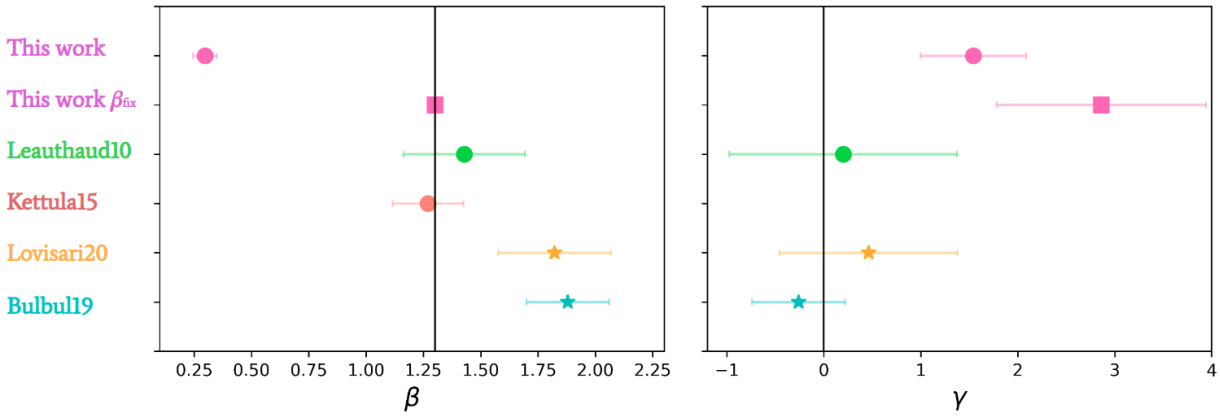


Figure 4.6: Values for the slope β and evolution parameter γ of the $M - L_X$ scaling relation. We compare the results obtained in this work (with and without a fixed slope value) with some of the ones available in the literature, i.e. Leauthaud et al. (2010); Kettula et al. (2015); Lovisari et al. (2020) and Bulbul et al. (2019). The vertical lines represent the self-similar prediction and the star marker differentiates the works that used hydrostatic mass estimates in their analysis.

In Leauthaud et al. (2010), the analysis is done using 206 galaxy groups (i.e. $M_{200c} \lesssim$

$10^{14} M_{\odot}$) of the COSMOS catalog, where the gravitational lens mass estimate was performed through stacking in 9 bins of X-ray luminosity and redshift. The sample employed in their analysis covers a wide range of mass and X-ray luminosity values, which is imperative for obtaining reliable slope estimates. Their work found a slope slightly steeper but still consistent with self-similarity, $\beta = 1.42 \pm 0.18$.

Furthermore, Leauthaud et al. (2010) has also analyzed the temporal evolution of the $L_X - M$ relation in this group regime and, making the proper correspondence to our parameter γ and propagating the errors accordingly, they found a value of $\gamma = 0.20 \pm 1.14$. Comparing this result with our work, we can see that, while we also found a positive value for γ, our study contributed to constraining the evolution parameter, decreasing even more the error value.

Kettula et al. (2015) also does a similar study of the $M - L_X$ relation with weak lensing measurements, using their newly proposed subsample of 12 intermediate-mass clusters from the CFHTLenS and XMM-CFHTLS surveys, together with the same data used in Leauthaud et al. (2010) and with additional 48 high-mass clusters from the Canadian Cluster Comparison Project (CCCP) (Hoekstra et al., 2015). Their results are the closest to the predicted by the self-similar model. Although they implement a Bayesian correction for the Eddington bias (Vikhlinin et al., 2009), they do not detail any other selection function or adopt a statistical analysis to account for the different samples used.

4.2 The slope parameter

In the previous sections, we have presented puzzling results for the $M - L_X$ calibration, where the slope significantly deviates from the self-similar prediction and early works. We have noted that, when compared with different studies, our work analyzes high-mass cluster samples with a larger number of data points. Moreover, it stands out for being the first one to carefully account for selection effects maintaining statistical consistency between all subsamples. Even so, it is important to conjecture what environmental processes in galaxy clusters could justify our results.

As was discussed in sections 1.6 and 1.3.1, one would expect the impact of non-gravitational effects to be significantly greater in low-mass clusters, such that it would be reasonable to suggest a different behavior of scaling relations in the galaxy group and

galaxy cluster regimes (Pratt et al., 2009; Kravtsov and Borgani, 2012; Kettula et al., 2015). For the $M - L_X$ relation, one would expect a steepening when studying low-mass systems, as was reported by the galaxy group analysis in Leauthaud et al. (2010). When including both groups and clusters, the slope value tends to align with the self-similar theory (Kettula et al., 2015), which could suggest that the high-mass end of the cluster population presents a shallower slope.

In addition, we are attempting to analyze the features of intricate systems like galaxy clusters through a straightforward relation between only two quantities (i.e. mass and X-ray luminosity), where we already have indications of hidden parameters. The study from Fujita and Aung (2019) for example, has shown the importance of the concentration parameter when calibrating scaling relations, as it correlates with both the mass and X-ray luminosity. These characteristics are thought to flatten the $M - L_X$ relation and increase the scatter in luminosity, such that, given our limited range in mass, this could present a significant effect on our results.

4.3 The evolution parameter

Even though our slope estimate disagrees with the self-similar prediction, we wanted to test whether or not we would find any temporal evolution trend if we constrained the β value. Indeed, there is a possibility that the γ estimate we found was perhaps a compensation for a misestimation of the slope. With this in mind, we sampled our posterior distribution just as before, changing only the slope prior into a more informative one; a Gaussian distribution centered in the self-similar value of 1.3, with a deviation of 0.15. The best fit is illustrated in figure 4.7 and the values of each parameter are displayed in table 4.2.

The γ values found in the main-fit and the fit with fixed slope deviate from the self-similar prediction by 2.8σ and 2.6σ, respectively. Hence, this new result also supports the statistically significant evidence of a mild temporal evolution in the scaling relation. All the different analyses done in this work present a positive value for the evolution parameter and, except for the fitting models (i) and (ii), all diverge from the non-evolution prediction by over 2σ. This suggests that there are hints of positive temporal evolution of $M - L_X$, i.e. for a fixed mass, one can expect a small increase in X-ray luminosity with redshift.

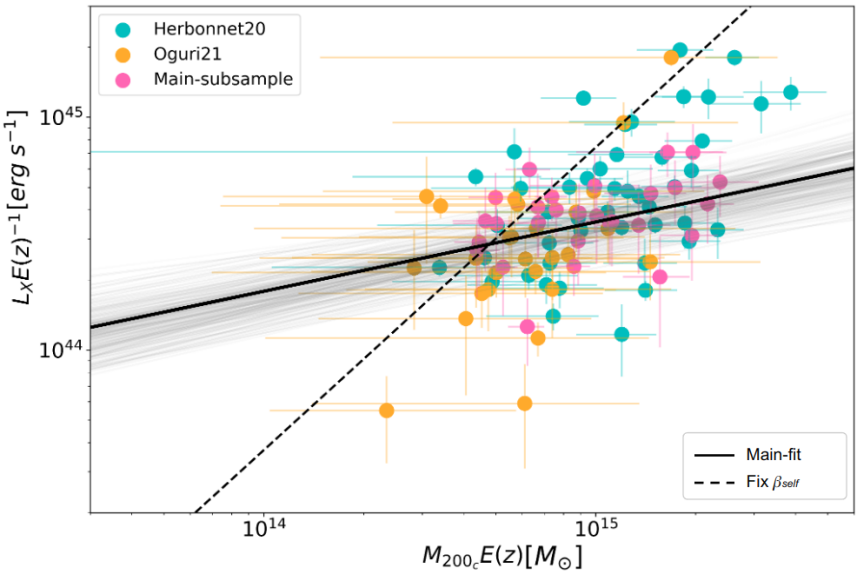


Figure 4.7: The *WL sample*, the main-fit, and the fit obtained by applying our full likelihood with a highly informative self-similar prior to the slope parameter.

From what was discussed in section 1.4.2.2, one possible physical explanation for this small decrease in X-ray luminosity at lower redshifts, given a fixed mass, would be the increase in the population of jet-mode AGNs in this regime. Although the feedback from AGNs is generally thought to balance the cooling flow in galaxy clusters, its effect might be larger than expected, causing a decrease in the expected X-ray luminosity.

We also acknowledge a potential bias in the γ parameter arising from our modeling of the X-ray luminosity distribution $P(\tilde{L}_X|L_X)$ as a Gaussian function, as described in Section 3.2. While assuming a normal distribution is reasonable for low redshifts, this assumption no longer holds for high values of z, where X-ray luminosity is expected to follow a Poisson distribution. The latter offers a more faithful representation of the high redshift regime where we have fewer photon counts. However, incorporating this model dependency with redshift would require complex changes in the construction of our likelihood function and is beyond the scope of this work.

4.4 Summary

This work analyses eight different sets of linear fitting for the calibration of the $M - L_X$ scaling relation and its evolution, generated through combinations between different fitting models and different samples. All the setups and results are presented in table 4.2.

Table 4.2 - Summary of all fits done in this work. The first column specifies the fitting method, whether it used a simple linear fit function, our full likelihood model, or some modification of it. More specifically, methods (i), (ii), and (iii) describe the linear fit function, the likelihood model without the HMF and the selection functions, and the likelihood model disregarding only the selection functions, respectively, as described in 4.1.1. The second column shows to which cluster sample the fitting was applied. The other columns present the best parameters found.

Fitting model	Sample	α	β	γ	σ_{intr}
Full likelihood	WL sample	40.1 ± 0.8	0.30 ± 0.05	1.54 ± 0.55	0.21 ± 0.01
Full likelihood	WL but *main-subsample*	40.2 ± 0.8	0.29 ± 0.06	2.11 ± 0.78	0.23 ± 0.02
Full likelihood	WL but *Herbonnet20*	42.8 ± 1.2	0.10 ± 0.09	2.16 ± 0.91	0.21 ± 0.01
Full likelihood	WL but *Oguri21*	39.0 ± 1.3	0.37 ± 0.09	1.41 ± 0.59	0.21 ± 0.01
(i)	WL sample	37.0 ± 1.3	0.50 ± 0.09	-	-
(ii)	WL sample	38.2 ± 1.2	0.43 ± 0.08	0.34 ± 0.62	0.22 ± 0.02
(iii)	WL sample	39.8 ± 0.9	0.32 ± 0.06	1.23 ± 0.54	0.23 ± 0.02
Fix β_{self}	WL sample	25.4 ± 0.2	1.3	2.86 ± 1.08	0.38 ± 0.05

From what was discussed and the results presented above, we can summarize our main discoveries and conclusions as:

- When using the full likelihood function to fit the *WL sample* we found a much shallower slope than the self-similar expectation. Moreover, although the γ parameter is consistent with the theory within 3σ, it is statistically significant to infer a minor positive evolution in the $M - L_X$ relation. Notably, compared to other previous works, our analysis contributed to constraining the γ parameter, reducing the error below the parameter value. The intrinsic scatter aligns with expectations for a high redshift and high mass sample.

- We also used the *WL sample* to calibrate the relation adopting different fitting models. When comparing the fit using the full likelihood with the fits with or without the selection functions or the HMF - fits (ii) and (iii) - we observe no strong difference between them. We can see, however, that the main effect of the selection functions and the HMF in the final result is a small flattening of the scaling relation, with the HMF being the most influential term.

- When using a Python function LMFIT.MINIMIZE for linear fitting, we still encounter a slope value smaller than the self-similar prediction. This suggests that the main cause for the strong deviation from self-similarity found in our results is not the statistical analysis, but whether the galaxy cluster sample analyzed or the presence of hidden properties not included in the scaling relation.

- By comparing the fits (ii), (iii), and the main-fit, we found that accounting for the HMF helps constrain the evolution parameter to an error below the measured value. We also observe that the inclusion of the selection functions decreases the error by almost ten percentage points.

- When analyzing the contribution of each subsample to our work, we observed that *Herbonnet20* subsample has the most impact on the results, most likely due to its larger number of clusters. We found no major contribution of the *main-subsample* for constraining γ and found the *Oguri21* subsample to have a small impact in reducing the slope value.

- To check if the γ value found in our main-fit was merely compensation from a misestimation of the slope, we also fitted the *WL sample* fixing the slope to the self-similar value. The value encountered differs from the no evolution statement at a 2.6σ confidence level, maintaining the moderately relevant evidence of a positive redshift evolution.

- When comparing our results to the two previously published works that used weak lensing mass estimates (i.e. Leauthaud et al., 2010; Kettula et al., 2015), we observe significant differences regarding the slope value. We point out that those works analyzed either only the galaxy group regime or the group plus cluster regime, and that the literature supports a difference in scaling relation between these two regimes. These considerations suggest that small slope values could be considered consistent for a high-mass cluster analysis.

- As for the γ parameter, the consistency of the results presented in this work suggests that there are hints of a small positive temporal evolution of the $M - L_X$ relation. We raise the conjecture that this feature might be caused by jet-mode AGN feedback processes, which are considered to be more frequent at low redshifts. We also point

out a possible evolution bias due to a redshift dependence of the X-ray luminosity distribution. Additionally, when compared with the studies of Leauthaud et al. (2010); Lovisari et al. (2020) and Bulbul et al. (2019), our work was the first to further constrain the parameter to errors below the measured value.

4.5 Perspectives

This work sought to study the $M - L_X$ relation and its temporal evolution for high-mass galaxy clusters in the CODEX catalog. To meet this goal, we constructed a detailed hierarchical Bayesian model that accounts for measurement errors, intrinsic scatter, and selection effects. Nonetheless, we have revealed unexpected outcomes regarding the slope and evolution parameter, which opens room for further investigations and refinements.

As mentioned in previous sections, our current likelihood function assumes a normal distribution for the X-ray luminosity throughout the entire redshift range, when in truth, it should have a temporal dependency that better represents the Poissonian behavior of the observed photon counts. By not considering this feature, we might be introducing a bias in the evolution parameter. We intend to correct this issue in future works, which would require deep knowledge of the X-ray luminosity behavior in our cluster sample, as well as modifications in the selection functions, especially in the simulated CODEX selection for RASS's images, detailed in section 3.2.1.

Another adjustment intended for future studies is to consider the correlation between the cluster's mass and richness. At present, our statistical analysis adopts the measured richness as the cluster's true richness, without accounting for measurement errors or intrinsic scatter. (Kiiveri et al., 2021) have already investigated the $\lambda - M$ relation for clusters in the CODEX catalog, which could be implemented in our study. Although incorporating this would introduce a new parameter for the likelihood integration, leading to an increase in computational time, it is essential for ensuring a statistically robust analysis.

In conclusion, when carefully accounting for uncertainties, biases, and selection effects, we get closer to revealing the astrophysical properties of our data. As we progress, the insights gained from this study may pave the way for future research pursuits, particularly in constraining the calibration and understanding the evolution of the $M - L_X$ relation, contributing to a more profound understanding of galaxy clusters.

Bibliography

Allen S. W., Evrard A. E., Mantz A. B., Cosmological Parameters from Observations of Galaxy Clusters, ARA&A, 2011, vol. 49, p. 409

Angelinelli M., Ettori S., Dolag K., Vazza F., Ragagnin A., Redshift evolution of the baryon and gas fraction in simulated groups and clusters of galaxies, A&A, 2023, vol. 675, p. A188

Ansarifard S., Rasia E., Biffi V., Borgani S., Cui W., De Petris M., Dolag K., Ettori S., Movahed S. M. S., Murante G., Yepes G., The Three Hundred Project: Correcting for the hydrostatic-equilibrium mass bias in X-ray and SZ surveys, A&A, 2020, vol. 634, p. A113

Arevalo F., Cid A., Moya J., AIC and BIC for cosmological interacting scenarios, European Physical Journal C, 2017, vol. 77, p. 565

Ascasibar Y., Sevilla R., Yepes G., Müller V., Gottlöber S., Adiabatic scaling relations of galaxy clusters, MNRAS, 2006, vol. 371, p. 193

Bartelmann M., Schneider P., Weak gravitational lensing, Phys. Rep., 2001, vol. 340, p. 291

Best P. N., Ker L. M., Simpson C., Rigby E. E., Sabater J., The cosmic evolution of radio-AGN feedback to z = 1, MNRAS, 2014, vol. 445, p. 955

Blumenthal G. R., Faber S. M., Primack J. R., Rees M. J., Formation of galaxies and large-scale structure with cold dark matter., Nature, 1984, vol. 311, p. 517

Braspenning J., Schaye J., Schaller M., McCarthy I. G., Kay S. T., Helly J. C., Kugel R., Elbers W., Frenk C. S., Kwan J., Salcido J., van Daalen M. P., Vandenbroucke B., The FLAMINGO Project: Galaxy clusters in comparison to X-ray observations, arXiv e-prints, 2023, p. arXiv:2312.08277

Briel U. G., Henry J. P., Lumb D. H., Arnaud M., Neumann D., Aghanim N., Gastaud R., Mittaz J. P. D., Sasseen T. P., Vestrand W. T., A mosaic of the Coma cluster of galaxies with XMM-Newton, A&A, 2001, vol. 365, p. L60

Esra Bulbul et al. X-Ray Properties of SPT-selected Galaxy Clusters at 0.2 ¡ z ¡ 1.5 Observed with XMM-Newton, ApJ, 2019, vol. 871, p. 50

Burns J. O., The Radio Properties of cD Galaxies in Abell Clusters. I. an X-ray Selected Sample, AJ, 1990, vol. 99, p. 14

Chen Y., Reiprich T. H., Böhringer H., Ikebe Y., Zhang Y. Y., Statistics of X-ray observables for the cooling-core and non-cooling core galaxy clusters, A&A, 2007a, vol. 466, p. 805

Chen Y., Reiprich T. H., Böhringer H., Ikebe Y., Zhang Y. Y., Statistics of X-ray observables for the cooling-core and non-cooling core galaxy clusters, A&A, 2007b, vol. 466, p. 805

Cibirka N., Um estudo de aglomerados de galáxias via lentes gravitacionais, São Paulo: Universidade de São Paulo, 2017, Ph.D. Thesis

N. Cibirka et al. CODEX weak lensing: concentration of galaxy clusters at z $\sim$ 0.5, MNRAS, 2017, vol. 468, p. 1092

Cowie L. L., Binney J., Radiative regulation of gas flow within clusters of galaxies: a model for cluster X-ray sources., ApJ, 1977, vol. 215, p. 723

Cox R., The algebra of probable inference. The John Hopkins Press, 1961

Damsted S., Finoguenov A., Clerc N., Davalgaitė I., Kirkpatrick C. C., Mamon G. A., Ider Chitham J., Kiiveri K., Comparat J., Collins C., CODEX: Role of velocity substructure in the scaling relations of galaxy clusters, A&A, 2023, vol. 676, p. A127

Dawson C., Introduction to Research Methods 5th Edition: A Practical Guide for Anyone Undertaking a Research Project. Little, Brown Book Group, 2019

Dutton A. A., Macciò A. V., Cold dark matter haloes in the Planck era: evolution of structural parameters for Einasto and NFW profiles, MNRAS, 2014, vol. 441, p. 3359

Eckmiller H. J., Hudson D. S., Reiprich T. H., Testing the low-mass end of X-ray scaling relations with a sample of Chandra galaxy groups, A&A, 2011, vol. 535, p. A105

Erlund M. C., Fabian A. C., Blundell K. M., Celotti A., Crawford C. S., Extended inverse-Compton emission from distant, powerful radio galaxies, MNRAS, 2006, vol. 371, p. 29

Fabian A. C., Observational Evidence of Active Galactic Nuclei Feedback, ARA&A, 2012, vol. 50, p. 455

Fabian A. C., Nulsen P. E. J., Subsonic accretion of cooling gas in clusters of galaxies., MNRAS, 1977, vol. 180, p. 479

Fabian A. C., Sanders J. S., Taylor G. B., Allen S. W., Crawford C. S., Johnstone R. M., Iwasawa K., A very deep Chandra observation of the Perseus cluster: shocks, ripples and conduction, MNRAS, 2006, vol. 366, p. 417

Fabozzi F. e. a., The Basics of Financial Econometrics. John Wiley & Sons, Inc., 2014

Feigelson E., Babu J., Modern Statistical Methods for Astronomy. Cambridge University Press, 2012

Finetti de B., La prévision : ses lois logiques, ses sources subjectives. Annales de l'Institut Henri Poincaré, 1937

A. Finoguenov et al. CODEX clusters. Survey, catalog, and cosmology of the X-ray luminosity function, A&A, 2020, vol. 638, p. A114

Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., emcee: The MCMC Hammer, PASP, 2013, vol. 125, p. 306

Fujita Y., Aung H., Halo Concentrations and the New Baseline X-Ray Luminosity–Temperature and Mass Relations of Galaxy Clusters, The Astrophysical Journal, 2019, vol. 875, p. 26

Gelman A., Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper), Bayesian Analysis, 2006, vol. 1, p. 515

Gianfagna G., De Petris M., Yepes G., De Luca F., Sembolini F., Cui W., Biffi V., Kéruzoré F., Macías-Pérez J., Mayet F., Perotto L., Rasia E., Ruppin F., Exploring the hydrostatic mass bias in MUSIC clusters: application to the NIKA2 mock sample, MNRAS, 2021, vol. 502, p. 5115

Goodman J., Weare J., Ensemble samplers with affine invariance, Communications in Applied Mathematics and Computational Science, 2010, vol. 5, p. 65

Gregory P., Bayesian Logical Data Analysis for the Physical Sciences. Cambridge University Press, 2005

Gruen D., Brimioulle F., Selection biases in empirical p(z) methods for weak lensing, MNRAS, 2017, vol. 468, p. 769

Gruen D., Seitz S., Becker M. R., Friedrich O., Mana A., Cosmic variance of the galaxy cluster weak lensing signal, MNRAS, 2015, vol. 449, p. 4264

Heckman T. M., Best P. N., The Coevolution of Galaxies and Supermassive Black Holes: Insights from Surveys of the Contemporary Universe, ARA&A, 2014, vol. 52, p. 589

Herbonnet R., Sifón C., Hoekstra H., Bahé Y., van der Burg R. F. J., Melin J.-B., von der Linden A., Sand D., Kay S., Barnes D., CCCP and MENeaCS: (updated) weak-lensing masses for 100 galaxy clusters, MNRAS, 2020, vol. 497, p. 4684

Hoekstra H., Herbonnet R., Muzzin A., Babul A., Mahdavi A., Viola M., Cacciato M., The Canadian Cluster Comparison Project: detailed study of systematics and updated weak lensing masses, MNRAS, 2015, vol. 449, p. 685

Hubble E., A Relation between Distance and Radial Velocity among Extra-Galactic Nebulae, Proceedings of the National Academy of Science, 1929, vol. 15, p. 168

Hubble E. P., Cepheids in spiral nebulae, The Observatory, 1925, vol. 48, p. 139

Hubble E. P., Extragalactic nebulae., ApJ, 1926, vol. 64, p. 321

Ivezić Ž., Connolly A. J., VanderPlas J. T., Gray A., Statistics, Data Mining, and Machine Learning in Astronomy. A Practical Python Guide for the Analysis of Survey Data, Updated Edition, 2020

Jaynes E., Probability, Frequency and Reasonable Expectation, Report 421, Microwave Laboratory, Stanford University (published 1988, in Maximum entropy and Bayesian methods in science and engineering, Vol. 1, Erickson, G.J., Smith, C.R. eds, Dordrecht: Kluwer Academic Pub-lishers, pp. 1-24)., 1957

Kaiser N., Evolution and clustering of rich clusters., MNRAS, 1986, vol. 222, p. 323

Kelly B. C., Some Aspects of Measurement Error in Linear Regression of Astronomical Data, ApJ, 2007, vol. 665, p. 1489

K. Kettula et al. CFHTLenS: weak lensing calibrated scaling relations for low-mass clusters of galaxies, MNRAS, 2015, vol. 451, p. 1460

K. Kiiveri et al. CODEX weak lensing mass catalogue and implications on the mass-richness relation, MNRAS, 2021, vol. 502, p. 1494

M. Klein et al. A new RASS galaxy cluster catalogue with low contamination extending to z $\sim$ 1 in the DES overlap region, MNRAS, 2019, vol. 488, p. 739

Kravtsov A. V., Borgani S., Formation of Galaxy Clusters, ARA&A, 2012, vol. 50, p. 353

C. Laigle et al. The COSMOS2015 Catalog: Exploring the $1 < z < 6$ Universe with Half a Million Galaxies, ApJS, 2016, vol. 224, p. 24

Alexie Leauthaud et al. A Weak Lensing Study of X-ray Groups in the Cosmos Survey: Form and Evolution of the Mass-Luminosity Relation, ApJ, 2010, vol. 709, p. 97

Lemaître G., Un Univers homogène de masse constante et de rayon croissant rendant compte de la vitesse radiale des nébuleuses extra-galactiques, Annales de la Société Scientifique de Bruxelles, 1927, vol. 47, p. 49

Lima Neto G., Notas de Aula: Astronomia Extragalática e Cosmologia, 2020

Limber D. N., The Analysis of Counts of the Extragalactic Nebulae in Terms of a Fluctuating Density Field. II., ApJ, 1954, vol. 119, p. 655

Lovisari L., Schellenberger G., Sereno M., Ettori S., Pratt G. W., Forman W. R., Jones C., Andrade-Santos F., Randall S., Kraft R., X-Ray Scaling Relations for a Representative Sample of Planck-selected Clusters Observed with XMM-Newton, ApJ, 2020, vol. 892, p. 102

Lynds R., Petrosian V., Luminous Arcs in Clusters of Galaxies, ApJ, 1989, vol. 336, p. 1

Markevitch M., The L_X-T Relation and Temperature Function for Nearby Clusters Revisited, ApJ, 1998, vol. 504, p. 27

Maughan B. J., Giles P. A., Randall S. W., Jones C., Forman W. R., Self-similar scaling and evolution in the galaxy cluster X-ray luminosity-temperature relation, MNRAS, 2012, vol. 421, p. 1583

McCarthy I. G., Schaye J., Ponman T. J., Bower R. G., Booth C. M., Dalla Vecchia C., Crain R. A., Springel V., Theuns T., Wiersma R. P. C., The case for AGN feedback in galaxy groups, MNRAS, 2010, vol. 406, p. 822

McNamara B. R., Nulsen P. E. J., Heating Hot Atmospheres with Active Galactic Nuclei, ARA&A, 2007, vol. 45, p. 117

Meneghetti M., Introduction to Gravitational Lensing: With Python Examples, 2022

L. Miller et al. Bayesian galaxy shape measurement for weak lensing surveys - III. Application to the Canada-France-Hawaii Telescope Lensing Survey, MNRAS, 2013, vol. 429, p. 2858

Mo H., van den Bosch F., White S., Galaxy Formation and Evolution. Cambridge University Press, 2020

Nagai D., The Impact of Galaxy Formation on the Sunyaev-Zel'dovich Effect of Galaxy Clusters, ApJ, 2006, vol. 650, p. 538

Nagai D., Vikhlinin A., Kravtsov A. V., Testing X-Ray Measurements of Galaxy Clusters with Cosmological Simulations, ApJ, 2007, vol. 655, p. 98

Narayan R., Bartelmann M., Lectures on Gravitational Lensing, arXiv e-prints, 1996, pp astro–ph/9606001

Navarro J. F., Frenk C. S., White S. D. M., The Structure of Cold Dark Matter Halos, ApJ, 1996, vol. 462, p. 563

Navarro J. F., Frenk C. S., White S. D. M., A Universal Density Profile from Hierarchical Clustering, ApJ, 1997, vol. 490, p. 493

Oguri M., Miyazaki S., Li X., Luo W., Mitsuishi I., Miyatake H., More S., Nishizawa A. J., Okabe N., Ota N., Plazas Malagón A. A., Utsumi Y., Hundreds of weak lensing shear-selected clusters from the Hyper Suprime-Cam Subaru Strategic Program S19A data, PASJ, 2021, vol. 73, p. 817

Padovani P., High energy emission from AGN and unified schemes. In Vulcano Workshop 1998: Frontier Objects in Astrophysics and Particle Physics , vol. 65, 1999, p. 159

Peterson J. R., Kahn S. M., Paerels F. B. S., Kaastra J. S., Tamura T., Bleeker J. A. M., Ferrigno C., Jernigan J. G., High-Resolution X-Ray Spectroscopic Constraints on Cooling-Flow Models for Clusters of Galaxies, ApJ, 2003, vol. 590, p. 207

Planelles S., Schleicher D. R. G., Bykov A. M., Large-Scale Structure Formation: From the First Non-linear Objects to Massive Galaxy Clusters, Space Sci. Rev., 2015, vol. 188, p. 93

Pratt G. W., Croston J. H., Arnaud M., Böhringer H., Galaxy cluster X-ray luminosity scaling relations from a representative local sample (REXCESS), A&A, 2009, vol. 498, p. 361

Press W. H., Schechter P., Formation of Galaxies and Clusters of Galaxies by Self-Similar Gravitational Condensation, ApJ, 1974, vol. 187, p. 425

Rich J., Fundamentals of Cosmology. Springer-Verlag, 2010

F. Ruppin et al. Redshift Evolution of the Feedback-Cooling Equilibrium in the Core of 48 SPT Galaxy Clusters: A Joint Chandra-SPT-ATCA Analysis, ApJ, 2023, vol. 948, p. 49

Rybicki G. L. A., Radiative Processes in Astrophysics. Wiley-VCH, 1985

Rykoff E. S., Rozo E., Busha M. T., Cunha C. E., Finoguenov A., Evrard A., Hao J., Koester B. P., Leauthaud A., Nord B., Pierre M., Reddick R., Sadibekova T., Sheldon E. S., Wechsler R. H., redMaPPer. I. Algorithm and SDSS DR8 Catalog, ApJ, 2014, vol. 785, p. 104

Sarazin C. L., X-ray emission from clusters of galaxies. Cambridge University Press, 1988

Schneider P., Extragalactic Astronomy and Cosmology. Springer, 2015

Morag I. Scrimgeour et al. The WiggleZ Dark Energy Survey: the transition to large-scale cosmic homogeneity, MNRAS, 2012, vol. 425, p. 116

Sereno M., A Bayesian approach to linear regression in astronomy, MNRAS, 2016, vol. 455, p. 2149

Smith S., The Mass of the Virgo Cluster, ApJ, 1936, vol. 83, p. 23

Tinker J., Kravtsov A. V., Klypin A., Abazajian K., Warren M., Yepes G., Gottlöber S., Holz D. E., Toward a Halo Mass Function for Precision Cosmology: The Limits of Universality, ApJ, 2008, vol. 688, p. 709

Veilleux S., Cecil G., Bland-Hawthorn J., Galactic Winds, ARA&A, 2005, vol. 43, p. 769

Vikhlinin A., Burenin R. A., Ebeling H., Forman W. R., Hornstrup A., Jones C., Kravtsov A. V., Murray S. S., Nagai D., Quintana H., Voevodkin A., Chandra Cluster Cosmology Project. II. Samples and X-Ray Data Reduction, ApJ, 2009, vol. 692, p. 1033

Vikhlinin A., Kravtsov A., Forman W., Jones C., Markevitch M., Murray S. S., Van Speybroeck L., Chandra Sample of Nearby Relaxed Galaxy Clusters: Mass, Gas Fraction, and Mass-Temperature Relation, ApJ, 2006, vol. 640, p. 691

Vikhlinin A., McNamara B. R., Forman W., Jones C., Quintana H., Hornstrup A., A Catalog of 200 Galaxy Clusters Serendipitously Detected in the ROSAT PSPC Pointed Observations, ApJ, 1998, vol. 502, p. 558

W. Voges et al. The ROSAT all-sky survey bright source catalogue, A&A, 1999, vol. 349, p. 389

Voit G. M., Tracing cosmic evolution with clusters of galaxies, Reviews of Modern Physics, 2005, vol. 77, p. 207

Walter G. A. T., Bayesian Linear Regression — Different Conjugate Models and Their (In)Sensitivity to Prior-Data Conflict. In: Kneib, T., Tutz, G. (eds) Statistical Modelling and Regression Structures. Physica-Verlag HD, 2010

Wambsganss J., Gravitational Lensing in Astronomy, Living Reviews in Relativity, 1998, vol. 1, p. 12

White S. D. M., Rees M. J., Core condensation in heavy halos: a two-stage theory for galaxy formation and clustering., MNRAS, 1978, vol. 183, p. 341

Witte R., Witte J., Statistics. Wiley, 2017

Donald G. York et al. The Sloan Digital Sky Survey: Technical Summary, AJ, 2000, vol. 120, p. 1579

Zwicky F., Die Rotverschiebung von extragalaktischen Nebeln, Helvetica Physica Acta, 1933, vol. 6, p. 110

Zwicky F., On the Masses of Nebulae and of Clusters of Nebulae, ApJ, 1937, vol. 86, p. 217

Appendix

Appendix A

CODEX galaxy cluster subsamples

Table A.1 - The CODEX-CFHT weak lensing sample of 35 galaxy clusters. The first 28 clusters pass our selection and comprise the *Main-subsample*. The last seven clusters do not pass the selection.

Cluster ID	z_{spec}	$< \log_{10} M_{200c} >$	$med(\log_{10} M_{200c})$	L_X	λ_{SDSS}
		$M_\odot$	$M_\odot$	$[10^{44}\ h_{70}^{-2}\ erg\ s^{-1}]$	
12451	0.5840	14.67	$14.8^{+0.2}_{-0.3}$	5.3 ± 1.8	68.6 ± 3.9
13390	0.623	14.80	$15.0^{+0.3}_{-0.4}$	9.9 ± 2.1	120.6 ± 5.8
16566	0.382	14.53	$14.6^{+0.1}_{-0.3}$	2.8 ± 1.1	108 ± 7
24865	0.486	14.84	$14.9^{+0.1}_{-0.2}$	4.6 ± 1.6	138 ± 23
24872	0.402	14.64	$14.7^{+0.2}_{-0.3}$	4.9 ± 1.2	149 ± 10
24925	0.2920	14.81	$14.9^{+0.2}_{-0.2}$	2.6 ± 0.6	64.8 ± 3.5
24981	0.411	14.56	$14.6^{+0.2}_{-0.3}$	7.4 ± 1.8	123 ± 12
25953	0.478	14.61	$14.6^{+0.2}_{-0.2}$	4.6 ± 1.3	131 ± 19
29283	0.549	14.93	$15.0^{+0.1}_{-0.3}$	6.3 ± 2.1	129 ± 30
29284	0.550	14.19	$14.3^{+0.3}_{-0.6}$	4.8 ± 1.9	122 ± 33
29811	0.488	15.18	$15.2^{+0.1}_{-0.2}$	5.5 ± 1.8	143.8 ± 4.9
35361	0.414	14.70	$14.7^{+0.2}_{-0.1}$	5.7 ± 1.2	103 ± 9
35399	0.516	14.80	$14.8^{+0.2}_{-0.2}$	5.0 ± 1.9	153 ± 31
36818	0.581	14.98	$15.0^{+0.2}_{-0.2}$	2.8 ± 1.4	104.5 ± 5.2
41843	0.434	14.28	$14.4^{+0.2}_{-0.5}$	3.7 ± 1.3	119 ± 13
41911	0.386	14.78	$14.8^{+0.2}_{-0.2}$	3.6 ± 1.3	104 ± 7
43403	0.422	14.89	$14.9^{+0.1}_{-0.2}$	4.5 ± 1.6	130 ± 10
46647	0.261	14.6	$14.7^{+0.2}_{-0.3}$	1.4 ± 0.5	66.2 ± 2.6
46649	0.619	15.08	$15.1^{+0.1}_{-0.2}$	9.9 ± 3.2	85 ± 31
47981	0.543	14.54	$14.7^{+0.3}_{-0.6}$	6.8 ± 2.5	136 ± 33
50492	0.527	15.19	$15.2^{+0.1}_{-0.1}$	7.0 ± 2.0	163 ± 30
50514	0.466	14.23	$14.3^{+0.3}_{-0.5}$	3.4 ± 1.3	82 ± 13
52480	0.565	15.00	$15.0^{+0.2}_{-0.2}$	6.8 ± 2.1	106 ± 54
54795	0.428	14.34	$14.4^{+0.3}_{-0.4}$	5.7 ± 1.6	125 ± 35
55181	0.547	14.51	$14.6^{+0.2}_{-0.4}$	5.5 ± 2.2	149 ± 43
56934	0.459	14.93	$15.0^{+0.2}_{-0.3}$	4.4 ± 1.6	79.6 ± 4.5

Continue in next page...

Table A.1 - Continued

Cluster ID	z_{spec}	$< \log_{10} M_{200c} >$	$med(\log_{10} M_{200c})$	L_X	λ_{SDSS}
		$M_\odot$	$M_\odot$	$[10^{44}\ h_{70}^{-2}\ erg\ s^{-1}]$	
59915	0.475	15.14	$15.2^{+0.1}_{-0.2}$	4.0 ± 1.4	143 ± 25
64232	0.529	14.19	$14.3^{+0.4}_{-0.7}$	4.7 ± 1.8	112 ± 37
24877	0.592	15.23	$15.2^{+0.2}_{-0.1}$	3.9 ± 1.6	44.3 ± 4.4
25424	0.509	14.37	$14.4^{+0.2}_{-0.3}$	5.3 ± 2.1	59.7 ± 2.9
25428	0.209	13.5	-	6.0 ± 2.5	3.6 ± 1.3
27974	0.475	14.67	$14.7^{+0.1}_{-0.2}$	8.1 ± 3.6	37.1 ± 6.2
54796	0.159	14.14	$14.3^{+0.3}_{-0.4}$	2.7 ± 0.4	30.1 ± 2.3
54799	0.418	14.3	-	1.7 ± 0.9	5.4 ± 2.9
55184	0.445	14.3	-	4.0 ± 2.6	3.8 ± 1.4

Table A.2 - The 83 CODEX galaxy clusters in Herbonnet et al. (2020). The first 46 clusters pass our selection and comprise the *Herbonnet20* subsample. The last 37 clusters do not pass the selection.

Cluster ID	z	M_{200c}	L_X	λ_{SDSS}
		$10^{14} M_\odot$	$[10^{44}\ h_{70}^{-2}\ erg\ s^{-1}]$	
6588	0.173344	13.3 ± 4.0	4.485 ± 0.215	147.794 ± 5.338
9037	0.288085	16.5 ± 4.0	3.390 ± 0.551	107.109 ± 3.818
9481	0.327063	11.4 ± 2.6	5.406 ± 0.688	112.983 ± 4.442
12734	0.136272	11.2 ± 2.7	3.589 ± 0.277	80.169 ± 3.934
16819	0.140942	14.1 ± 3.4	3.693 ± 0.309	68.911 ± 6.436
17346	0.278227	15.0 ± 3.1	5.794 ± 0.700	139.679 ± 4.325
20505	0.128743	5.9 ± 2.4	2.225 ± 0.291	112.159 ± 7.218
20766	0.407627	18.8 ± 3.7	4.073 ± 1.020	138.031 ± 5.001
22133	0.225804	8.2 ± 2.1	13.520 ± 0.733	195.604 ± 6.526
24614	0.544257	28.8 ± 8.1	17.200 ± 2.831	254.152 ± 8.481
25247	0.204750	10.3 ± 2.6	5.468 ± 0.624	77.789 ± 5.370
25809	0.232092	14.1 ± 3.8	7.564 ± 0.651	200.012 ± 6.177
26063	0.167284	11.3 ± 2.8	10.100 ± 0.569	106.759 ± 4.006
27969	0.103536	4.4 ± 2.3	2.617 ± 0.262	67.538 ± 4.629
29227	0.170439	4.0 ± 2.3	6.033 ± 0.550	91.809 ± 5.673
29249	0.286628	7.2 ± 2.4	5.797 ± 0.905	92.144 ± 4.678
29315	0.281980	11.1 ± 3.9	11.032 ± 1.490	136.063 ± 4.952
29657	0.126764	7.0 ± 2.6	1.485 ± 0.194	62.733 ± 3.693
29715	0.122491	8.5 ± 3.6	3.453 ± 0.304	62.367 ± 3.760
30718	0.228355	7.9 ± 2.7	4.124 ± 0.636	124.958 ± 6.199
31149	0.225728	18.7 ± 4.3	8.851 ± 0.627	170.341 ± 6.440

Continue in next page...

Table A.2 - Continued

Cluster ID	z	M_{200c}	L_X	λ_{SDSS}
		$10^{14} M_\odot$	$[10^{44}\ h_{70}^{-2}\ erg\ s^{-1}]$	
32750	0.191902	5.4 ± 2.5	5.438 ± 0.664	129.214 ± 5.314
35445	0.114230	3.2 ± 2.2	2.392 ± 0.208	86.607 ± 4.290
35646	0.426069	14.8 ± 3.6	4.398 ± 0.998	124.500 ± 5.884
35839	0.159558	6.7 ± 2.4	3.116 ± 0.276	70.497 ± 4.537
39497	0.142196	10.8 ± 3.1	7.399 ± 0.500	115.912 ± 6.153
41892	0.545885	23.4 ± 7.4	15.367 ± 3.896	140.075 ± 5.887
45788	0.230294	16.4 ± 3.0	13.758 ± 1.472	156.047 ± 7.816
45867	0.109716	13.4 ± 3.3	1.901 ± 0.190	78.369 ± 4.047
46421	0.250618	11.0 ± 2.6	5.460 ± 1.245	96.034 ± 3.638
48016	0.167845	10.0 ± 2.4	4.246 ± 0.469	177.160 ± 5.392
49187	0.516370	4.3 ± 4.3	9.394 ± 2.461	61.039 ± 5.674
53349	0.252525	15.8 ± 4.1	22.090 ± 1.762	167.809 ± 8.460
53527	0.225154	10.7 ± 2.9	1.307 ± 0.446	80.657 ± 3.575
55475	0.229432	6.4 ± 2.5	4.389 ± 0.631	94.280 ± 4.337
57879	0.185832	23.9 ± 4.4	19.791 ± 1.590	184.055 ± 5.524
58218	0.120026	4.6 ± 2.3	2.083 ± 0.275	66.424 ± 4.204
59487	0.295476	16.7 ± 3.5	6.863 ± 1.091	139.008 ± 4.165
60202	0.189176	4.6 ± 2.6	3.767 ± 0.647	61.161 ± 4.304
60634	0.540393	16.3 ± 4.5	16.396 ± 3.264	158.808 ± 5.674
64520	0.2080	6.4 ± 2.7	2.114 ± 0.365	103.479 ± 6.664
64521	0.2130	7.0 ± 2.5	2.050 ± 0.367	103.311 ± 3.188
64532	0.208744	8.5 ± 2.4	6.052 ± 0.629	169.655 ± 5.368
65130	0.247569	9.1 ± 4.0	6.793 ± 0.912	116.140 ± 5.219
66237	0.3062	12.0 ± 3.0	2.766 ± 0.755	104.760 ± 9.506
70574	0.110	6.9 ± 2.6	2.491 ± 0.264	60.704 ± 3.970
32567	0.094953	6.1 ± 2.5	1.616 ± 0.183	80.474 ± 4.364
64209	0.0562	8.4 ± 3.3	8.38 ± 2.97	71.06 ± 4.11
55260	0.0370	7.8 ± 3.2	7.89 ± 0.49	62.93 ± 3.54
72640	0.0564	4.1 ± 2.7	2.43 ± 0.15	44.38 ± 3.02
20507	0.1269	3.8 ± 2.4	3.63 ± 0.36	44.81 ± 3.70
65975	0.0652	6.5 ± 3.0	1.40 ± 0.06	38.00 ± 3.57
43549	0.1405	1.6 ± 0.41	3.43 ± 0.35	112.21 ± 12.67
25322	0.1474	5.0 ± 2.5	6.07 ± 0.54	41.24 ± 2.86
57829	0.0956	1.05 ± 0.29	5.69 ± 0.39	88.16 ± 7.37
62335	0.0869	8.3 ± 3.5	5.88 ± 0.37	82.33 ± 8.42
35088	0.0580	0.15 ± 0.15	1.60 ± 0.14	37.99 ± 2.68
35137	0.0645	1.39 ± 0.33	18.65 ± 5.86	60.00 ± 5.64
35342	0.0890	4.4 ± 2.3	1.46 ± 0.14	39.27 ± 4.03
44634	0.0586	3.7 ± 2.9	1.14 ± 0.10	49.72 ± 3.51
53574	0.0783	1.81 ± 0.38	1.92 ± 0.64	11.48 ± 3.31
53571	0.0760	3.2 ± 2.4	1.42 ± 0.13	63.37 ± 4.12
53572	0.0960	2.9 ± 2.2	2.74 ± 0.25	40.78 ± 3.77
17773	0.0865	2.5 ± 2.4	1.31 ± 0.14	22.61 ± 2.69

Continue in next page...

Table A.2 - Continued

Cluster ID	z	M_{200c} $10^{14} M_{\odot}$	L_X $[10^{44}\ h_{70}^{-2}\ erg\ s^{-1}]$	λ_{SDSS}
35471	0.0668	1.2 ± 0.31	2.68 ± 0.17	110.86 ± 4.95
35543	0.0912	1.45 ± 0.34	2.02 ± 0.64	154.77 ± 6.07
68066	0.0981	8.4 ± 3.0	3.92 ± 0.33	98.63 ± 4.40
68049	0.1031	6.8 ± 2.8	4.08 ± 0.39	44.12 ± 3.75
50350	0.0925	0.21 ± 0.21	2.61 ± 0.22	32.30 ± 3.46
68241	0.0929	3.9 ± 2.4	9.00 ± 0.56	32.80 ± 4.01
41713	0.1203	3.0 ± 2.3	2.07 ± 0.21	42.84 ± 3.50
68307	0.0743	8.8 ± 3.2	1.42 ± 0.15	52.33 ± 3.75
60040	0.1515	2.6 ± 2.1	2.23 ± 0.30	19.62 ± 2.16
24691	0.1171	2.3 ± 2.3	2.91 ± 0.25	34.28 ± 3.04
64081	0.0886	3.8 ± 2.7	2.42 ± 0.20	50.47 ± 3.08
53063	0.0910	5.1 ± 3.1	4.18 ± 0.25	83.02 ± 7.24
40075	0.2566	1.01 ± 0.21	2.10 ± 0.42	53.09 ± 2.86
26180	0.3865	0.38 ± 0.25	2.50 ± 0.58	48.01 ± 4.29
39357	0.1668	6.90 ± 2.60	2.60 ± 0.28	51.12 ± 2.78
39356	0.1944	6.10 ± 2.50	2.50 ± 0.38	151.94 ± 16.25
53810	0.1409	1.68 ± 0.33	3.30 ± 0.47	100.24 ± 12.79
55927	0.1998	3.10 ± 0.23	2.30 ± 0.33	39.47 ± 5.33
17449	0.4485	8.70 ± 3.80	3.80 ± 0.36	45.59 ± 3.62

Table A.3 - The 84 CODEX galaxy clusters in Oguri et al. (2021). The first 27 clusters pass our selection and comprise the *Oguri21* subsample. The last 57 clusters do not pass the selection.

Cluster ID	z	M_{200c} $10^{14} M_{\odot}$	L_X $[10^{44}\ h_{70}^{-2}\ erg\ s^{-1}]$	λ_{SDSS}
22080	0.5243	$8.18^{+11.61}_{-4.65}$	4.42 ± 1.26	82.92 ± 4.87
52342	0.3267	$6.23^{+7.33}_{-5.21}$	2.97 ± 1.02	109.38 ± 5.18
52344	0.2032	$5.96^{+7.20}_{-4.70}$	2.41 ± 0.48	83.68 ± 4.19
52413	0.2777	$5.08^{+6.07}_{-4.11}$	4.84 ± 0.85	84.82 ± 6.61
52418	0.3477	$6.87^{+8.08}_{-5.71}$	3.08 ± 0.90	104.67 ± 4.47
52985	0.1358	$6.20^{+7.09}_{-5.33}$	3.56 ± 0.33	99.13 ± 4.35
54774	0.2974	$3.74^{+4.66}_{-2.91}$	2.88 ± 1.12	60.38 ± 4.47
55040	0.3742	$5.07^{+7.81}_{-2.46}$	2.98 ± 0.94	68.06 ± 2.91
55515	0.1968	$3.09^{+3.84}_{-2.42}$	4.58 ± 0.54	64.35 ± 2.87
56923	0.1954	$2.13^{+3.09}_{-1.18}$	0.61 ± 0.25	62.14 ± 4.23
57004	0.3276	$4.65^{+5.64}_{-3.69}$	3.59 ± 0.86	112.65 ± 6.09

Continue in next page...

Table A.3 - Continued

Cluster ID	z	M_{200c}	L_X	λ_{SDSS}
		$10^{14} M_\odot$	$[10^{44}\ h_{70}^{-2}\ erg\ s^{-1}]$	
57070	0.3643	$121.25^{+139.14}_{-104.19}$	2.88 ± 0.91	122.91 ± 5.04
57132	0.3736	$6.09^{+7.82}_{-4.43}$	2.22 ± 0.85	68.93 ± 3.54
57165	0.4563	$3.19^{+4.43}_{-2.04}$	1.74 ± 0.92	65.89 ± 3.82
57323	0.4799	$2.19^{+4.44}_{-0.64}$	2.92 ± 1.34	72.33 ± 3.86
57507	0.5822	$87.49^{+106.12}_{-69.86}$	13.12 ± 2.92	109.76 ± 5.99
57654	0.2282	$5.44^{+6.61}_{-4.20}$	0.66 ± 0.31	75.07 ± 2.72
57879	0.1843	$15.39^{+16.81}_{-14.05}$	19.79 ± 1.59	184.05 ± 5.52
58013	0.1323	$6.26^{+7.25}_{-5.32}$	1.20 ± 0.20	86.59 ± 4.57
58014	0.5540	$4.22^{+6.24}_{-2.41}$	6.01 ± 2.18	72.31 ± 3.72
58106	0.4053	$4.51^{+6.86}_{-2.95}$	3.78 ± 1.26	66.16 ± 3.39
58165	0.2953	$4.07^{+5.79}_{-2.55}$	2.12 ± 0.71	63.37 ± 3.87
59444	0.3884	$7.13^{+8.89}_{-5.10}$	4.79 ± 1.56	80.70 ± 6.03
59524	0.4882	$2.38^{+2.96}_{-1.80}$	5.92 ± 2.87	77.87 ± 4.49
59533	0.2817	$8.56^{+9.87}_{-7.25}$	5.55 ± 0.88	124.32 ± 4.23
59607	0.2771	$3.95^{+5.21}_{-2.84}$	2.01 ± 0.58	88.99 ± 3.92
60155	0.1885	$4.57^{+5.22}_{-3.94}$	2.36 ± 0.52	85.36 ± 3.36
21636	0.5003	$2.90^{+4.25}_{-1.75}$	10.68 ± 2.06	55.57 ± 3.77
21642	0.2261	$1.68^{+2.16}_{-1.22}$	1.50 ± 0.32	29.30 ± 1.94
21643	0.2940	$5.67^{+7.58}_{-3.77}$	1.16 ± 0.36	42.25 ± 3.01
21644	0.0946	$0.97^{+1.38}_{-0.58}$	1.40 ± 0.12	27.14 ± 2.03
21765	0.2721	$2.75^{+3.36}_{-2.17}$	1.94 ± 0.42	32.30 ± 2.60
21887	0.2543	$2.95^{+3.89}_{-2.14}$	0.83 ± 0.26	45.85 ± 2.50
21908	0.3834	$3.46^{+4.33}_{-2.67}$	1.13 ± 0.50	51.52 ± 3.40
21909	0.2294	$3.76^{+4.66}_{-2.79}$	1.56 ± 0.36	31.54 ± 3.37
21999	0.2540	$2.44^{+4.01}_{-1.47}$	0.58 ± 0.28	34.19 ± 3.14
22086	0.1357	$1.78^{+2.64}_{-0.97}$	0.35 ± 0.11	9.25 ± 1.30
22094	0.1355	$2.09^{+2.55}_{-1.65}$	1.14 ± 0.15	29.93 ± 3.04
22100	0.1313	$3.67^{+4.81}_{-2.57}$	0.26 ± 0.09	18.04 ± 2.60
52422	0.1964	$2.21^{+2.86}_{-1.61}$	1.75 ± 0.39	62.17 ± 6.28
52427	0.2792	$2.42^{+4.13}_{-1.23}$	1.57 ± 0.56	57.93 ± 7.69
52589	0.1633	$1.98^{+2.67}_{-1.35}$	0.76 ± 0.21	44.75 ± 2.47
55044	0.1938	$0.19^{+0.56}_{-0.01}$	6.15 ± 2.43	11.89 ± 1.65
56910	0.1603	$1.63^{+2.66}_{-0.71}$	1.05 ± 0.35	13.98 ± 1.61
56916	0.1097	$1.74^{+2.30}_{-1.22}$	0.51 ± 0.12	22.72 ± 2.65
56920	0.2829	$2.29^{+2.89}_{-1.71}$	2.00 ± 0.65	41.10 ± 3.05
56969	0.2827	$3.10^{+3.83}_{-2.43}$	1.83 ± 0.67	52.41 ± 3.52
57043	0.1536	$4.92^{+5.62}_{-4.25}$	9.64 ± 1.97	53.08 ± 3.28
57081	0.1398	$0.16^{+0.20}_{-0.12}$	5.49 ± 1.42	19.21 ± 1.94
57083	0.3407	$2.87^{+3.61}_{-2.14}$	2.07 ± 0.66	31.89 ± 2.87
57134	0.3749	$2.81^{+3.61}_{-2.07}$	2.47 ± 0.84	17.52 ± 3.04
57143	0.1109	$1.80^{+2.48}_{-1.07}$	0.60 ± 0.11	10.50 ± 1.79
57163	0.1006	$1.86^{+2.88}_{-1.08}$	0.33 ± 0.08	41.36 ± 3.05
57175	0.4114	$5.82^{+7.94}_{-3.85}$	2.87 ± 1.05	46.87 ± 4.26

Continue in next page...

Table A.3 - Continued

Cluster ID	z	M_{200c}	L_X	λ_{SDSS}
		$10^{14} M_\odot$	$[10^{44}\ h_{70}^{-2}\ erg\ s^{-1}]$	
57260	0.2222	$1.80^{+3.18}_{-0.99}$	0.82 ± 0.33	27.36 ± 2.00
57399	0.1877	$2.56^{+3.22}_{-1.95}$	1.96 ± 0.40	41.92 ± 4.37
57511	0.1049	$0.60^{+1.07}_{-0.11}$	0.23 ± 0.09	16.25 ± 1.69
57524	0.3241	$3.68^{+4.65}_{-2.70}$	2.38 ± 0.71	53.61 ± 4.57
57528	0.3306	$3.64^{+5.52}_{-1.89}$	1.24 ± 0.54	46.64 ± 4.58
57585	0.3144	$3.56^{+4.58}_{-2.50}$	2.69 ± 0.73	55.81 ± 3.24
57600	0.1122	$2.43^{+3.12}_{-1.74}$	0.21 ± 0.09	7.72 ± 0.94
57608	0.1417	$2.40^{+2.90}_{-1.93}$	1.14 ± 0.21	34.25 ± 3.05
57684	0.1709	$3.97^{+5.10}_{-2.80}$	1.09 ± 0.26	39.53 ± 3.10
57686	0.1848	$1.07^{+1.58}_{-0.49}$	6.33 ± 2.42	27.66 ± 2.43
57732	0.4495	$4.52^{+6.06}_{-3.00}$	6.09 ± 1.63	48.24 ± 4.52
57764	0.1417	$1.46^{+2.34}_{-0.69}$	0.67 ± 0.35	38.17 ± 2.51
58008	0.1072	$1.22^{+1.36}_{-0.85}$	0.50 ± 0.11	13.96 ± 2.07
58022	0.1528	$5.94^{+7.72}_{-4.21}$	0.75 ± 0.19	53.99 ± 3.75
58031	0.4044	$5.98^{+8.34}_{-3.90}$	2.30 ± 0.97	52.69 ± 2.95
58083	0.4980	$3.98^{+4.87}_{-3.15}$	5.04 ± 1.97	56.46 ± 2.94
58098	0.1395	$2.30^{+3.36}_{-1.26}$	0.85 ± 0.19	44.17 ± 3.02
58158	0.4121	$2.35^{+3.40}_{-1.19}$	3.51 ± 1.31	43.89 ± 3.25
59364	0.2060	$1.69^{+2.46}_{-0.94}$	0.99 ± 0.44	39.58 ± 5.04
59400	0.0591	$2.23^{+2.68}_{-1.76}$	0.95 ± 0.13	54.13 ± 4.65
59454	0.3762	$3.18^{+4.23}_{-2.11}$	2.22 ± 0.98	28.93 ± 2.65
59461	0.2603	$1.81^{+2.42}_{-1.23}$	0.56 ± 0.37	12.31 ± 1.86
59469	0.1116	$1.33^{+1.88}_{-0.85}$	0.64 ± 0.17	19.68 ± 1.98
59517	0.3843	$3.41^{+5.25}_{-2.07}$	2.27 ± 0.97	12.68 ± 1.79
59526	0.2808	$3.45^{+5.55}_{-2.06}$	2.52 ± 0.70	47.40 ± 3.14
59604	0.2702	$5.81^{+8.06}_{-3.77}$	1.28 ± 0.47	39.35 ± 2.44
59610	0.2860	$5.71^{+8.79}_{-3.52}$	5.84 ± 1.16	53.84 ± 3.42
60061	0.2506	$0.60^{+1.09}_{-0.07}$	0.98 ± 0.39	19.38 ± 1.99
60076	0.1551	$1.94^{+2.68}_{-1.23}$	1.23 ± 0.24	55.35 ± 3.41
60091	0.2889	$2.11^{+2.94}_{-1.33}$	1.46 ± 0.52	24.38 ± 3.19