Galaxy Cluster Science with *Chandra*

Crispin Logan

A thesis submitted to the University of Bristol
in accordance with the requirements of the degree of
Doctor of Philosophy
in the Faculty of Science

Department of Physics,
H.H. Wills Physics Laboratory,
Tyndall Avenue,
Bristol,
BS8 1TL

March 2020

$\sim 44,000$ words

Abstract

Galaxy clusters are the largest gravitationally bound structures in our Universe, and hence can be used to constrain the physical laws that govern cosmic evolution. In order to use clusters as cosmological probes, large samples of bona-fide clusters are needed, in addition to a method with which to estimate cluster masses accurately. Additionally, clusters can be studied to understand the physical processes within the clusters themselves. As clusters form via hierarchical growth, growing ever larger through mergers with smaller systems and accretion of matter from the surrounding environment, understanding the physical processes in the outer regions of a cluster can tell us more about how they grow.

The first project in my PhD tackled the issue of constructing reliable samples of clusters for cosmology, specifically analysing *Chandra* observations of high redshift cluster candidates detected in the XXL survey, to probe not only their AGN contamination, but also to ensure that they were indeed genuine clusters.

The second project in my PhD focussed on the calibration of the X-ray hydrostatic cluster mass estimation method, by comparing the X-ray hydrostatic masses of a sample of 44 clusters with masses for the same sample of clusters measured with an entirely independent method based on information on galaxy dynamics obtained from optical observations.

The final project in my PhD used the same sample of clusters as the mass calibration work to study the outer regions of clusters, specifically calculating the gas fraction at large radii of these clusters, to gain insight into the state of the cluster gas in the cluster outskirts.

Acknowledgements

I would like to thank my PhD supervisors, Ben Maughan and Malcolm Bremer for their help and support throughout my PhD. I am especially grateful for having had such an excellent primary supervisor in Ben, and could not have asked for more from Ben at any point during my PhD.

I would also like to thank my parents for their endless support throughout my life, though words cannot express my gratitude for all they do.

Declaration

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's Regulations and Code of Practice for Research Degree Programmes and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

I note here any contributions to the work from collaborators, along with any work which has previously been published for each of the individual chapters.

Chapter 2 – *Chandra* Constraints on the AGN Contamination of z $>$ 1 XXL Galaxy Clusters – The XXL consortium is a large collaboration of over 100 members. The first paper in the series, which includes the builders of the project is Pierre et al. (2016). This chapter includes results from the large body of work done by members of the XXL consortium, however the precise work in the chapter itself is my own, with the following exceptions: Nicolas Clerc carried out the growth curve analysis mentioned in §2.3; Mark Birkinshaw used his radio maps of the cluster regions to identify if there were any radio counterparts present in the same position as any suspected X-ray point sources in §2.5; I also incorporated comments from the XXL internal reviewer, Adam Mantz, and comments from the anonymous referee from the Astronomy and Astrophysics journal to which the work was submitted.

Note: This chapter is a reproduction of a paper with the same title, published in the Astronomy & Astrophysics Journal. The reference for this paper is Logan C.H.A., Maughan B.J., Bremer M.N., Giles P., Birkinshaw M., Chiappetti L., Clerc N., Faccioli L., Koulouridis E., Pacaud F., Pierre M., Ramos-Ceja M.E., Vignali C., Willis J., 2018, A&A, 620, A18. The manuscript is reproduced here under the non-exclusive right of re-publication granted by ESO to the author(s) of the paper.

Contents

List of Figures

List of Tables

1

Introduction to Clusters of Galaxies

1.1 Introduction

Clusters of galaxies are the largest gravitationally bound structures in the Universe. They have formed, and continue growing, via the hierarchical growth of smaller structures, starting from the initial density perturbations soon after the Big Bang.

Galaxy clusters typically contain 100s of galaxies, which comprise $\sim 5\%$ of their total mass, and are visible in the optical regime. In addition, all clusters have an atmosphere of hot gas (at 10^{7-8}K) that occupies much of a cluster's volume and fills the space between cluster galaxies. It accounts for $\sim 15\%$ of the total cluster mass, and is seen as extended X-ray emission. Dark matter is responsible for the rest ($\sim 80\%$) of the mass in a cluster, and thus measuring cluster mass is a non-trivial matter, as most of the mass is 'invisible'.

Indeed, obtaining accurate cluster masses is an important part of cluster science, as they can be used to constrain the cosmology of our Universe. This is possible, as, being the largest gravitationally bound structures in the Universe, clusters are tracers of large scale structure. Thus, by studying the evolution of clusters over cosmic time, we can constrain the cosmological parameters, whose values determine the rate of growth of structure in our Universe. However, clusters are not only studied for their use as cosmological probes. Due to their size (Mpc scale, masses of $\sim 10^{15}\ M_{\odot}$), as well as their makeup, clusters are unique astrophysical laboratories,

and galaxy evolution, feedback processes in the cluster itself or in galaxies within the cluster, plasma physics and many other processes can be studied.

The first observations of galaxy clusters were carried out in the optical band. It was in the 18th century that William Herschel identified the Virgo cluster of galaxies. However, it was in the 20th century where significant progress in our understanding of clusters began. Abell (1958) and Zwicky et al. (1961) compiled catalogues of large numbers (1000s) of galaxy clusters, that were used as a resource for systematic studies of galaxy clusters and their properties, and are still a valuable resource for cluster science today. Individual galaxy velocities can be measured (from Doppler-shifted spectra), and can subsequently be used, in conjunction with their radius from the cluster centre, to estimate a mass of the cluster using the virial theorem. It was using this method that Zwicky (1937) first measured a cluster mass, finding that if the luminous material (the galaxies) was the only contributor to the cluster's gravitational potential, then the individual galaxy velocities were too high, and would have escaped the cluster's gravitational potential long ago, providing the first evidence for dark matter. In Zwicky's words: "the total mass ... considerably exceeds the sum of the masses of individual galaxies".

Perhaps the biggest advances in our understanding of galaxy clusters, came with the advent of X-ray astronomy. It was in the 1960s and 1970s, with the first X-ray observations of clusters, that a portion of the 'missing mass' was first detected. X-ray emission was observed when looking at known locations of galaxy clusters in the sky (Byram et al., 1966; Bradt et al., 1967; Fritz et al., 1971; Gursky et al., 1971; Meekins et al., 1971), and it was suggested that all clusters may be sources of X-ray emission (Cavaliere et al., 1971). Further observations of other clusters confirmed this. Due to the X-ray spectra of clusters, the leading suggestion for the emission mechanism was thermal bremsstrahlung from hot gas. It was assumed that this gas would have fallen into the cluster's potential from the space surrounding them, however when the X-ray Fe emission line was detected in several clusters (Mitchell et al., 1976; Serlemitsos et al., 1977), it was inferred that a significant amount of this gas had to come from the iron produced during the lifecycle of stars in the cluster galaxies (Bahcall & Sarazin, 1977). This hot cluster gas is commonly referred to as the intra-cluster medium (ICM), and its high temperature is caused by the energy released during the shocks and compression that occur as the gas falls into the cluster's deep potential well during the process of cluster formation. These early observations were the start of a wave of breakthroughs in our understanding of clusters that led to the launch of further X-ray observatories in the later decades of the 20th century. Currently the two most

important X-ray observatories are *XMM-Newton* and *Chandra* (both launched in 1999).

X-ray observations can also be used to measure the mass of clusters. One can estimate the mass of a cluster by investigating the effect it has on the temperature and gas density profiles of the hot gas in the cluster (under the assumption that the ICM is in hydrostatic equilibrium - see §1.7 for more detail), or can use a mass proxy (such as luminosity or temperature) to estimate the mass. With the large collecting area of *XMM-Newton* and the exquisite spatial resolution of *Chandra* working in synergy (e.g. Logan et al., 2018), the purity of cluster samples and the precision of their X-ray mass measurements has reached such a level that constraining cosmological parameters using these mass measurements has become possible.

Perhaps one of the most rapidly developing fields in cluster science today is the discovery of clusters using the Sunyaev-Zeldvich (SZ) effect. The SZ effect occurs when cosmic microwave background (CMB) photons are Compton scattered to higher energies through interactions with the fast moving electrons present in the atmosphere of hot gas that pervades the cluster. One advantage of using the SZ effect to detect and study clusters is that the strength of the SZ cluster signal is independent of redshift (Birkinshaw, 1999), meaning it is an especially useful tool to detect clusters at high redshift.

1.2 Cosmology with Clusters

The study of cosmology is the study of the formation and growth of our Universe. In this section I give a brief overview of our current understanding of cosmology, and how the study of clusters can be used to probe the cosmology of our Universe.

1.2.1 Hubble equation and expanding Universe

One of the first pieces of evidence that we were in an expanding Universe was provided by Hubble (Hubble, 1929). By measuring the velocity with which galaxies were receding from us, and plotting this against their distance, he realised that there was a clear linear relationship between these two quantities, where the recession velocity, $\overrightarrow{v}$, is proportional to the distance, $\overrightarrow{r}$, to the object (Hubble's law):

$$\overrightarrow{v} = H_0 \overrightarrow{r} \tag{1.1}$$

where H_0 is the Hubble constant (where the subscript '0' is necessary to define the value of the Hubble parameter, H, at the current epoch as H changes with time).

The fact that everything in the Universe is moving away from everything else (ignoring peculiar motions), was the first piece of evidence that suggested that the Universe began with the Big Bang, undergoing a period of rapid expansion from a high-density and high-temperature state, followed by continuing, but less rapid, expansion today.

Due to this expansion of our Universe, we often use a co-moving coordinate system when modelling how our Universe is evolving:

$$\overrightarrow{r} = a(t)\overrightarrow{x} \tag{1.2}$$

where $a(t)$ is the scale factor, which varies with time and has a value such that the co-moving coordinate system tracks the expansion rate of the Universe, and that objects with no peculiar velocity will be at the same coordinate values a set time later.

Using the co-moving coordinate system, where the expansion of the Universe is given by $a(t)$, we can write Hubble's law as

$$v(t) = H(t)d(t) \tag{1.3}$$

where $H(t)$ = $\dot{a}/a$ and $d(t)$ is the distance.

1.2.2 Friedmann equation

One of the most important equations in cosmology is the Friedmann equation. We can derive the Friedmann equation under the assumption of the cosmological principle, which states that the Universe looks the same wherever you are (i.e. it is homogeneous and isotropic). Homogeneity states that the Universe looks the same at each point, and isotropy states that the Universe looks the same in every direction.

For this derivation / motivation of the Friedmann equation, a sketch of our setup is shown in Figure 1.1. If we consider a test particle (which can represent any particle in our Universe due to our assumption of the cosmological principle), and say that it is at radius r from the 'centre' of the Universe (again, the 'centre' of the Universe can be any point, as the Universe looks the same from anywhere), then under the rule of energy conservation:

$$U = T + V \tag{1.4}$$

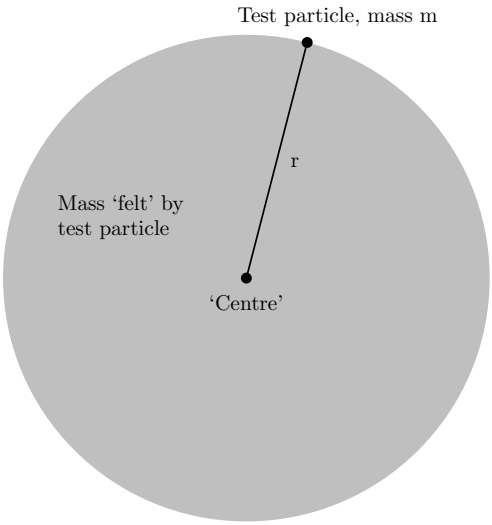

Figure 1.1: A sketch showing the setup that we use to derive the Friedmann equation. The test particle, with mass m, is a distance r from the 'centre' of the Universe, and the gravitational potential, V (see equation 1.4), of the test particle is proportional to the mass within the sphere bounded by radius r.

where U is a constant (that depends on the specific test particle that we choose), T is the kinetic energy of the particle, and V is the gravitational potential energy of the test particle. We can expand this equation to obtain how the separation between the test particle and 'centre' of the Universe changes:

$$U = \frac{1}{2}m\dot{r}^2 - \frac{4\pi}{3}G\rho r^2 m \tag{1.5}$$

where m is the mass of the test particle, and ρ is the average density within the sphere bounded by radius r, and $\dot{r}$ is the velocity of the test particle. G is the gravitational constant.

The above can apply to any test particle and 'centre' in our Universe, because of the assumption of homogeneity, and we can use the co-moving coordinate system (see equation 1.2) to rewrite equation 1.5 as:

$$U = \frac{1}{2}m\dot{a}^2x^2 - \frac{4\pi}{3}G\rho a^2x^2 m \tag{1.6}$$

which via rearrangement and multiplying both sides by 2 / ma^2x^2 can be written as:

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

where $kc^2 = -2U/mx^2$, and the value of k dictates whether the Universe is closed, flat or open ($k > 0$, k = 0 and $k < 0$ respectively), which correspond to the curvature of the Universe (spherical, flat or hyperbolic respectively). This is the basic form of the Friedmann equation, and can also be written as

$$H(t)^2 = \frac{8\pi G}{3}\rho - \frac{kc^2}{a^2} \tag{1.8}$$

using the definition $H(t) = \dot{a}/a$ from §1.2.1.

For a flat Universe (k = 0), we have a simpler form of the Friedmann equation

$$H(t)^2 = \frac{8\pi G\rho}{3} \tag{1.9}$$

which has a given density for a given value of $H(t)$, also known as the critical density, ρ_c where

$$\rho_c = \frac{3H^2}{8\pi G} \tag{1.10}$$

The critical density is the density at which the Universe is flat (for a given $H(t)$).

In cosmology, a density parameter Ω is used where

$$\Omega(t) = \frac{\rho}{\rho_c} \tag{1.11}$$

and, clearly, for the case when Ω = 1, we have a flat Universe (k=0). Ω can be broken down into different components, commonly as $\Omega = \Omega_m + \Omega_\Lambda$ where Ω_m is the matter density of our Universe and Ω_Λ is the effective mass density of the dark energy of our Universe. This Ω_Λ term appears due to the cosmological constant that Einstein added to the Friedmann equation to counteract the effects of gravity to achieve a static Universe (which is what was believed to be the case at the time, before Hubble's discovery that the Universe was expanding, after which the cosmological constant was assumed to be zero). The cosmological constant Λ appears in the Friedmann equation as

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{kc^2}{a^2} + \frac{\Lambda}{3} \tag{1.12}$$

It is currently believed to be non-zero, after the discovery in the late 20^{th} century that the expansion of our Universe is in fact accelerating.

1.2.3 Distances measures in cosmology

There are a number of distance measures in cosmology that can be used to measure the distance between two points. Though not strictly speaking a distance measure, redshift can be used to approximate distances to distant objects. The redshift of an object, z, is defined as the fractional Doppler shift of light emitted from the object whose motion relative to the observer is purely due to the expansion of the Universe:

$$z = \frac{\nu_{\rm e}}{\nu_{\rm o}} - 1 = \frac{\lambda_{\rm o}}{\lambda_{\rm e}} - 1 \tag{1.13}$$

where $\nu_{\rm o}$ and $\nu_{\rm e}$ refer to the observed and emitted frequency of light, respectively, and $\lambda_{\rm o}$ and $\lambda_{\rm e}$ refer to the observed and emitted wavelengths, respectively.

Redshift can also be related to the scale factor $a(t)$ of the Universe, as

$$1 + z = \frac{a(t_{\rm o})}{a(t_{\rm e})} \tag{1.14}$$

The co-moving distance, $D_{\rm M}$, is the distance between two points which remains constant with time, as long as the two points are moving with the Hubble flow. For instance, if we define the proper distance as the distance between two points that would be measured at the time they are observed, and then divide this distance by the ratio of the scale factor of then to now, we obtain the co-moving distance (alternatively it could be the proper distance multiplied by 1 + z). The idea of a co-moving distance follows naturally from the co-moving coordinate system.

The co-moving distance is defined as

$$D_{\rm M} = \frac{c}{H_0} \int \frac{dz}{E(z)} \tag{1.15}$$

where

$$E(z) = \sqrt{(1+z)^2(1+\Omega_{\rm m} z) - z(2+z)\Omega_\Lambda} \tag{1.16}$$

and $E(z)$ is used to describe the dependence on redshift of the Hubble parameter as

$$H(z) = H_0 E(z) \tag{1.17}$$

Related to the co-moving distance are two further distance measures, the angular diameter distance, $D_{\rm A}$, and the luminosity distance, $D_{\rm L}$. The angular diameter distance is defined as the ratio of an object's physical size, l, to its angular size, θ (in radians):

$$D_{\rm A} = \frac{l}{\theta} \tag{1.18}$$

and is related to the co-moving distance as

$$D_{\rm A} = \frac{D_{\rm M}}{(1+z)} \tag{1.19}$$

An interesting point to note is that the angular diameter distance does not increase indefinitely with increasing redshift, but in fact, at $z \sim 1$, it starts to turn over, meaning that objects of the same physical size will have larger apparent angular sizes at higher redshift, after this turnover point. This is due to the finite speed of light and the fact that the Universe is expanding.

The luminosity distance defines how the flux of an object is related to its bolometric luminosity, with the relationship as follows:

$$D_{\rm L} = \frac{D_{\rm M}}{(1+z)^2} \tag{1.20}$$

The $1/(1+z)^2$ dependence of the luminosity distance upon the co-moving distance is due to two separate $1/(1+z)$ dependences, both due to the expansion of the Universe. One is due to the fact that each photon's wavelength will be redshifted upon arrival and hence the photon itself will be of lower energy when observed, and the other is due to the fact that the arrival of photons will be less frequent.

1.2.4 Cluster formation, evolution, and use as cosmological probes

Clusters are believed to be formed from the initial density perturbations in the moments after the Big Bang (Kravtsov & Borgani, 2012). These grew into more massive structures, accreting more matter with time. Their growth came via mergers of smaller systems, and they are seen today as galaxy clusters. The values of the cosmological parameters in the cosmological model used to describe our Universe affect its evolution (see Figure 1.2), and as clusters are the largest gravitationally bound objects in our Universe, clusters are a perfect tracer of this evolution of structure (for example, Borgani & Guzzo, 2001; also see Voit, 2005 for a review).

A cluster mass function describes the number density of clusters as a function of mass, and can be constructed by measuring the masses of a representative sample of clusters. By comparing the observed cluster mass function to the predicted mass function (given certain cosmological parameters) in different redshift bins, the evolution of structure in the Universe can be probed,

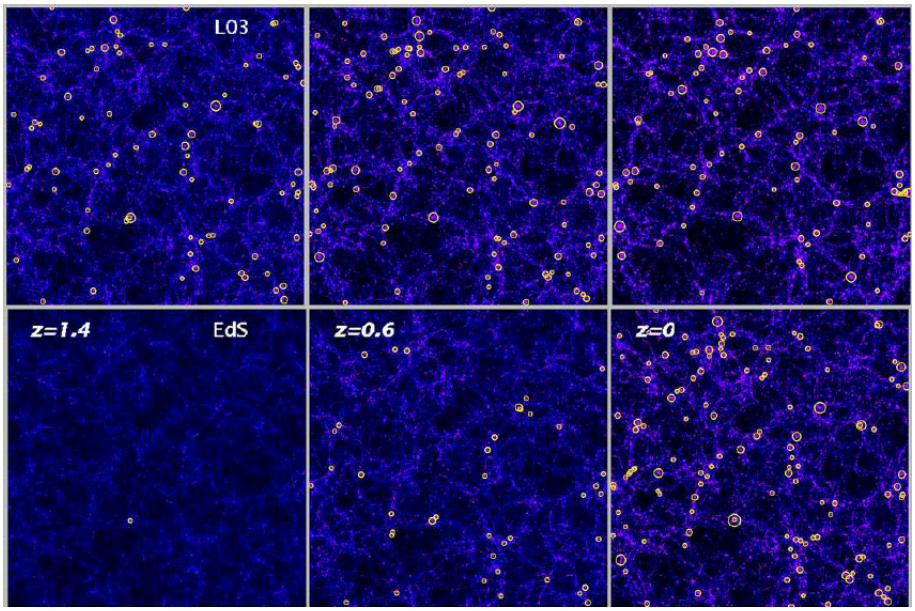

Figure 1.2: The evolution (with decreasing redshift, or increasing cosmic time) of large scale structure from N-body simulations with two different cosmological models used is shown. The structure seen is the dark matter distribution. The upper panels show $\Omega_{\rm m} = 0.3$ and $\Omega_\Lambda = 0.7$, and the bottom panels show $\Omega_{\rm m} = 1$ (and k = 0 for both). Yellow circles show clusters with $T > 3$ keV, with bigger circles showing higher temperatures. Taken from Borgani & Guzzo (2001).

and the cosmological parameters of our Universe constrained (e.g. Vikhlinin et al., 2009). In order to do this accurately, the cluster sample itself needs to be well understood, to ensure that it is as close to a representative sample as possible; if it is not, this needs to be accounted for. In addition, our mass estimation techniques need to be well understood, including any biases that there might be in the technique used.

Clusters can also be used as cosmological probes using an entirely different method. As clusters are so large, they can be used as 'standard buckets', meaning that the make-up of a cluster is assumed to be representative of the Universe as a whole. Thus, by measuring the baryon fraction, cosmological parameters can be constrained (e.g. Ettori et al., 2009). The baryon fraction is the baryon mass, consisting predominantly of the gas mass and the stellar mass, divided by total cluster mass, and will be equivalent to Ω_b / Ω_m where Ω_b and Ω_m are cosmological parameters. By using a prior constraint on the Hubble constant, and using the value of Ω_b from, for example, big bang nucleosynthesis calculations, the value of Ω_m can be constrained.

In order to estimate masses using X-ray data for large samples of clusters to carry out cosmological studies, observational mass proxies are often used. The reason for this is that in order to do a full hydrostatic mass analysis (see Chapter 3), spatially resolved temperature measurements are needed, which require long observation times for typical clusters. Thus, mass proxies are used, which require shorter observation times to be measured. Examples of these mass proxies are the temperature or X-ray luminosity of the ICM. The use of these mass proxies in cluster cosmology means that the study of X-ray scaling relations are highly important, especially those relating cluster mass with another easier to measure quantity, such as temperature or luminosity. X-ray scaling relations are presented in more detail in §1.5, and can also be used on their own to study processes in clusters (see §1.5.2).

1.3 Galaxy Clusters at Different Wavelengths

Galaxy clusters look very different depending on the wavelength band in which they are observed, and in each wavelength band we may learn different things about a cluster. Indeed, in two different wavelength bands we may in fact be looking at completely different components of a cluster. This can be seen in Figure 1.3, where we show a combined X-ray and optical image of a cluster.

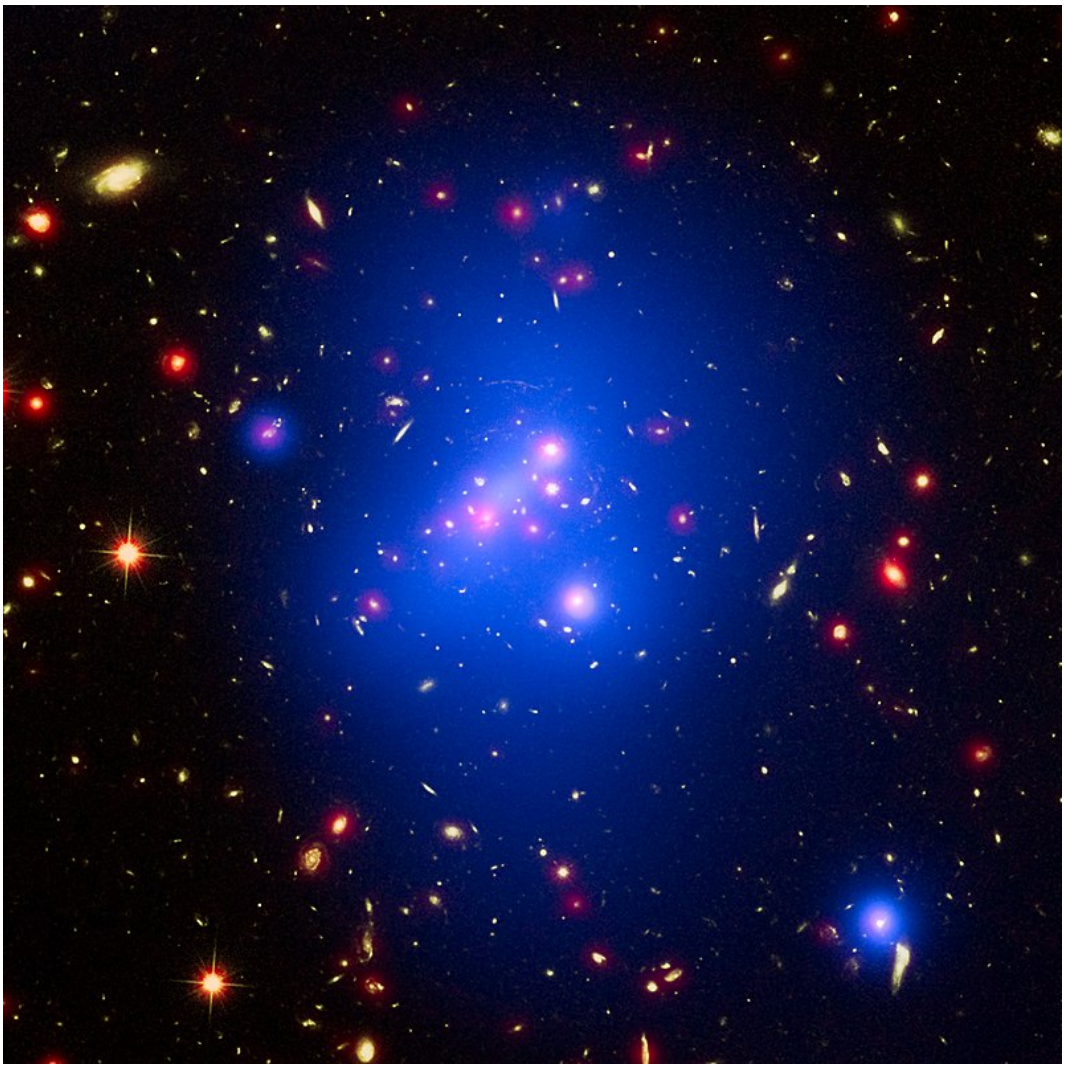

Figure 1.3: An image of a galaxy cluster with X-ray emission shown in blue, overlaid on an optical and infrared image (shown in red and green respectively). The X-ray emission originates from the hot ICM and is clearly extended. The optical and infrared image shows an over-density of galaxies in the centre of the cluster. The X-ray data is taken with *Chandra*, the optical data is from *Hubble*, and the infrared data is from *Spitzer*.

1.3.1 X-ray

When observing a cluster in the X-ray, we see extended X-ray emission. This is from thermal emission of the ICM, and is due to two emission mechanisms, detailed below.

a) Continuum Emission

The ICM is sufficiently hot (10^{7-8}K), such that it is almost completely ionized. The primary emission mechanism in clusters, thermal bremsstrahlung emission (free-free), arises when a fast moving electron interacts with a positive ion. This interaction results in the electron being deflected, causing it to radiate electromagnetic energy, predominantly in the form of X-rays (due to the high temperatures of the gas involved), in order to maintain the conservation of energy. The hotter the gas, the higher the average energy of the electrons in the gas, and the faster moving they will be, such that when an electron is deflected by a positive ion, the higher the energy of the X-ray photon that is produced. Because the energies (and therefore speeds) of the electrons in a hot gas have a specific distribution, the sum of the radiation from each of these deflections can be predicted, and is solely dependent on the temperature of the gas. The shape of this total radiation emitted is called a thermal bremsstrahlung continuum (see Figure 1.4). The intensity, I, of the emission at an energy E, for a gas with temperature T, is:

$$I(E,T) = AG(E,T)Z^2 n_{\rm e} n_{\rm i} (kT)^{1/2} e^{-E/kT} \tag{1.21}$$

where k is Boltzmann's constant, A is a proportionality constant, and $G(E,T)$ is the Gaunt factor, which corrects for quantum mechanical effects and the effects of distant collisions, and is a slowly decreasing function of E. Z is the charge of the positive ions, and $n_{\rm e}$ and $n_{\rm i}$ are the electron and ion densities respectively. It is clear that the intensity is proportional to the density squared of the gas, assuming that $n_{\rm e}$ and $n_{\rm i}$ are similar.

Other continuum emission mechanisms are radiative recombination (free-bound), which is when an electron is captured by a positive ion, and a photon is emitted. This free-bound emission only becomes dominant over bremsstrahlung emission at low temperatures (kT $<<$ 0.1keV Kaastra et al., 2008), and such low temperatures are not seen in typical clusters. Another continuum emission mechanism is two-photon emission, where a photon pair is emitted following a single electron transition between two specific energy states (2s-1s), to circumvent quantum mechanical selection rules that would otherwise forbid this radiative transition. However, two-photon emission is not strong enough to be seen in cluster spectra (see Figure 1.4).

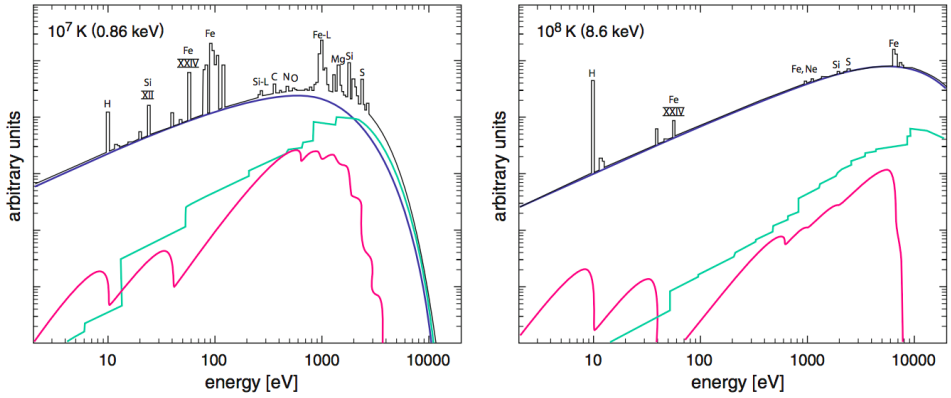

Figure 1.4: From Boehringer & Werner (2009). This figure shows the X-ray spectrum (with intensity plotted on the y-axis) for a gas at 10^7K (left panel) and 10^8K (right panel) with solar abundance. Bremsstrahlung emission (blue), recombination radiation (green), and two-photon radiation (red) are shown. The line emission is also shown, and the elements from which the emission arises are labelled.

b) Line Emission

Another important source of radiation from a hot gas is from line emission (bound-bound). In the ICM, the elements heavier than hydrogen are not completely ionized, and line emission provides a significant contribution to overall cluster emission, at least for temperatures $\lesssim 5 \times 10^7$K (or 4.3 keV). A fast moving electron colliding with an ion of one of these heavier elements will often transfer some of its energy to one of the bound electrons in that ion, causing it to move to a higher energy level. This state is short lived and when the excited electron moves back to its original energy level, it will emit a photon of a characteristic energy, equal to the energy difference between the two different energy levels of this transition. This leads to 'extra' emission above the bremsstrahlung continuum at this characteristic energy for the heavier element in question. Line emission is increasingly important at lower temperatures in the range of temperatures typically seen in clusters (i.e. the lower end of 1 - 10 keV). This is evident in Figure 1.4, where the contribution from line emission to the emission spectrum is higher at lower temperature.

c) X-ray Spectra and Spectral Fitting

By combining the contributions from the different X-ray emission mechanisms from the ICM, we can obtain a predicted emission spectrum for a cluster. There are various theoretical models

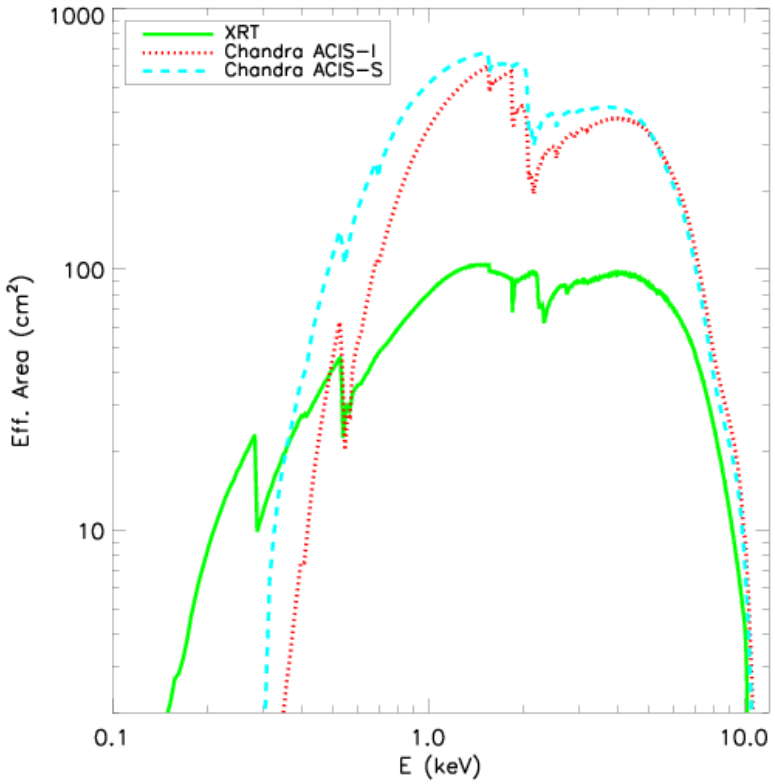

Figure 1.5: The effective area versus incident photon energy is shown for the *Chandra* ACIS-I and ACIS-S chips (see §1.4), and for the X-ray Telescope (XRT) of another X-ray observatory, *SWIFT*. Taken from Tundo et al. (2012).

that combine these different emission mechanisms into one total emission spectrum; amongst them is the Astrophysical Plasma Emission Code (APEC, Smith et al., 2001). The purpose of these models is to then fit them to an observed X-ray cluster spectrum to determine the properties of the cluster in question. This is done by varying the model parameters, such as the temperature or metallicity, until the model fits the observed spectrum well. However, before this fitting process occurs, the model spectrum needs to be convolved with the instrument response (such that it appears as though it has been observed through an X-ray telescope).

The two instrumental effects that are taken into account are the telescope effective area (see Figure 1.5), and the CCD (charged coupled device) response, which are both a function of the energy of the photons and the position on the CCD where the photon is detected. The information on these two instrumental effects are stored in the Ancillary Response File (ARF) and the

Redistribution Matrix File (RMF). The ARF tells us about the quantum efficiency (fraction of incident photons detected) of the CCD and effective area of the telescope as a function of photon energy. Once the model spectrum is multiplied by the ARF, it is effectively showing what an observed spectrum would look like were it observed by a detector with perfect energy resolution. We then convolve it with the RMF, which describes the probability that a photon will be detected in a given energy channel. These "response files", with which we convolve the model spectrum, are specific to the telescope used to obtain the observed data to which we want to fit our model data.

Once we have convolved the model spectrum with the ARF and RMF, we can fit it to the observed cluster spectrum. One fitting statistic used to judge whether the predicted model is a good fit to the data is called the χ^2 statistic. It is defined by:

$$\chi^2 = \sum \frac{(c_{\rm i} - m_{\rm i})^2}{\sigma_i^2} \tag{1.22}$$

where $c_{\rm i}$ is the observed number of counts in bin i, and σ is its error. $m_{\rm i}$ is the predicted model counts. To use the χ^2 statistic, each bin in which the χ^2 statistic is being used should contain a sufficiently large number of photons (> 20), meaning that the uncertainty on the number of photons in each bin can be considered to be Gaussian.

Another method for judging the goodness of fit is called the *C*-statistic[1], and is based on the cash-statistic (Cash, 1979). It has a major advantage over the χ^2 statistic in that we can use fewer photons in each bin, as it assumes a Poissonian distribution of the photons. This means that when using the *C*-statistic to fit a model to data in different radial bins (in order to obtain a radial profile of whatever is being fitted for), one can use finer radial bins, leading to increased resolution in the profile. The *C*-statistic is defined as:

$$C = 2\sum_{i=1}^{N}(tm_{\rm i}) - S_i \ln(tm_{\rm i}) + \ln(tm_{\rm i}) \tag{1.23}$$

where N is the total number of bins, S_i are the observed counts and $m_{\rm i}$ are the predicted count rates (based on the current model and instrument response) in bin i, and t is the exposure time. In this work we use the *C*-statistic (see §4.3.3 and 4.3.8) for all spectral fitting.

[1]found under 'cstat' on `https://heasarc.gsfc.nasa.gov/xanadu/xspec/manual/XSappendixStatistics.html`

1.3.2 Optical

When observing a cluster of galaxies in the optical waveband, we see an over-density of galaxies in the cluster region. However, due to projection effects, simply seeing an over-density of galaxies does not necessarily mean there is a cluster present. Weak lensing (WL) is a powerful technique that can be used to detect the presence of a galaxy cluster, and probe its properties. WL uses the fact that, as predicted by Einstein's theory of general relativity, a massive object will distort the space around it such that a light ray passing through the gravitational potential will be deflected. For a background galaxy, this will mean its shape is distorted, and through measuring the overall distortions of a number of background galaxies, the presence of a cluster and its mass distribution can be estimated. Due to the small level of distortion in individual galaxies, the average distortion over a large number of galaxies needs to be measured for this method to be effective.

Strong gravitational lensing, which can occur due to the more concentrated gravitational potential in the cores of clusters, leads to more strongly lensed images of background galaxies, and can be used to infer the mass of the inner regions of clusters.

Galaxy dynamics can also tell us about the properties of clusters, and if the line of sight velocities for a number of cluster galaxies can be measured, the velocity dispersion method can also be used to measure cluster masses.

1.3.3 Microwave and Radio

Galaxy clusters can be observed in the microwave and radio wavelengths by using the SZ effect (Birkinshaw, 1999). The SZ effect occurs when CMB photons are upscattered to higher energy when they pass through the ICM of a cluster, leading to the overall CMB spectrum to be shifted to a higher frequency. Thus, when observing large areas of the sky, clusters will leave 'imprints' on the CMB map, due to this upscattering of the CMB photons by the cluster gas. These 'imprints' are detected using telescopes with observing frequencies that correspond to the decrement, null and increment region of the SZ effect blackbody spectrum of the CMB. The null is at the frequency where the SZ effect CMB blackbody spectrum and unaffected CMB blackbody spectrum have the same intensity, and the decrement and increment are below and above this frequency respectively. The strength of the SZ cluster signal is solely dependent on

the properties of the ICM (and not the distance to the cluster), making it a powerful tool in detecting clusters at high redshift, which is the regime where methods using X-ray and optical data perform poorly.

Observing clusters in the radio regime can also reveal radio halos and relics. Radio halos are diffuse sources found in the central regions of merging clusters and might be a result of the (re)-acceleration of electrons due to the merger. Radio relics (or shocks) are diffuse extended sources that trace particles undergoing (re)-acceleration at shock wave boundaries (van Weeren et al., 2019).

1.4 X-ray Observatories

As most cosmic X-rays cannot penetrate even the low density outer layers of the Earth's atmosphere, observing X-rays from ground based instruments is impossible. In order to observe any X-rays at all, it is necessary to be above 99 per cent of the Earth's atmosphere, and to observe cosmic X-rays in the energy band where they are most prominent, it is necessary to be above 99.9999 per cent of the Earth's atmosphere.

In order to make a useful observation of an X-ray source, an X-ray observatory must first capture X-rays, and then detect them and measure their energies. Modern X-ray telescopes use mirrors to collect and focus X-rays. This is a difficult task for two reasons, both due to the high energies of X-rays. Firstly, the mirror itself must be very smooth; due to their high energies, X-rays have wavelengths of nanometres, and a 1nm bump on a mirror will negatively affect their reflection efficiency. Secondly, X-rays will only be reflected from a smooth surface if the angle of incidence is small ($<1^\circ$). In fact, X-rays will be reflected with a high efficiency until the angle of incidence hits a critical angle θ_c, where the efficiency drops considerably. The critical angle is related to the energy of the incident X-ray, E, and the electron density in the material that it is incident upon, ρ, as follows:

$$\theta_c \propto \sqrt{\rho}/E \tag{1.24}$$

It is this sudden drop-off in reflection efficiency combined with an inverse dependence of the critical angle on energy that leads X-ray telescopes to have an abrupt high energy cut-off, above which they cannot detect X-rays efficiently. Clearly, a denser material (higher ρ), will have a higher critical angle, and thus be able to detect X-rays to a higher energy.

1.4.1 X-ray Detectors

After being focussed, a detector measures an X-ray's properties. The early X-ray missions used proportional counters to detect X-rays, and most modern missions use charge coupled devices (CCDs).

Proportional counters contain a gas in a sealed chamber, which, when an X-ray passes through the gas, causes ionisation of the gas atoms, and the resultant electrons are attracted to a positively charged wire within the chamber. The signal measured at the wire due to these electrons is proportional to the X-ray photon's energy.

X-ray CCDs are made up of an array of semiconductor (often silicon) pixels. When an X-ray hits a pixel, it will liberate a number of electrons that is dependent on the energy of the X-ray photon and the energy gap of the semiconductor. With a certain frequency (up to 100s of times per second), the number of electrons in each pixel are "read out"; electrode gates are used to to enable the electrons in each pixel to be transferred along a row of pixels for this read-out process.

The cutting-edge of X-ray detectors in terms of spectral capability are calorimeters. Calorimeters consist of a crystal kept at a very low temperature; when an X-ray is absorbed by the crystal, its temperature will rise, and this temperature rise is proportional to the X-ray energy. The energy resolution that can be obtained with calorimeters are an order of magnitude better than the highest energy resolution of CCDs. A calorimeter was flown on the *Suzaku* mission, but unfortunately the cooling mechanism failed before data were obtained. More recently, in 2016, the *Hitomi* satellite was launched, and carried a calorimeter. Unfortunately this satellite span out of control soon after being placed in orbit. However, before this occurred, an observation of the Perseus cluster was made using the calorimeter, and the data sent back to Earth. The data received were exquisite (see Figure 1.6), and promise a future for calorimeters in future X-ray missions.

1.4.2 Early X-ray Observatories

In the 1950s and 1960s, X-ray detectors were flown on balloons and rockets to observe the X-ray sky. These early missions provided detections of some of the brightest X-ray objects in the sky, and opened a new window to the cosmos. The first X-ray satellite, *Uhuru*, was launched in 1970. Proportional counters were used in this mission, and lightcurves of the strongest X-ray sources were measured. The launch of *Einstein* in 1979, the first satellite that focussed

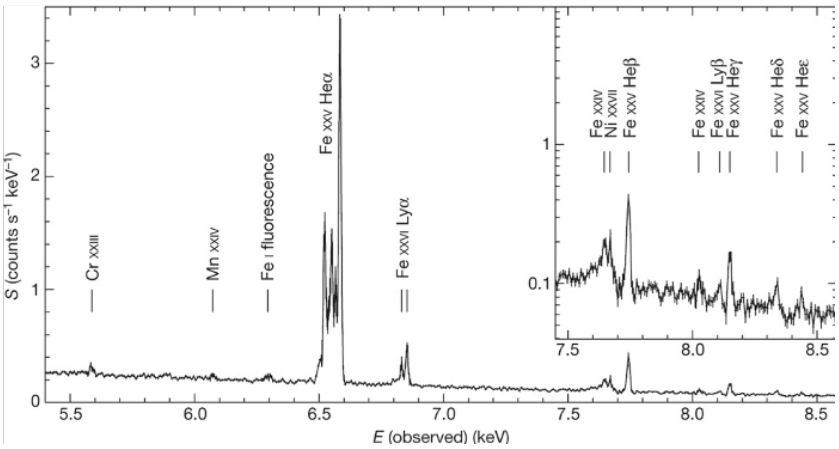

Figure 1.6: The spectrum of the Perseus cluster as taken with *Hitomi*. Taken from Hitomi Collaboration et al. (2016)

X-rays using a mirror system, enabled a huge jump in the sensitivity of X-ray observations, and also provided superior spatial resolution. In 1990, the ROSAT satellite was launched, as a collaboration between Germany, the USA and the UK, and undertook an all-sky survey that is still used today. In 1999 *Chandra* and *XMM-Newton* were launched, and are still in use today, with each year of observations bringing further scientific advances.

1.4.3 *Chandra*

Chandra operates in the 0.1-10 keV energy band, with a sub-arcsecond spatial resolution. This exquisite detail with which *Chandra* can study objects is an order of magnitude better than previous X-ray observatories. To achieve this resolution, *Chandra*'s mirrors are the smoothest and cleanest ever produced. The mirror design (see Figure 1.7) is called a Wolter type 1 design, named after Hans Wolter, who developed it in the 1950s. It uses sections of a paraboloid and a hyperboloid in its setup. *Chandra* has four nested mirrors (see Figure 1.8) in this setup to increase the effective area of the telescope.

Due to the finite precision with which the mirrors can be aligned in their assembly, and the presence of any dust on the surface of the mirrors, an incident X-ray photon will not fall at exactly the focal point of the mirror setup. Instead, an X-ray telescope will have a point spread function (PSF), which tells us the chance that an incident X-ray will land at a certain point on the detector (which is placed at the focal point of the mirror setup). *Chandra*'s PSF is complex, and is a function of X-ray energy and the position of the source on the X-ray detector, increasing

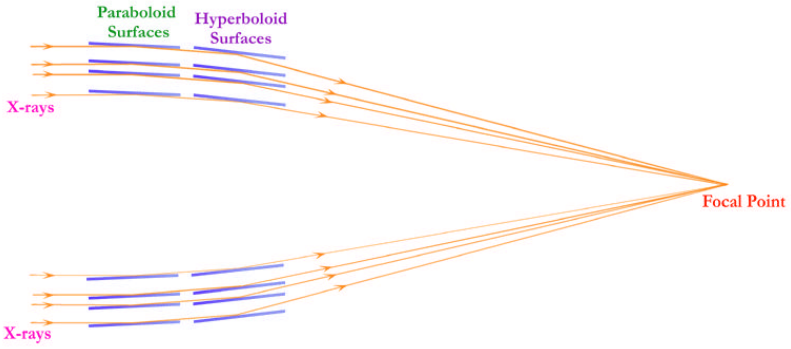

Figure 1.7: A schematic showing the Wolter 1 mirror setup. Taken from `http://chandra.harvard.edu/resources/illustrations/teleSchem.html`.

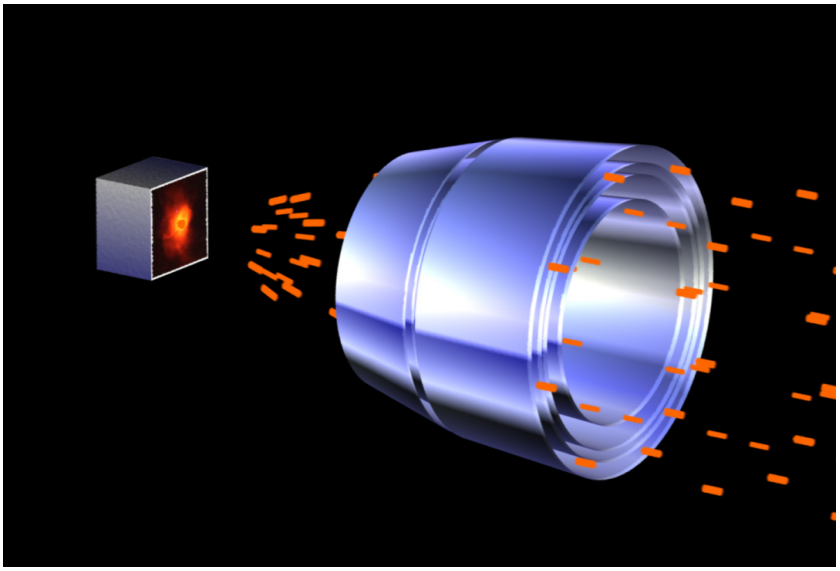

Figure 1.8: An image showing *Chandra*'s four nested mirrors focussing X-rays onto the focal plane, where the ACIS flight focal plane is found (see Figure 1.9). Taken from `http://chandra.harvard.edu/resources/illustrations/teleSchem.html`.

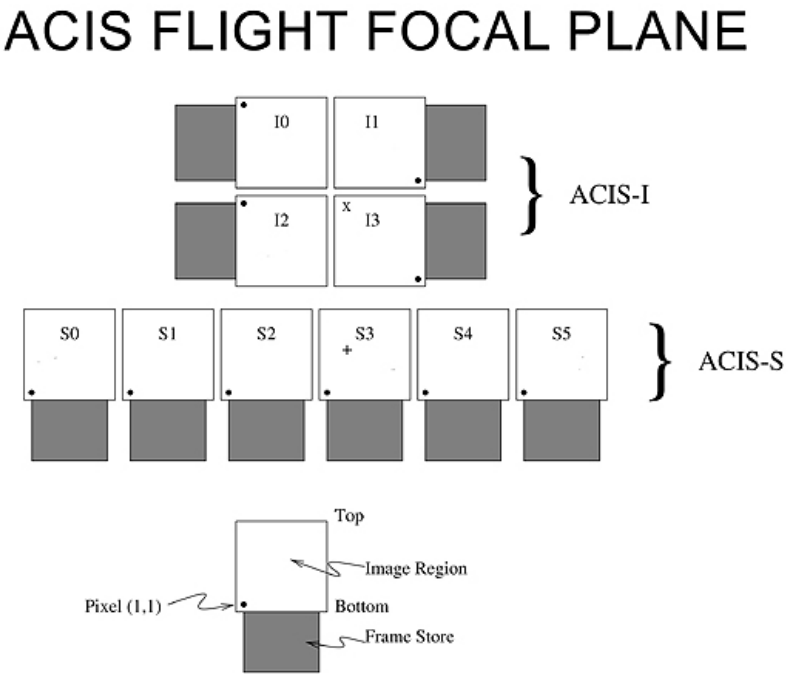

Figure 1.9: A schematic of the ACIS flight focal plane showing the CCD chip layout. The chips are labelled according to whether they are ACIS-I or ACIS-S chips. The '+' and 'x' represent the aimpoints for the S array and I array chip setup. The chips S1 and S3 are back-illuminated, and the other chips are front-illuminated. Taken from `http://chandra.harvard.edu/resources/illustrations/instrumentsSchema.html`.

in size with distance from the centre and decreasing with energy.

Chandra's detectors, called the ACIS (Advanced CCD Imaging Spectrometer) camera, are placed at the focal point of the telescope (see Figure 1.7). Their layout is shown in Figure 1.9. The ACIS camera consists of 10 front illuminated (FI) chips, and two back-illuminated (BI) chips (S3 and S1). Back illuminated chips have the readout electronics on the side of the chip that is not exposed, meaning that they have a higher background, but are more responsive at lower energies. Each chip is 1024 by 1024 pixels with a pixel size of 0.492 arcseconds, and has an 8.3 by 8.3 arcminute field of view.

Observations can be undertaken with up to six of the chips turned on at once. Using the four ACIS-I chips (and the S3 chip) is a popular chip combination for cluster studies, as the ACIS-I chips provide a larger field of view (of 16.9 by 16.9 arcminutes) and lower background than a chip combination using predominantly ACIS-S chips.

The *Chandra* observatory also carries other instruments (the High Resolution Camera and the High Resolution Spectrometers), but for the work presented in this thesis the *Chandra* data that are used were taken using the ACIS camera.

1.4.4 *XMM-Newton*

XMM-Newton operates in the 0.2-12.0 keV energy range, and uses the Wolter 1 mirror design as is used in *Chandra*. *XMM-Newton* has three telescopes, each with 58 nested mirrors in contrast to *Chandra*'s four mirrors, meaning it has a much higher effective area and thus sensitivity. In addition, *XMM-Newton*'s field of view is 30 arcminutes, and has a PSF of $\sim 6''$ (FWHM).

Each of *XMM-Newton*'s three X-ray telescopes has a CCD instrument mounted at the focus. These three instruments are termed the European Photon Imaging Camera (EPIC[2]), and they are all imaging spectrometers: one is the pn camera (called as such as it uses pn CCDs), and the other two use Metal-Oxide Semiconductor (MOS) CCD arrays. Each MOS camera has seven front-illuminated CCDs, with one CCD located at the focal point of the telescope, and the other six around it, at an angle so as to roughly follow the curvature of the focal plane, in order to improve the focus for off-axis sources. Each CCD consists of 600 $\times$ 600 pixels, and each pixel is $1.1''$ wide. The pn camera consists of 12 back-illuminated CCDs of 200 $\times$ 64 pixels in a 6 $\times$ 2 array, with a pixel width of $4.1''$. The pn camera has a faster readout time and higher effective

[2] https://xmm-tools.cosmos.esa.int/external/xmm_user_support/documentation/uhb/epic.html

area than the MOS cameras, and the MOS cameras have a higher spectral resolution. *XMM-Newton* also has two Reflection Grating Spectrometers (RGS) instruments that are mounted in the two telescopes that have MOS cameras (and use $\sim$ 1/2 of the incoming X-ray photons in each telescope), and an optical/UV monitor.

Clearly *XMM-Newton* and *Chandra* have very different strengths, with one providing exquisite spatial resolution, and the other providing extremely high sensitivity. These two observatories complement each other perfectly, and work best in synergy with each other. One example of this is presented in Chapter 2, where we use *Chandra* to follow up high redshift cluster candidates that are detected with *XMM-Newton* in the XXL cluster survey. *XMM-Newton* is best used as a survey telescope due to its high sensitivity, yet cannot provide the resolution necessary to resolve point sources in distant clusters, hence the use of *Chandra* to vet the selected distant clusters identified from the XXL survey.

1.4.5 Future X-ray Observatories

In mid-2019 *eROSITA*, a joint German and Russian mission, was launched and will map the entire X-ray sky in the 0.3 - 10 keV band over a 7-year period. *eROSITA* has a PSF similar in size to that of *XMM-Newton*'s, and is expected to be $\sim$ 25 times more sensitive than *ROSAT* in the 0.5 - 2 keV band, and should detect $\sim$ 100,000 clusters. Current expectations are that the data will be made public 2 years after it is obtained.

XRISM, which has an expected launch date of mid-2022, is a collaboration between JAXA and NASA, and is intended to be a successor to the short-lived *Hitomi* mission, carrying a calorimeter spectrometer with which it will be able to obtain high resolution spectra.

Another major X-ray observatory, ATHENA, is currently due for launch in 2031, and will be 100 times more sensitive than *XMM-Newton*. ATHENA will be able to detect even the faintest of clusters due to its high sensitivity, and enable it to carry out one of the main aims of the mission, which is to investigate how galaxy groups and clusters evolve.

1.5 Self-similarity and scaling relations

Galaxy clusters are often described as self-similar. Self-similarity refers to the fact that clusters are effectively scaled up or scaled down versions of each other. There are two types of self-similarity. Strong self-similarity means that clusters of differing masses are identical scaled

versions of each other. Weak self-similarity means that clusters of differing redshifts are identical scaled versions of each other, when the changing density of the Universe with redshift is accounted for. Self-similarity has two main assumptions: that clusters form via a single gravitational collapse, and that the only energy input into the cluster comes from the gravitational collapse of the cluster. Neither of these assumptions are completely true, which leads to departures from the theoretical scaling relations when calculating scaling relations from observations of clusters.

1.5.1 Theoretical scaling relations

Scaling relations are power law relations between cluster properties, and the theoretical scaling predictions rely on self-similarity holding true.

Theoretical scaling relations are easily derived (see e.g. Giodini et al., 2013). We show the derivation for the mass - temperature relation below. For the following derivation the following conventions are used: the density can be written as $\rho = \triangle \rho_{\rm c}$ where $\rho_{\rm c}$ is the critical density (see equation 1.10) at the cluster redshift; $R_{\triangle}$ corresponds to the radius within which the mean density is $\triangle$ times the critical density at the cluster's redshift; and $M_{\triangle}$ is the mass enclosed within $R_{\triangle}$.[3]

For a dynamically relaxed cluster, the gas and galaxies will be related by the virial theorem

$$2K = -U \tag{1.25}$$

where K is the kinetic energy and U is the potential energy. The average kinetic energy per particle for a monatomic gas with temperature, T, is

$$\langle K_i \rangle = \frac{3}{2} kT \tag{1.26}$$

where k is the Boltzmann's constant. The total kinetic energy of a gas with N particles is $N\langle K_i \rangle$, so

$$K \propto NkT \propto M_{gas,\triangle} kT \tag{1.27}$$

[3]Using $R_{\triangle}$ as a radius within which cluster quantities can be measured means that clusters can be compared to each other in a consistent way regardless of their redshift and mass (assuming self-similarity). Common values of $\triangle$ used to report cluster quantities in literature are 2500, 500, and 200.

where M_{gas} is the gas mass. For self-similar clusters, $M_{gas,\triangle} \propto M_{\triangle}$ where $M_{\triangle}$ is the total cluster mass, so

$$K \propto NkT \propto M_{\triangle}kT \tag{1.28}$$

We now have the kinetic energy of the system and want the potential energy U which we can write as

$$U \propto \frac{GM_{\triangle}^2}{R_{\triangle}} \tag{1.29}$$

Writing the virial theorem in terms of equation 1.28 and 1.29 we obtain

$$M_{\triangle}kT \propto \frac{M_{\triangle}^2}{R_{\triangle}} \tag{1.30}$$

Then, writing $R_{\triangle}$ in terms of the cluster density such that $R_{\triangle} \propto M_{\triangle}^{1/3}\rho^{-1/3}$ and substituting this into equation 1.30 we obtain

$$M_{\triangle} \propto (kT)^{3/2}\rho^{-1/2} \tag{1.31}$$

Using equation 1.10 and writing the redshift dependence of H as $H(z) = E(z)H_0$, where the factor $E(z)$ is an increasing function of redshift z that depends on the cosmological parameters as defined in equation 1.16, it is clear that $\rho \propto \triangle E(z)^2$, which when substituted into equation 1.31 gives us the mass - temperature scaling relation

$$M_{\triangle} \propto kT^{3/2}\triangle^{-1/2}E(z)^{-1} \tag{1.32}$$

Cluster temperature is therefore expected to increase with total cluster mass, and for a fixed mass, the cluster temperature increases with redshift.

Some other X-ray scaling relations (see e.g. Giodini et al., 2013) are the L - T (luminosity - temperature) relation, and the L - M (luminosity - mass) relation. Another is the M-Y_X relation (Kravtsov et al., 2006), where Y_X = $M_{\rm gas,500} \times T$ and is related to the thermal energy of the ICM. The M-Y_X relation has a significantly lower scatter than the $L - M$ and $M - T$ relation, and so Y_X is sometimes used preferentially as a mass proxy. There also exist scaling relations where one (or both) of the parameters is not derived from X-ray data, such as the σ - T relation, where the σ refers to the statistical velocity dispersion about the mean for the member galaxies of a cluster.

1.5.2 Observed scaling relations

Scaling relations from observations are useful for two main reasons. Firstly, large cluster samples with mass measurements are needed to accurately constrain cosmological parameters, but performing a full hydrostatic mass analysis from X-ray data requires high quality data, and is a complicated process (see Chapter 3). Thus, if there is only low quality data available for a set of clusters, the L - M relation can be used to get an estimate of the cluster masses (e.g. Reiprich & Böhringer, 2002; Pratt et al., 2009), as the luminosities of the clusters can be accurately calculated from lower quality data. For estimating mass from an X-ray observable such as cluster luminosity, clearly we would want to use the L - M scaling relation derived from observations of clusters, and not from theory. Secondly, observed departures from the theoretical models provide information on the processes that have been neglected in the derivation of the theoretical scaling models. For example, the theoretical model of the L - T relation predicts $L \propto kT^2$, but observationally a steeper slope is found (e.g. Maughan et al., 2012). It is likely that non-gravitational processes not included in the self-similar model are responsible for this, by raising the cluster temperature and lowering the ICM density. The effect is stronger compared to gravitational energy in lower mass systems, which leads to a steepening of the relation.

1.6 Cluster morphology and cool cores

1.6.1 Morphology

Understanding the morphological state of galaxy clusters is extremely important in cluster studies. One reason is related to the importance of obtaining accurate mass estimates of a large representative sample of clusters, required in order to do cosmology with clusters (see §1.2). As mentioned in §1.7, the X-ray hydrostatic mass method relies on the assumption that the cluster is in hydrostatic equilibrium. The validity of this assumption is highly dependent on the dynamical state of the cluster and its ICM, and for a dynamically disturbed cluster, this assumption is likely to be less valid than in a dynamically relaxed cluster. Thus, for accurate cosmological studies, the breakdown of this assumption for the mass estimates of these disturbed clusters needs to be taken into account. Additionally, the state of a cluster can affect the detection probability of a cluster in an X-ray survey (as relaxed clusters are more likely to have a cool core, see §1.6.2), such that if a representative sample of clusters in our Universe is desired for further study, there

will be a bias towards selecting cool core systems if using X-ray selection, which will need to be accounted for (see §1.9).

Clusters have different dynamical states due to the nature of their evolution, growing via the hierarchical merging of smaller objects. A disturbed cluster might have recently undergone a merger or a significant interaction with another cluster, and is likely to not be spherically symmetric. In a merger, the kinetic energy from the gas of both of the clusters involved is converted into thermal energy through shocks and turbulence, and on a $\sim$Gyr timescale the combined gas from both clusters will fall into hydrostatic equilibrium with the deeper gravitational potential of the now merged clusters. This will eventually lead to a relaxed cluster, provided there are no more interactions with other clusters in this time period. There are various diagnostics to determine if a cluster is disturbed or not. One of the many methods is measuring the standard deviation of the distance between the X-ray peak of the cluster and the centroid where this distance is measured in different aperture sizes (e.g. Poole et al., 2006, also presented in more detail in §4.3.10). The power ratio is also a useful diagnostic - it parametrises the amount of substructure in the ICM and relates it to the morphology of a cluster (Buote & Tsai, 1996; Weißmann et al., 2013).

1.6.2 Cool core clusters

From X-ray observations of galaxy clusters, many clusters (at low redshift) are observed to have a central surface brightness peak. The inferred radiative cooling times of these surface brightness peaks are significantly less than the ages of the clusters. It was proposed by Fabian & Nulsen (1977) that this was due to a cooling flow, where the central gas cools quickly due to its high density (as bremsstrahlung emission is proportional to the square of the gas density, see §1.3.1), causing the gas to flow inwards from outside this central region and cool. However, this would imply large amounts of star formation in the central regions, as the cool gas would condense and form stars. This is not seen in the level predicted by the cooling flow model, but there still is some star formation and cold gas present (e.g. Edge, 2001; McDonald et al., 2012). Additionally, the very low ICM temperatures predicted by this model are not seen either, although there is a significant temperature drop in the central region of these clusters (see Figure 1.10), and a significant increase in gas density in the cores of these clusters. It is now believed that there is a feedback mechanism in the central region which re-heats the cool gas sufficiently such that it suppresses star formation. There have been various feedback mechanisms proposed,

the most probable being energy input due to AGN activity and supernovae, and turbulence and conduction within the ICM (see Peterson & Fabian, 2006 for a comprehensive review). In the case of AGN feedback, the general model now proposed is that cool gas clumps form due to ICM cooling, and one of these clumps might feed the AGN, which in turn reheats this cool gas (Pizzolato & Soker, 2005). ICM cooling thus enhances AGN activity.

The revision of the cooling flow model due to lack of observational support led to the proposal of cool-core (CC) clusters (Molendi & Pizzolato, 2001). CC clusters are believed to be dynamically relaxed (Leccardi et al., 2010), as the effect of undergoing a merger has a large impact on the cluster core, which either disrupts the CC region more directly (e.g. Fabian & Daines, 1991; Russell et al., 2012; Rossetti et al., 2013; Wang et al., 2016) or mixes the cool gas with the surrounding hotter gas (ZuHone et al., 2010). Thus, CC clusters are often used for determining cluster masses for cosmological studies, as the assumption of hydrostatic equilibrium is more likely to be satisfied. There are many diagnostics to split cluster samples into CC and non cool-core (NCC) clusters (see Hudson et al., 2010 for a comprehensive study of different diagnostics). The most widely used X-ray methods include measuring the drop in temperature in the cluster core (O'Hara et al., 2006), the central cooling time (e.g. Bauer et al., 2005; Santos et al., 2010), the cuspiness of the 3D gas density profile in the cluster core (Vikhlinin et al., 2007), or the concentration parameter (Santos et al., 2008). We present the central cooling time and cuspiness diagnostic in more detail in §4.3.10.

We note that other wavelength regimes can provide additional information that can be used to select CC clusters. Optical data can be used to locate the BCG, and by measuring the offset between the brightest cluster galaxy (BCG) and the X-ray peak (Hudson et al., 2010), one can select clusters with negligible offsets to select CCs; additionally, infrared data can be used to measure the star formation rate in the BCG, which, if high, can also indicate the presence of a CC (Rawle et al., 2012).

1.7 Estimating Galaxy Cluster Masses

1.7.1 X-ray Hydrostatic Masses

The X-ray hydrostatic mass method can be the most precise method of obtaining cluster masses providing long observation times are used. However, the method assumes that the cluster's gravitational potential and the gas pressure originating solely from the ICM are in hydrostatic

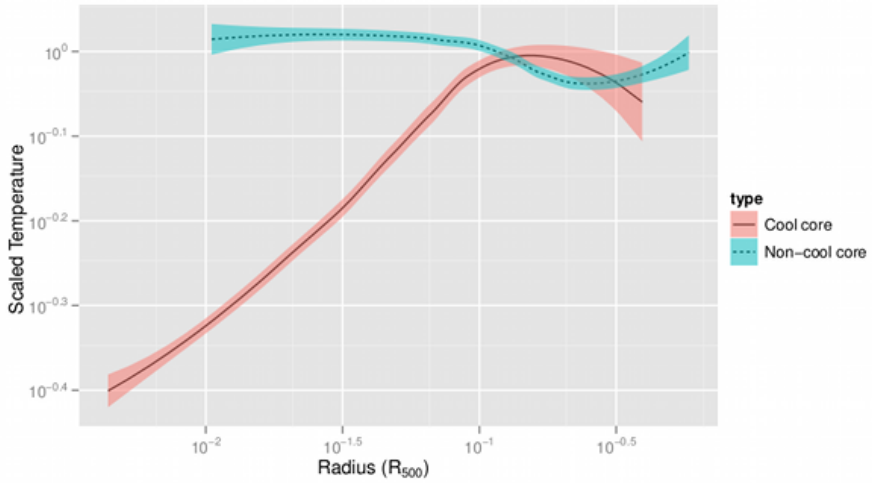

Figure 1.10: Mean (deprojected) temperature profiles for CC (red solid line) and non-CC (blue dotted line) clusters for a sample of 20 clusters from Sanderson et al. (2006). The shaded regions represent the 1σ errors. Taken from Sanderson et al. (2006).

equilibrium, as we can only measure the thermal pressure from X-ray data. This means that if there are any non-thermal pressure sources in the cluster, this can lead to a systematic underestimation of the true cluster mass. The work presented in Chapter 4 investigates the level of this bias.

We present two methods of obtaining the cluster mass from X-ray data in detail in Chapter 3.

1.7.2 Optical Masses

Galaxy cluster masses are obtained from optical observations of clusters using two main methods. One method utilizes the position and velocity information of the cluster galaxies (galaxy dynamics), and the other utilizes the apparent distortion in the shapes of background galaxies due to the bending of light rays by the cluster's gravitational potential. One advantage of using optical masses over X-ray masses is that they are independent of the state of the ICM, unlike X-ray masses, which rely on the assumption of hydrostatic equilibrium of the ICM. Additionally, the caustic and WL mass estimation technique are the only methods that allow accurate cluster mass estimation out to large cluster radius.

a) Caustic masses

Caustic masses (see Diaferio, 2009 for a review) are one example of dynamical masses. It is a relatively new method, developed in the late 20^{th} century, and uses spectroscopic redshift measurements of a large number ($\gtrsim$50) of member galaxies to determine the line of sight velocities (relative to the centre of mass of the cluster). By plotting the line of sight velocity versus the projected distance from the cluster centre for each of the member galaxies (referred to as a redshift diagram), an overpopulated region of this parameter space is clearly seen (see Figure 1.11), and is bounded by a 'trumpet shape'. The boundaries between galaxies within this trumpet shape and outside it are known as caustics. The galaxies within this overpopulated parameter space are assumed to be bound by the cluster's gravitational potential, and their velocities are assumed to be less than the escape velocity of the cluster, such that the over-density of galaxies lie within the escape velocity of the cluster. The amplitude, $A(r)$, of this trumpet shape is related to the cluster's escape velocity, and decreases as a function of projected distance from the cluster centre. Diaferio & Geller (1997) show that this caustic amplitude, $A(r)$, can be related to the cluster mass within a radius, r:

$$GM(< r) = \mathcal{F}_\beta \int_0^r A^2(r)\, dr \tag{1.33}$$

where $\mathcal{F}_\beta$ is a filling factor that is constant with radius. The filling factor $\mathcal{F}_\beta$ comes from filling function $\mathcal{F}_\beta(r)$, which combines the cluster density profile and the anisotropy of the density field, both of which vary as a function of r, and the approximation of this function to a constant does lead to overestimating the cluster mass within $0.5R_{200}$ by $\sim 15\%$ (Serra et al., 2011). Beyond this radius, the filling function is indeed roughly constant with radius, so using a constant filling factor does not lead to significant overestimation or underestimation of mass at larger radii.

b) Weak Lensing Masses

Perhaps the most widely-used mass estimation method that uses optical data is the WL method (see Hoekstra et al., 2013 for a review). Light rays from galaxies that pass through the mass distribution of a galaxy cluster will be bent, as predicted by general relativity. WL masses can be obtained by measuring the distortions of the shapes of background galaxies, as a more massive cluster will bend the light rays more strongly, thus leading to greater average galaxy

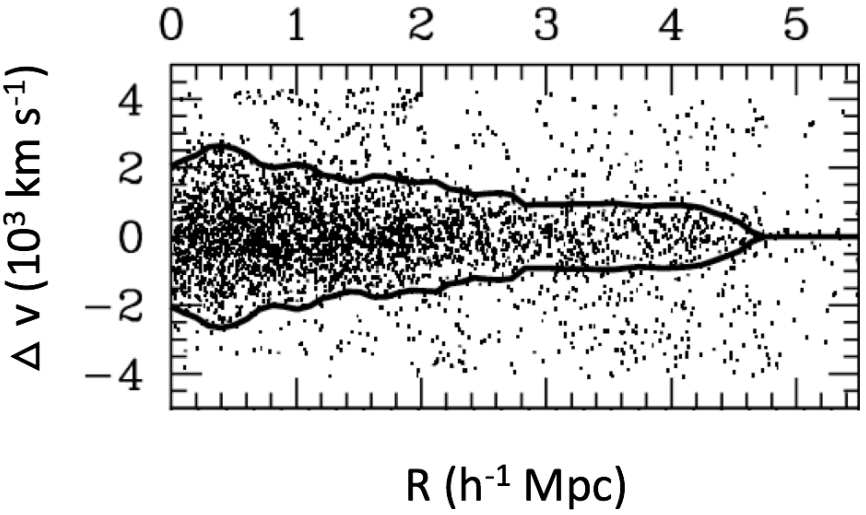

Figure 1.11: Redshift versus projected radius from the cluster centre for member galaxies for an ensemble of clusters. The solid black lines show the caustics. Adapted from Rines et al. (2013).

distortion. As this distortion is very subtle for an individual galaxy, and galaxies themselves have different shapes (before the distortion), the shapes of a large number of galaxies are measured. The redshift distribution of the background galaxies also needs to be measured, as the lensing effect also depends on the redshift of each background galaxy, and the foreground needs to be ignored. This statistical nature of WL is still currently its limiting factor, limiting the precision with which masses can be measured, and WL mass estimates have significantly larger errors than X-ray mass estimates. In addition, WL is susceptible to projection along the line of sight - the method assumes that the only mass causing distortions in the background galaxies is the cluster mass, but if there are other massive objects along the line of sight that are not accounted for, then the distortions will also be affected by these, and will lead to overestimates of the cluster mass.

1.7.3 SZ Masses

As the SZ signal arises from the interaction of the CMB photons and the ICM, the SZ effect will be larger in more massive clusters as there will be more gas in more massive clusters to upscatter the CMB photons. However, there is not a direct relation between the SZ signal and the cluster mass; in fact, as the SZ signal just gives the integral of the pressure along the line of sight, it is necessary to calibrate the SZ signal as a mass proxy using some other mass measurement

method, such as X-ray hydrostatic masses. A current weakness of this method is that the results of cluster cosmology studies (see §1.2) using cluster masses derived from their SZ signal are highly reliant on the correct calibration of these SZ masses. Recently, there has been a focus on understanding a tension between the cosmological results from the *Planck* CMB experiments and the *Planck* cluster number counts experiments (e.g. Douspis et al., 2019), and the source of this tension could largely originate from the uncertain calibration of the SZ signal as a mass proxy.

1.8 Galaxy Cluster Surveys

Galaxy clusters can be detected in the X-ray band, the optical band, and the radio band, using the methods described in §1.3, and each detection method has its own strengths and weaknesses.

1.8.1 X-ray Surveys

X-ray surveys are an efficient method to detect clusters of galaxies, with a low contamination rate. The first X-ray survey was carried out with the *Uhuru* satellite (Giacconi et al., 1972), in the early 1970s, and showed a correlation between the presence of overdensities of galaxies in the optical band, and extended X-ray emission. The *HEAO-1* Observatory (Rothschild et al., 1979) followed *Uhuru*, and undertook an all-sky survey (Piccinotti et al., 1982), from which 30 galaxy clusters were identified. However, due to their poor spatial resolution and sensitivity, these early X-ray surveys suffered from confusion, and were only able to identify clusters at redshifts $z < 0.1$. The *Einstein* observatory (*HEAO-2*) followed (Giacconi et al., 1979), with increased sensitivity and spatial resolution, leading to the detection of 100s of X-ray sources, and enabling the identification of clusters out to higher redshifts (EMSS, Gioia et al., 1990b).

Following these observatories' successful missions, came the launch of *ROSAT* in 1990, with significantly improved sensitivity and spatial resolution, and its all-sky survey led to the production of cluster samples with 100s of clusters (e.g. BCS Ebeling et al., 1998, REFLEX Böhringer et al., 2001, and MACS Ebeling et al., 2001 among others) that are still useful today. In fact, because of the low instrumental background of the *ROSAT-PSPC* detector, *ROSAT* observations are still used to study the outskirts of clusters (e.g. Eckert et al., 2012, 2015).

Another leap forward in X-ray astronomy came with the launch of the *XMM-Newton* and *Chandra* observatories in 1999 (see §1.4 for more detail on these observatories). Due to their

smaller field of view (especially *Chandra*), an all-sky survey has not been carried out with either observatory. This does not mean that they cannot carry out useful X-ray cluster surveys; instead, these observatories typically look at smaller regions of the sky in more detail.

Cluster surveys carried out using *XMM-Newton* include the *XMM*-LSS survey (Pierre et al., 2006), a pilot for the larger *XMM*-XXL survey (Pierre et al., 2016) which has detected hundreds of clusters out to a redshift $z \sim$1.2, and has extensive radio and optical follow-up observations. All *XMM-Newton* cluster surveys mentioned thus far target specific areas of the sky, however there currently exists the XCS (Mehrtens et al., 2012), a cluster survey that uses all public *XMM-Newton* data available to compile a cluster catalogue for further study. Other serendipitous X-ray cluster surveys include RCDS (Rosati et al., 1998), 400D (Burenin et al., 2007) and WARPS (Scharf et al., 1997), which all use archival *ROSAT* data, as well as EMSS (Gioia et al., 1990b), which uses data from *Einstein*.

eROSITA (Pillepich et al., 2012) is the latest addition to currently operational X-ray telescopes, and was launched in mid 2019. It is currently carrying out an all sky survey, and is expected to detect $\sim$100,000 clusters up to redshift $z \sim 1.3$, enabling the mapping of large scale structure in our Universe such that its evolution over cosmic time can be studied.

1.8.2 Optical Surveys

Optical surveys can be carried out by looking for overdensities of galaxies on the sky, as this is how a galaxy cluster will appear. However, this method can be prone to projection effects, as structure along the line of sight will also appear as an over-density of galaxies on the sky. This over-density search method was used to identify clusters as early as the 18^{th} century when Herschel found large concentrations of objects (called nebulae at the time, but now understood to be galaxies) in the region of the sky where the Virgo cluster is found. In the 20^{th} century, the catalogues of optical galaxies compiled by Abell (1958) and Zwicky et al. (1961) enabled the identification of a large number of clusters. More recently, the Sloan Digital Sky Survey (SDSS, Hao et al., 2010) has found tens of millions of galaxies, leading to the identification of tens of thousands of clusters (e.g. Wen et al., 2009). The cluster identification methods that use SDSS data typically utilize either colour information or the photometric redshift (or spectroscopic redshift if available) of each galaxy to check that the galaxies in the overdense region have similar redshifts, decreasing the probability that the cluster detection is, in fact, due to projection effects.

Clusters can also be detected using the WL technique by measuring the average distortion of a large number of galaxies over a region of the sky (see §1.7 b) for more detail). If the average distortion is high, it is inferred that a cluster is present. WL surveys are free from assumptions about the state of the ICM (unlike X-ray surveys), however, due to projection effects (described in §1.7 b)), false detections are possible. WL surveys include the Deep Lensing Survey (Wittman et al., 2001), the Subaru WL Survey (Miyazaki et al., 2007), and lensing surveys using the CFHT (Gavazzi & Soucail, 2007; Erben et al., 2013). WL surveys have thus far been carried out over a maximum of a few hundred square degress, due to the number of galaxies required for a single WL cluster detection. Euclid, a space mission with a current launch date of mid-2022, and LSST, a ground based telescope with full science operations survey expected to begin in 2022/2023, will detect billions of galaxies and enable a new generation of lensing surveys, with unprecedented scale and accuracy.

1.8.3 SZ Effect Surveys

Notable SZ surveys have been carried out using the South Pole Telescope (SPT), and the *Planck* satellite. The SPT 2500 deg^2 survey has detected hundreds of clusters (Bleem et al., 2015), and over a thousand clusters have been identified with *Planck* (Planck Collaboration et al., 2016c) in an all-sky survey. The Atacama Cosmology Telescope (ACT) has also undertaken surveys detecting 100s of clusters (Hasselfield et al., 2013; Hilton et al., 2018).

1.9 Issues facing X-ray Surveys and X-ray observations of clusters

1.9.1 Biases

The two main biases affecting X-ray cluster surveys are referred to as the Malmquist and Eddington biases. The Eddington bias comes from the fact that all cluster surveys are flux limited (due to instrumental constraints), and because lower mass clusters are more common than higher mass clusters, more lower mass clusters will be observed above the flux limit than one might expect given their underlying population, due to the intrinsic scatter of their luminosity with mass and measurement errors. More generally, the Eddington bias can apply to any survey with a detection threshold, and if the detection property has scatter with mass, then there will be an Eddington bias in the observed sample of objects. The Malmquist bias, which also affects all surveys, also originates from the flux limit, and refers to the fact that due to this flux limit, at

high redshift, only the most luminous clusters will be selected, leading us to over-sample the high luminosity end of the cluster population. These two effects need to be accounted for when trying to determine the true underlying cluster population, as needs to be done for accurate cosmological studies (see §1.2).

1.9.2 Cool cores

A cool core cluster (see §1.6.2 for more detail) is so named due to the observed temperature drop in the core of the cluster. This drop in temperature is generally accompanied by an increase in the ICM density, leading to a very luminous core (as the intensity of the emission increases as a function of the square of the gas density - see §1.3.1 a)). X-ray surveys are more likely to detect cool core clusters, which are generally also morphologically relaxed. This is because X-ray surveys are surface brightness limited, and for cool core clusters their luminosity is concentrated in a smaller area (Pesce et al., 1990; Fabian, 1994). Hence, an X-ray selected cluster sample will preferentially select a certain population of clusters (i.e. those that are relaxed). The fact that morphology affects cluster detection probability in X-ray cluster surveys also needs to be accounted for when trying to determine the true underlying cluster population for cosmological studies.

1.9.3 Point source contamination

Active Galactic Nuclei (AGN) will vastly outnumber clusters in any given X-ray observation. Some of these AGN will appear in the same region of sky as a cluster, and if unresolved (due to the relatively poor spatial resolution of X-ray telescopes) will lead to an artificially high luminosity being measured for the cluster. Additionally there may be AGN in the cluster itself that are unresolved. Unresolved point sources will also affect the detection probability of a cluster, either enhancing it, or leading to the cluster being rejected altogether. Thus, the unresolved AGN contamination of the clusters in an X-ray selected sample needs to be well understood.

Point sources can also be an issue in X-ray cluster detection as multiple closely-spaced point sources may be mistaken for a cluster. This is especially true for high redshift ($z \gtrapprox 1$) clusters, which appear smaller on the sky. This effect also needs to be accounted for when trying to determine the true underlying cluster population and we present one method to address this issue in detail in Chapter 2.

1.10 Simulations

Although this thesis focuses on using observational data to study clusters, the use of simulations is incredibly important in cluster science. Simulations of clusters have progressed to a point where they very accurately reproduce the scaling relations obtained from real cluster observations, as well as ICM density profiles (see Borgani & Kravtsov, 2011 for a review). Simulations of large volumes can be used to model the evolution of structure in our Universe and what we should observe in our Universe, given different cosmological parameters, aiding cosmological studies using clusters. Current large scale cosmological simulations include the BAHAMAS simulations (McCarthy et al., 2017), and the IllustrisTNG simulations (Nelson et al., 2019). Perhaps the most impactful simulation produced is the Millenium simulation (Springel et al., 2005) - to date over 1000 publications have made use of the Millenium simulation data [4].

Higher resolution simulations, focussing on smaller volumes, can also be used to test the impact of various physical processes in clusters, as in simulations, these can turned on or off, and completely controlled. An example of this is probing the processes in the central region of simulated cool-core clusters (Rasia et al., 2015). Another example is that by creating mock X-ray images of clusters from simulations, where the true mass, for example, is known, biases in X-ray based mass measurement techniques can be studied (Nagai et al., 2007); this can also be done at any other wavelength to probe the biases present in other mass estimation methods. Another use of simulations is to select certain interesting systems, such as galaxy cluster mergers, and make mock X-ray images of them to investigate what observational signatures they would have, in order to then use these observational signatures to detect merging systems from real X-ray observations (Molnar, 2015).

1.11 Structure of this thesis

The outline of this thesis is as follows. In Chapter 2 I present the results of *Chandra* follow up of 21 high redshift galaxy clusters detected in the XXL survey. In Chapter 3 I detail how to estimate cluster masses using X-ray data. In Chapter 4 I present work that uses X-ray hydrostatic and caustic mass profiles of a sample of 44 clusters in order to probe the hydrostatic bias. In Chapter 5 I use the same data and sample of clusters as in the previous chapter, now focussing

[4] https://wwwmpa.mpa-garching.mpg.de/millennium/

on gas fraction profiles out to large cluster radii. In Chapter 6 I present the conclusions of the thesis, and outline future work.

Throughout this thesis we assume a WMAP9 cosmology of $H_0 = 69.7$ km/s/Mpc, $\Omega_\Lambda = 0.718$, and $\Omega_m = 0.282$ (Hinshaw et al., 2013). All measurement uncertainties correspond to 68% confidence level, unless explicitly stated otherwise. Throughout this chapter, when referring to base 10 logarithm we use '$\log$', and use '$\ln$' if referring to the natural logarithm. For the calculation of the values of the mean molecular weight in the ICM, μ (0.611 atomic mass units, a.m.u.), and the mean molecular weight per electron and proton, $\mu_{\rm e}$ (1.173 a.m.u.) and $\mu_{\rm p}$ (1.407 a.m.u.) respectively, that we use in Chapters 3 and 4 , we use data on the composition of astrophysical plasmas from Cox (2000), assuming a 0.3 solar abundance plasma.

2

Chandra Constraints on the AGN Contamination of z $>$ 1 XXL Galaxy Clusters

Note: This chapter is a reproduction of a paper with the same title, published in the Astronomy & Astrophysics Journal. The reference for this paper is Logan C.H.A., Maughan B.J., Bremer M.N., Giles P., Birkinshaw M., Chiappetti L., Clerc N., Faccioli L., Koulouridis E., Pacaud F., Pierre M., Ramos-Ceja M.E., Vignali C., Willis J., 2018, A&A, 620, A18. The manuscript is reproduced here under the non-exclusive right of re-publication granted by ESO to the author(s) of the paper.

2.1 Abstract

Context: The XMM-XXL survey has used observations from the *XMM-Newton* observatory to detect clusters of galaxies over a wide range in mass and redshift. The moderate PSF (FWHM $\sim 6''$ on-axis) of *XMM-Newton* means that point sources within or projected onto a cluster may not be separated from the cluster emission, leading to enhanced luminosities and affecting the

selection function of the cluster survey. Aims: We present the results of short *Chandra* observations of 21 galaxy clusters and cluster candidates at redshifts $z > 1$ detected in the XMM-XXL survey in X-rays or selected in the optical and infra-red. Methods: With the superior angular resolution of *Chandra* we investigate whether there are any point sources within the cluster region that were not detected by the XMM-XXL analysis pipeline, and whether any point sources were misclassified as distant clusters. Results: Of the 14 X-ray selected clusters, nine are free from significant point source contamination, either having no previously unresolved sources detected by *Chandra*, or with less than about 10% of the reported XXL cluster flux being resolved into point sources. Of the other five sources, one is significantly contaminated by previously unresolved AGN, and four appear to be AGN misclassified as clusters. All but one of these cases are in the subset of less secure X-ray selected cluster detections and the false positive rate is consistent with that expected from the XXL selection function modelling. We also considered a further seven optically-selected cluster candidates associated with faint XXL sources that were not classed as clusters. Of these, three were shown to be AGN by *Chandra*, one is a cluster whose XXL survey flux was highly contaminated by unresolved AGN, while three appear to be uncontaminated clusters. By decontaminating and vetting these distant clusters, we provide a pure sample of clusters at redshift $z > 1$ for deeper follow-up observations, and demonstrate the utility of using *Chandra* snapshots to test for AGN in surveys with high sensitivity but poor angular resolution.

2.2 Introduction

Galaxy cluster surveys provide us with large, well-controlled samples of clusters that enable us to place constraints on cosmological models through tests of the growth of structure. For the tightest constraints on the cosmological parameters, we need a large look-back time, with samples that include clusters at $z > 1$. These high-redshift clusters enable the study of the astrophysical processes that drive galaxy and cluster evolution over cosmic time.

Although galaxy cluster surveys can be carried out at different wavelengths (e.g. Rosati et al., 1998; Böhringer et al., 2004; Gladders & Yee, 2005; Eisenhardt et al., 2008; Rozo et al., 2010; Sehgal et al., 2013; Planck Collaboration et al., 2014; Stanford et al., 2014), searching for extended X-ray emission has the advantage that the cluster candidates that are identified are

much less likely to suffer from projection effects than selecting clusters based on galaxy overdensities which can contain projections of galaxies along the line of sight that are not associated with virialized systems. This is because a given amount of gas dispersed in clumps and filaments will be much fainter in X-rays than the same gas confined and compressed in a single potential well, as is the case in a cluster, where this gas is termed the intra-cluster medium (ICM). This is due to bremsstrahlung emissivity (the main emission mechanism in a cluster) being proportional to the square of the density of the gas.

X-ray surveys have proven very effective in identifying large numbers of galaxy clusters (e.g. Gioia et al., 1990a; Ebeling et al., 1998; Rosati et al., 1998; Böhringer et al., 2004; Pierre et al., 2004; Ebeling et al., 2010; Fassbender et al., 2011; Mehrtens et al., 2012; Willis et al., 2013; Pierre et al., 2016, hereafter XXL Paper I) including many at redshifts $z > 1$, with the most distant clusters found up to a redshift $z \approx 2$ (Nastasi et al., 2011; Santos et al., 2011; Willis et al., 2013; Mantz et al., 2014, hereafter XXL Paper V).

While X-ray surveys are effective at finding clusters of galaxies, clusters are vastly outnumbered by active galactic nuclei (AGN), which dominate extragalactic X-ray source counts. With sufficient angular resolution, clusters are resolved, allowing these two classes to be separated. However, for clusters at cosmological distances this becomes challenging because of the low surface brightness of the cluster emission and the fact that the detected emission from these distant clusters can have angular extents similar to (or smaller than) the PSF of most X-ray observatories. This can lead to AGN being misclassified as clusters or a compact cluster being misclassified as AGN (e.g. Somboonpanyakul et al., 2018).

It is also possible for a genuine cluster detection to be contaminated by X-ray emission from an unresolved AGN in, or projected onto the cluster, giving rise to various issues (e.g. Giles et al., 2012). Most importantly a cluster with AGN contamination will have its flux and temperature overestimated (Branchesi et al., 2007). This has implications for the use of luminosity or temperature as a mass estimator to carry out cosmological studies (reviewed by Allen et al., 2011), or for studies of the scaling relations between cluster properties (e.g. Pratt et al., 2009; Maughan et al., 2012; Giles et al., 2016, also known as XXL Paper III). Unresolved AGN in or projected onto clusters also alter the apparent surface brightness distribution of the cluster which can enhance or decrease its detection probability making it difficult to understand the selection function of cluster surveys at the level needed for cosmological studies. An additional complication is that AGN in galaxy clusters are significantly more common at higher redshift. Galametz

et al. (2009) found that X-ray selected AGN are at least three times more prevalent in clusters at $1 < z < 1.5$ than in clusters at $0.5 < z < 1$ (though we note that the sample used by Galametz et al., 2009 is effectively a volume limited survey of clusters and so has different selection to the XXL survey). This is a higher increase in AGN density than that seen in the field population of AGN (Martini et al., 2013). Other work using X-ray selected clusters (Ruderman & Ebeling, 2005) also show X-ray point sources to be more prevalent in clusters than the field. However, for low mass clusters ($\lesssim 3 \times 10^{14}$ $M_{\odot}$) at z < 1 there is evidence that the density of X-ray selected AGN in X-ray selected clusters is consistent with the field (Koulouridis et al., 2014). Optically selected AGN in optically selected clusters show similar agreement between the AGN fraction in clusters and the field (Marziani et al., 2017), but with some indication that the AGN fraction can be higher in compact groups (Martínez et al., 2010).

The problem of AGN contamination of X-ray cluster surveys can be addressed statistically by using realistic models of the population of AGN in and projected onto distant clusters in the calibration of the selection function. The state-of-the-art is the use of full cosmological hydrodynamical simulations which include self-consistent modelling of cluster and AGN populations (Koulouridis et al., 2017). The observational data upon which to base such models are sparse, and this project was the first systematic attempt to observationally survey the AGN content of distant X-ray selected galaxy clusters. Similar work can also now be found in Biffi et al. (2018). The AGN contribution to individual distant clusters has previously been studied (e.g. Hilton et al., 2010), and the cosmic evolution of AGN in clusters has been studied using optical/infrared-selected clusters, including z > 1 clusters (Galametz et al., 2009), but to our knowledge this is the first time that distant clusters detected in an X-ray survey have been looked at, so this work has particular bearing for X-ray cluster surveys.

Our work uses the XXL survey (XXL Paper I), which is the largest survey carried out by the *XMM-Newton* satellite and covers a total area of 50 deg^2 distributed over two fields (XXL-N and XXL-S). *XMM-Newton* has an on-axis half energy width (HEW) PSF of $\sim 15''$ which degrades and becomes increasingly asymmetric as a function of distance from the aimpoint. The XXL survey's primary aim is to investigate the large-scale structure of the Universe using the distribution of galaxy clusters (and AGN) as tracers of the matter distribution. The survey has detected several hundreds of galaxy clusters out to a redshift of $z \approx 2$ (365 in the most recent list, Adami et al., 2018, referred to as XXL Paper XX hereafter) above an X-ray flux limit of $\sim 5 \times 10^{-15}$ erg s^{-1} cm^{-2} in the 0.5 - 2 keV band. We study a set of 21 $z > 1$ clusters and

candidates using short *Chandra* observations to assess the level of AGN contamination. We use the term "candidates" in recognition of the fact that some of the sources without spectroscopic confirmation or flagged as less reliable by the X-ray detection pipeline may not be genuine clusters. The main aims of this work are to quantify the contribution of unresolved point sources to the XXL detection of extended ICM emission and flag for rejection those candidate clusters where the XXL detection is fully resolved into one or more point sources by *Chandra*. This decontamination is made possible by *Chandra*'s on-axis sub-arcsecond PSF. This work is especially important given the upcoming launch of *eROSITA* (Merloni et al., 2012). *eROSITA*'s all-sky survey is expected to detect $\sim 10^5$ clusters out to redshifts $z > 1$ (Pillepich et al., 2012) and will have on-axis spatial resolution similar to that of *XMM-Newton* and so will face the same challenges as *XMM-Newton* in resolving point sources in distant clusters.

The structure of the paper is as follows. In Section 2.3 we discuss the sample selection and data preparation. Section 2.4 details the data processing steps. Notes on individual clusters are given in Section 2.5. We discuss our results in Section 2.6. The conclusions are presented in Section 2.7. Throughout this paper we assume a WMAP9 cosmology of $H_0 = 70$ km/s/Mpc, $\Omega_\Lambda = 0.72$, and $\Omega_{\rm m} = 0.28$ (Hinshaw et al., 2013).

2.3 Sample and Data Preparation

Our sample was initially constructed to comprise the 15 $z > 1$ clusters and cluster candidates from the XMM-LSS survey (a $\sim$10 deg^2 precursor to, and subset of XXL; Willis et al., 2013). The redshifts of two of those clusters (XLSS J022252.3-041647 and XLSSU J021712.1-041059) were subsequently revised to be at z < 1, so were dropped. Two of the remaining Willis et al. (2013) clusters had existing *Chandra* archival data, the other 11 were targeted with new *Chandra* snapshot observations. We subsequently expanded our sample to include a further four $z > 1$ clusters detected in the wider XXL survey that have available Chandra data. The full 50 deg^2 XXL survey contains a further seven $z > 1$ clusters for which we have been awarded Chandra observations, four of which have been observed and are included in this work, while the remaining three clusters have yet to be observed. Our final sample thus contains 21 $z > 1$ clusters and candidates in total.

The XXL source detection pipeline XAMIN ranks clusters into classes (Pacaud et al. (2006), Pacaud et al. (2016) - hereafter XXL Paper II, Faccioli et al. (2018) - also known as XXL Paper

XXLID	ObsID	Class Willis	Class XXL	z	RA (J2000)	Dec. (J2000)	F_{60} (10^{-14} erg s^{-1} cm^{-2})	Chip configuration	Clean time (ksec)
XLSSC 072	18253	C1	C1	1.00	33.850	-3.726	4.1±0.4	ACIS-S	9.9
XLSSC 029	7185	C1	C1	1.05	36.017	-4.225	3.2±0.3	ACIS-S	31.9
XLSSC 005	18256	C1	C1	1.06	36.788	-4.301	0.9±0.2	ACIS-S	10.9
3XLSS J021825.9-045947	17306		C1	1.13$^{\diamond}$	34.608	-4.997	0.4±0.1	ACIS-I†	50.8
XLSSC 122	18263	C1	C1	1.99	34.433	-3.759	1.3±0.3	ACIS-S	10.6
XLSSC 048	18254	C1	C2	1.01	35.722	-3.473	1.1±0.3	ACIS-S	9.4
XLSSC 073	18255	C1	C2	1.03	33.744	-3.506	0.7±0.3	ACIS-S	17.9
3XLSS J022755.7-043119	20534		C2	1.05$^{\diamond}$	36.982	-4.522	0.3±0.3‡	ACIS-S	31.6
3XLSS J021320.3-053411	20535		C2	1.08$^{\diamond}$	33.334	-5.570	$0.1^{+0.2\ddagger}_{-0.1}$	ACIS-S	35.2
XLSSC 203	17304		C2	1.08	34.428	-4.989	0.2±0.1	ACIS-I†	44.7
XLSSC 634	11741		C2	1.08	355.691	-54.185	4.8±0.6	ACIS-I†	62.7
3XLSS J021325.0-042000	20536		C2	1.20$^{\diamond}$	33.355	-4.334	1.5±0.5‡	ACIS-S	9.9
3XLSS J022005.5-050826	13374	C2	C2	1.65$^{\diamond}$	35.023	-5.141	0.6±0.2‡	ACIS-I	75.7
3XLSS J022418.4-043956	18262	C2	C2	1.67$^{\diamond}$	36.077	-4.666	0.6±0.2‡	ACIS-S	11.9
XLSSC 034	20538		C3	1.04	35.372	-4.099	2.1±0.9	ACIS-S	9.9
3XLSS J022059.0-043922	18257	C2	C3	1.11$^{\diamond}$	35.246	-4.657	0.9±0.3‡	ACIS-S	9.8
XLSSC 046	18259	C2	C3	1.22	35.763	-4.606	0.7±0.2	ACIS-S	20.8
3XLSS J022351.3-041841	6390		C3	1.27$^{\diamond}$	35.963	-4.313	0.9±0.2‡	ACIS-S†	10.8
3XLSS J021700.4-034746	18260	C2	C3	1.54$^{\diamond}$	34.251	-3.796	0.7±0.2‡	ACIS-S	9.9
3XLSS J022812.3-043836	18261	C2	C3	1.67$^{\diamond}$	37.051	-4.644	0.4±0.1‡	ACIS-S	9.6
3XLSS J022554.3-045059	18264	C2	C3	2.24$^{\diamond}$	36.476	-4.850	0.2±0.2‡	ACIS-S	21.7

Table 2.1: Summary of the cluster sample and *Chandra* data. Column 1 is the cluster name; column 2 is the *Chandra* ObsID; column 3 is the cluster class (see section 2) from Willis et al. (2013) or blank if the cluster is not part of that sample; column 4 is the cluster class from the updated XXL pipeline; column 5 is the redshift of the cluster (from XXL Paper XX or for those not in that paper, the redshifts have not yet been published); columns 6 and 7 are the RA and Dec. coordinates of the *XMM* X-ray peak (from XXL Paper XX); column 8 is the cluster flux in the 0.5 - 2 keV energy band measured in the 60″ cluster region using XXL data (those that are not included in XXL Paper XX are marked with a ‡); column 9 is the CCD chip configuration for the observation where a † means that the cluster fell off-axis in the observation - the off-axis distance is given in Section 2.5; column 10 is the cleaned *Chandra* observation time. Redshifts that are photometric are marked with a $\diamond$

XXIV). Galaxy cluster candidates are selected from the XAMIN maximum likelihood outputs in EXT, EXT_STAT and EXT_DET_STAT, which correspond to the extent, likelihood of extent, and detection significance, respectively. A source is considered extended if it has measured EXT greater than 5″ and EXT_STAT greater than 15. The extended sources are then sorted into categories: the C1 class selects candidates with an EXT_STAT greater than 33 and a EXT_DET_STAT greater than 32; the C2 class comprises the remaining candidates. The C1 sample is expected to be mostly free of contamination by point sources. The C2 sample is expected to be about 50% comprised of misclassified AGN, image artefacts and other spurious detections (Pierre et al., 2006; Adami et al., 2011), though it is worth noting that the contamination of the final C2 sample is likely to be significantly lower than this, as all cluster candidates are visually inspected, and obvious spurious sources are rejected. There exists a third class, the C3 sample, which consists of clusters known from optical/IR catalogues, that are associated with some X-ray emission that is too weak to be characterised (see Pierre et al., 2006, or XXL Paper XX). However, despite this, not all cluster candidates are expected to be genuine clusters: it is possible that in some cases where a cluster has been identified by XXL, there could just be a galaxy overdensity coincident with one or more AGN. The classifications were calibrated by simulations where the pipeline was run on previous *XMM* observations with model clusters and randomly distributed AGN added (Pacaud et al., 2006, 2007; Clerc et al., 2012). These observations were restricted to low redshift clusters, and the purpose of this work is to extend this to lower signal-to-noise high redshift clusters which is more challenging due to the high redshift clusters often not being resolved, and there being poorer quality supporting data.

The XXL analysis pipeline has been upgraded since the work reported in Willis et al. (2013), leading to some changes in classification for individual objects (XXL Paper XXIV). For the present analysis, we are using cluster classifications and properties consistent with those in the latest data release (XXL Paper XX). Throughout this paper we often refer to the updated pipeline results, which are the results from XAMIN consistent with the version used in XXL Paper XX.

Our sample consists of five C1 clusters, nine C2 clusters and seven C3 clusters. Three C2 clusters (3XLSS J022755.7-043119, 3XLSS J021320.3-053411,3XLSS J021325.0-042000) and 1 C3 cluster are reported here for the first time. Table 2.1 shows the properties of the clusters in our sample. The cluster flux in the 0.5 - 2 keV energy band measured in the 60″ cluster region using XXL data, F_{60}, reported in Table 2.1 in column 8, was computed using a growth curve analysis as described in XXL Paper II (either taken from XXL Paper XX or recomputed directly

by us for objects not included in this paper). Two clusters (XLSSC 072 and XLSSC 029) are in the XXL 100 brightest galaxy cluster sample (XXL Paper II) and ten clusters (all C1s, 4 C2s - XLSSC 048, XLSSC 073, XLSSC 203, XLSSC 634 and 1 C3 - XLSSC 034) are in XXL Paper XX.

The clusters in our *Chandra* snapshot programme that were not covered by archival data were observed with the ACIS-S configuration with an exposure time designed to give a significant detection of a point source contributing $> 10\%$ of the 0.5 - 2 keV band XXL flux for C1s and spectroscopically confirmed C2s and $> 25\%$ for other cluster candidates. A minimum exposure time of 10 ks was imposed on all observations. The snapshot observations were not designed to detect significant emission from the ICM, although a borderline significant detection was expected in some cases. For those clusters already covered by archival data, two were in the ACIS-S configuration and four in the ACIS-I configuration (see Table 2.1). In some of the archived observations, the cluster fell relatively far from the optical axis, leading to a larger PSF than for an on-axis observation, which sometimes caused complications in the analysis (see Section 2.5).

All 21 clusters in our sample were analysed with the CIAO[1] 4.9 software package and CALDB[2] version 4.7.4 (Fruscione et al., 2006). The level 1 event files were reprocessed using the `chandra_repro` tool following the standard data reduction threads[3]. Periods of background flares were identified and removed using lightcurves analysed with the `deflare` tool. For observations taken in the ACIS-S configuration the cluster always fell on only the S3 chip, so a lightcurve was extracted from only the S3 chip. For the observations in the ACIS-I configuration a lightcurve was extracted from the four front illuminated (FI) chips, CCD_IDs I0-I3 (excluding any other chips in the observation). The CCDs not used for the lightcurve filtering were discarded from the rest of the analysis.

In Figures 2.1, 2.2 and 2.3 we show optical and *Chandra* images for the C1, C2 and C3 clusters respectively.

[1]See `http://cxc.harvard.edu/ciao`
[2]See `http://cxc.harvard.edu/caldb`
[3]See `http://cxc.harvard.edu/ciao/threads/index.html`

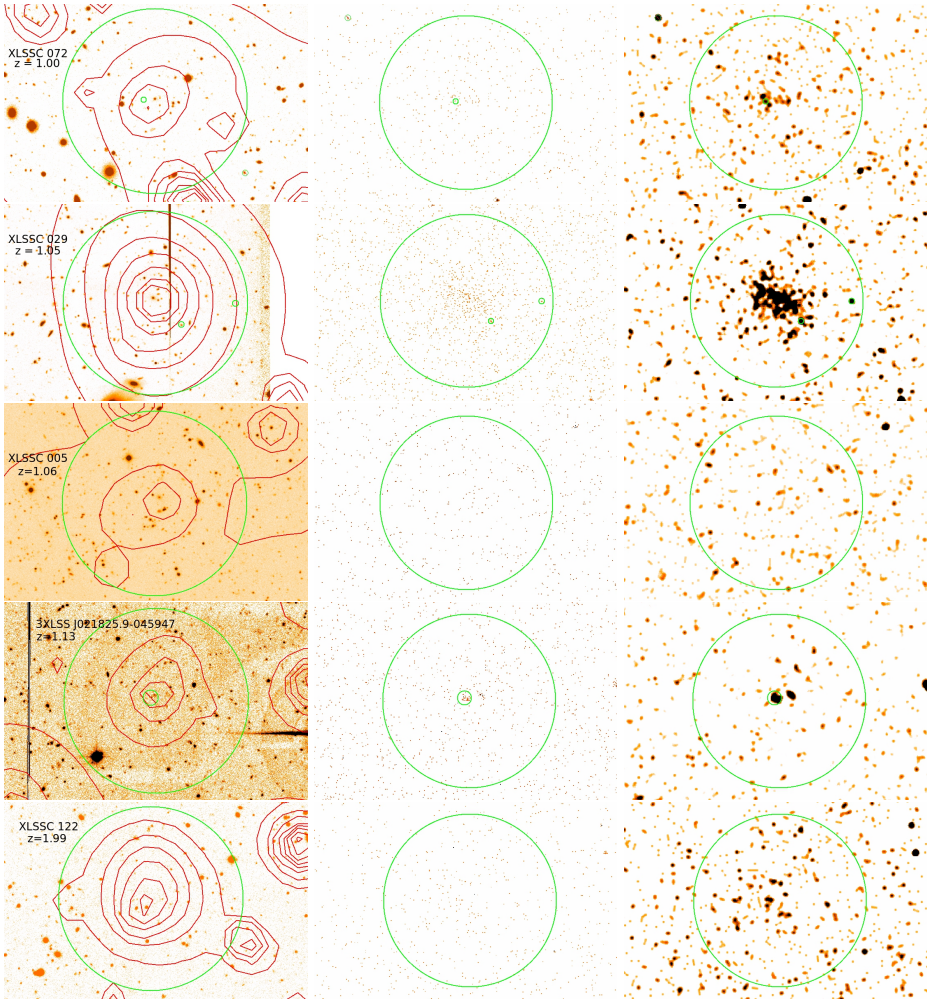

Figure 2.1: A comparison of the optical image with the *XMM-Newton* contours from the 0.5 - 2 keV band (red) superimposed (left) and the raw and smoothed (using a Gaussian with $\sigma \sim 2.5''$) *Chandra* (centre and right, respectively) images for all C1 clusters. All optical images are i-band images from the CFHTLS except for 3XLSS J021825.9-045947 which is r-band. *Chandra* images are in the 0.3 - 8.0 keV band. The green circle is the same in all images and is of radius $60''$ and centred on the *XMM* X-ray peak. Point sources within $60''$ of the *XMM* X-ray peak are marked by the smaller green circles in all images. In the raw *Chandra* images, if a *Chandra* point source was detected in XXL then it is circled in red.

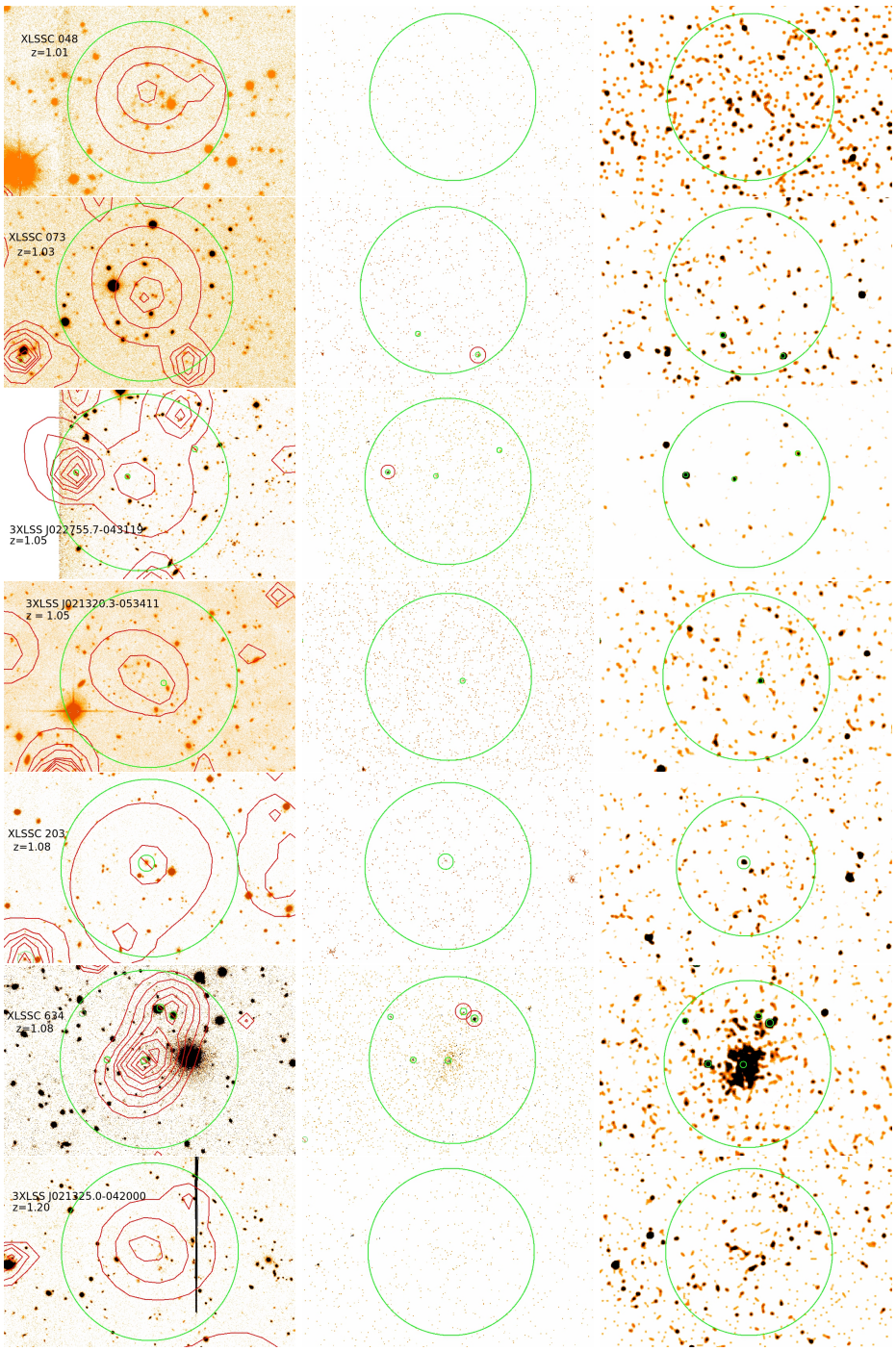

Figure 2.2: Same as Figure 2.1 but for all C2 clusters. All optical images are i-band images from the CFHTLS except for XLSSC 203 which is r-band and XLSSC 073 which is g-band. The green circle is the same in all images and is of radius 60″ and centred on the *XMM* X-ray peak. Point sources within 60″ of the *XMM* X-ray peak are marked by the smaller green circles in all images. In the raw *Chandra* images, if a *Chandra* point source was detected in XXL then it is circled in red.

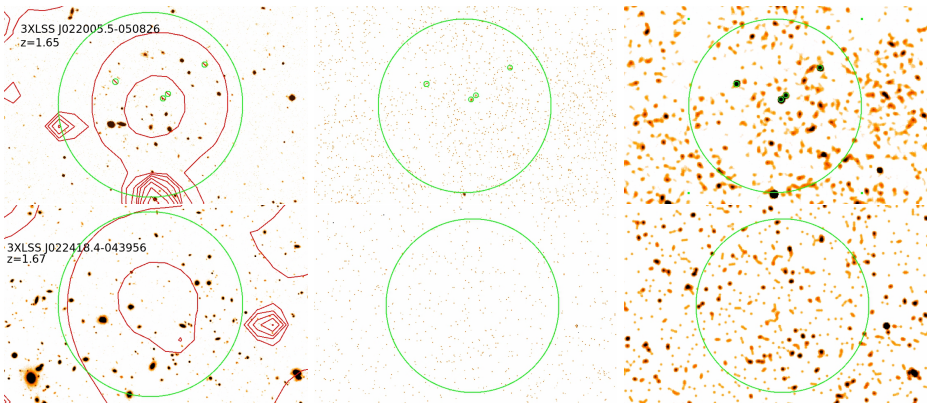

Figure 2.2: *- continued*

2.4 Data Processing

The main focus of our analysis is both to obtain flux constraints for detected sources, and to determine upper limits for possible sources that were not detected. For source detection we use the CIAO `wavdetect` tool, and for photometry the CIAO `srcflux` tool was used. The `srcflux` tool uses a Bayesian method to compute the background-marginalised posterior probability distribution of the source flux. `srcflux` has three possible outcomes: a "good measurement" where the probability distribution function (PDF) is not truncated at zero for the confidence interval specified, so the lower limit is given as well as the most probable flux and upper limit; "pdf truncated at zero" where the most probable flux and upper limit are given, but the lower limit is not given as the PDF is truncated at zero for the confidence interval specified; "mode of zero" where the most probable flux is zero and a lower limit is therefore not given, but an upper limit is still given.

In the following section we describe the detection and photometry of point sources in the *Chandra* data in or projected onto the cluster regions. We assume that all point sources detected are AGN, as AGN vastly outnumber any other contaminating point sources at this depth - the possibility that they could be X-ray bright stars is $\sim$3% (Galametz et al., 2009,Chiappetti et al., 2018 - also known as XXL Paper XXVII). For several clusters, point sources were detected in these regions by the XXL pipeline and excluded from the XXL cluster flux measurements. Since the goal of our analysis is to estimate the effects of AGN that were unresolved by *XMM* we do not include the point sources that were detected by XXL in the main body of this paper. These

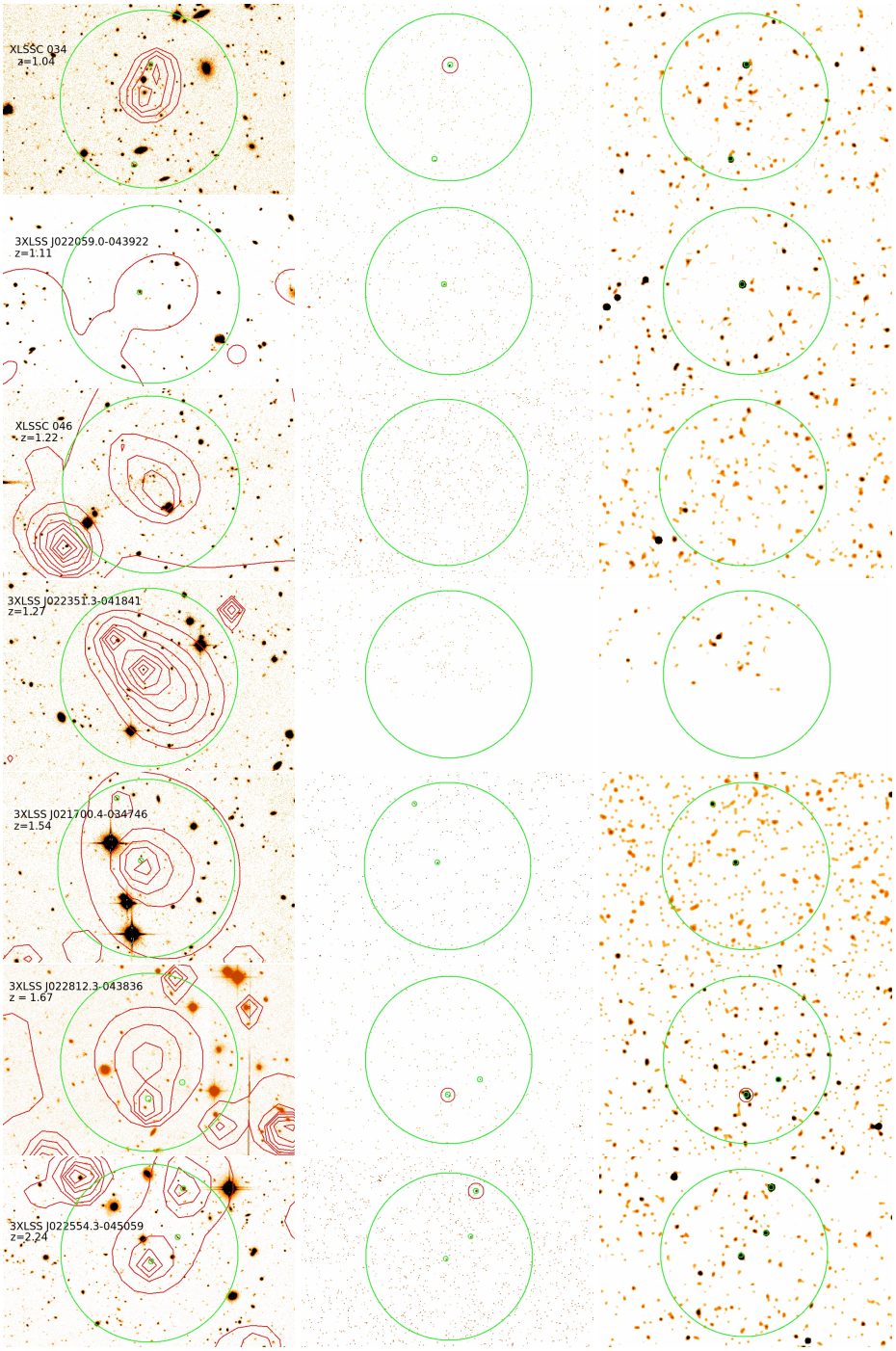

Figure 2.3: Same as Figure 2.1 but for all C3 clusters. All optical images are i-band images from the CFHTLS. The green circle is the same in all images and is of radius 60″ and centred on the *XMM* X-ray peak. Point sources within 60″ of the *XMM* X-ray peak are marked by the smaller green circles in all images. In the raw *Chandra* images, if a *Chandra* point source was detected in XXL then it is circled in red.

sources are detailed in Table 2.3.

2.4.1 Point Source Detection and Flux Calculation

For the purpose of point source detection, images and the appropriate exposure maps were produced in the 0.3 - 8 keV band (Kim et al., 2007). The CIAO `wavdetect` tool was used to search for point sources in these images. The `scales` parameter was set as $(\sqrt{2})^n$ with n = 0 - 8 and the `sigthresh` parameter was set to 1×10^{-6} such that there will be $\sim$4 false-positive source detections per image for the 4 FI chips in the ACIS-I observations and $\sim$1 for the S3 chip in the ACIS-S observations. Since we are considering only the 60″ region around the cluster (as this is the region within which cluster fluxes are measured in the XXL pipeline - we note that there is a possibility of point sources being present outside this 60″ region, but these should not affect the XXL measurement of the X-ray flux within the 60″ cluster region), the false positive rate will be $\sim$0.05 false-positive source detections per cluster, corresponding to $\sim$1 false positive in the full sample of clusters. The detection limit corresponds to $\sim$5 photons from the source aperture in `wavdetect`.

In some cases where the cluster fell off-axis, due to the observation being from pre-existing *Chandra* data not specifically designed to observe the cluster, there was ambiguity as to whether a detected source was a point source or ICM emission. There were also cases where no source was detected by `wavdetect` but a visual inspection suggested a possible point source in or projected onto the cluster region. In order to be conservative in our classification of whether point sources were present, we flagged as possible point sources any regions within 60″ of the *XMM* X-ray peak that possessed either (i) at least 4 counts in a single pixel, or (ii) at least 6 counts in a 1″ circle with at least one pixel containing 2 or more counts. This formalised our visual inspection enabling us to apply it to simulated images when determining upper limits as described below.

Multi-wavelength data were used to assist the classification of these possible point sources, and details for each are given in Section 2.5. For the optical band we used the Canada-France-Hawaii Telescope Legacy Survey (CFHTLS) data for XXL-N[4]. These images were taken with the wide field optical imaging camera MegaCam, a 340 Megapixel camera with a 1′ by 1′ field of view. For XLSSC 634 in XXL-S the image was taken from the BCS survey (Desai et al., 2012)

[4] `http://www.cfht.hawaii.edu/Science/CFHTLS`

with the Mosaic2 imager on the Blanco 4m telescope[5]. For radio data we used the NRAO VLA Sky Survey (Condon et al., 1998) and Tasse et al. (2008) for the XXL-N field and used Australia Telescope Compact Array (ATCA) data (Smolčić et al. (2016) - also known as XXL Paper XI, Butler et al. (2017) - also known as XXL Paper XVIII) for the XXL-S field (for XLSSC 634). We define an optical or radio source as a likely counterpart to a *Chandra* detected point source if it falls within $2''$ of the *Chandra* detected point source coordinates. The matching radius is chosen to be $2\ ''$ as it is large enough to account for any astrometry errors, and not too large so as to introduce many spurious matches.

Fluxes were then measured for all point sources detected within $60''$ of the *XMM* X-ray peak (as in XXL Paper XX), assuming a power law model with $\Gamma = 1.7$, consistent with the modelling used in other XXL papers (Fotopoulou et al. (2016) - also known as XXL Paper VI, XXL Paper XXVII); however, since we are measuring the flux in a relatively narrow band (compared to the full *Chandra* bandpass), without needing to extrapolate, and with too few counts to fit the spectral index, the exact choice of spectral index is not too important. The source region was set to be the 90% encircled energy radius of the PSF at 1 keV and the background region was an annulus centred on the same coordinates as the source region, with the inner radius equal to the source radius, and the outer radius five times greater than the inner radius. The `psfmethod` option in `srcflux` was set to `quick`, which uses the radius of the source circle to obtain the PSF fraction in the specified energy band, and assumes that the background region contains 0% of the source flux, so the effect of any source flux that falls in the background region is neglected - we note that this means a small amount of flux is lost, but the effect should not be large. The absorbing column, N_H, was fixed at the Galactic value (Kalberla et al., 2005): $\approx$2 - 2.5×10^{20}cm^{-2} for all clusters except XLSSC 634 which had $N_H\approx1.5\times10^{20}$cm^{-2}). All of the `wavdetect` detected point sources had "good measurements" from `srcflux`, except for XLSSC 072 which had "mode of zero" for its flux measurement so we report this as a 1σ upper limit.The fluxes are reported in column 6 in Table 2.2.

For those clusters that had no point sources detected within $60''$ of the *XMM* X-ray peak, we determined an upper limit on the flux of any undetected point source. For each cluster we simulated an image of a point source, using the *Chandra* PSF at the detector position of the *XMM* X-ray peak, and normalised to a particular point source flux. Poisson noise was added and the point source was added to the original *Chandra* image at the *XMM* X-ray peak. We then

[5] `http://www.ctio.noao.edu/noao/node/9`

applied the same detection method used on the original data and recorded whether the simulated point source was detected. This process was repeated for 100 realisations of the Poisson noise for a given point source flux. The source flux was then varied until the simulated source was detected in 68% of the realisations, and the corresponding flux was defined as the 1σ upper limit on the flux of an undetected point source. This value is reported in column 6 of Table 2.2. The upper limits are driven by the Poisson noise on the low number of counts expected from the faint point source and hence can be significantly larger than the measured flux for detected point sources in comparable observations.

To estimate the possible contribution of point sources to the cluster flux measured with *XMM*, we compute the AGN contamination fraction. The AGN contamination fraction is the contribution of the combined flux from all of the point sources detected by *Chandra* (or upper limits for those clusters with no point sources detected) within $60''$ of the *XMM* X-ray peak (that were not detected by XXL and excluded from the XXL flux calculation) as a fraction of F_{60} (see column 4, 6 and 7 in Table 2.2). These cluster fluxes are updated compared to those from Willis et al. (2013), and calculated using the updated version of the XXL analysis pipeline. Figures 2.1, 2.2 and 2.3 show images of the clusters in the sample, and indicate the positions of point sources that were detected by XXL and/or by the *Chandra* follow-up observations. Those detected by XXL were already excluded from the F_{60} values and so do not contribute to the AGN contamination fractions calculated here. As mentioned above, the contamination was calculated as the combined point source flux (or the upper limit in the case of clean clusters) of those point sources not previously resolved by XXL as a fraction of the cluster flux. Therefore, a cluster with a contamination $\gtrsim 1$ can be thought of as being a misclassified point source(s). Lower, but non-zero, values suggest that the XXL flux comes from a blend of cluster and point source emission.

2.4.2 Calculating Cluster Fluxes from the *Chandra* data

The *Chandra* snapshot observations were optimised to detect significant point source contamination in the XXL clusters, and are not expected to be deep enough to measure detailed ICM properties. Nonetheless we attempted to place constraints on the ICM flux from the *Chandra* data. All of the point sources in the image were masked using a circle with a radius necessary to include 90% of the flux at 1 keV, and the flux from each cluster was estimated using `srcflux`. A $60''$ radius circle was used as the source region (consistent with the XXL flux measurements),

and the background region used was an annulus with inner and outer radii of 120″ and 180″ respectively, as measured from the *XMM* X-ray peak. In some cases this background region went off chip and this was accounted for. An absorbed APEC thermal plasma model (Smith et al., 2001) was used to model the cluster flux. The absorption was set at the Galactic value (Kalberla et al., 2005), the metal abundance set to 0.3 solar, and the plasma temperature to 3.5 keV (typical of high redshift XXL clusters, XXL Paper XX). The redshifts used are in Table 2.1. If the 3σ lower bound on the PDF of the flux in this region was non-zero, then we treated this as a definite detection of ICM emission with *Chandra*. This was the case for five clusters. In 11 other cases, an ICM flux measurement was still possible, but the 3σ lower bound extended to zero flux. In the remaining cases the mode of the posterior distribution for the flux was zero, so only an upper limit was measured.

The effect of masking the point sources means some cluster emission is also lost from the masked region. The effect of this will be greatest for off-axis sources where the PSF and therefore the mask size is greatest. 3XLSS J021825.9-045947 has the largest PSF at the *XMM* X-ray peak of all observations where a point source is detected in the 60″ cluster region (see Figure 2.1). The masked region accounts for $\sim$0.5% of the cluster area in the 60″ region. Modelling the cluster emission as a beta-model (Cavaliere & Fusco-Femiano, 1976) with $\beta = 0.66$ and a core radius of 175 kpc and assuming that the point source mask is at the *XMM* X-ray peak (as this will maximise the amount of presumptive ICM flux lost) it is found that $\sim$2.5% of the total cluster emission from the 60″ region is masked. Thus we can ignore this effect as the difference is much smaller than our 1σ errors on the cluster fluxes (see Table 2.2).

2.5 Notes on Individual Clusters

In this section we note any instances where we departed from the analysis described in Section 3 and other points of interest. In all cases, when PSF sizes are reported, we give the 90% encircled energy radius at 1 keV.

For each cluster/cluster candidate below we give the name, *Chandra* ObsID, XXL class, and categorise its level of AGN contamination based on all of the data available. CC indicates a "clean cluster" with a low level of AGN contamination; PC indicates a cluster that is "partially contaminated" from the point sources previously unresolved in XXL; FC indicates a "fully contaminated" cluster (i.e. most likely a point source - or multiple point sources - that

was misclassified as extended). This information is also given in column 8 in Table 2.2.

3XLSS J021825.9-045947 / ObsID 17306 / C1 / FC - This cluster fell 2.8′ off-axis in an archived observation, where the PSF is 4.09″ compared with 0.83″ on-axis. A source was detected at the *XMM* X-ray peak but due to the larger PSF at the source position it is not clear whether this is a genuine point source or a detection of extended emission. However, the X-ray source is coincident with a radio source and an unresolved optical source so we conclude it is likely to be a radio-loud quasar, and treat it as a point source. In addition, our `dmstat` search method identified a potential point source that was undetected by `wavdetect`, $\sim 5''$ from the source that was detected at the *XMM* X-ray peak. From the optical data, there is a likely optical counterpart to this possible X-ray source that appears slightly extended in nature so is likely to be a galaxy. We thus conclude that this source (if real) is likely to be an AGN in that galaxy rather than a detection of the ICM. We do not include this undetected point source when calculating the cluster contamination, however if we were to include it the AGN contamination fraction would rise from 0.67 to 0.90. In either case it appears likely that the XXL detection is a misclassified AGN or pair of AGN and not a genuine extended source.

XLSSC 122 / ObsID 18263 / C1 / CC - This cluster is at z = 1.99 (based on results in Mantz et al., 2018b, hereafter XXL Paper XVII, using X-ray spectroscopy) and is the most distant cluster discovered by XXL to date (see XXL Paper XX). It has a Sunyaev-Zel'dovich effect detection (XXL Paper V) and deep *XMM* follow up (XXL Paper XVII). `wavdetect` found no point sources in the larger 60″ circular region around the *XMM* X-ray peak, and inspecting the image visually confirms this. We therefore computed an upper limit for contamination as described in Section 2.4.1. We first reported a 3σ upper limit on the flux contamination of 8% in XXL paper XVII. Using the same *Chandra* data, we here place a 1σ upper limit of 18% on the flux of any undetected point source. This weaker constraint is due to the more rigorous and conservative definition of an upper limit in the current work (see Section 2.4.1)

3XLSS J021320.3-053411 / ObsID 20535 / C2 / FC - This cluster has one point source detected in the 60″ cluster region by `wavdetect`. In addition, our `dmstat` search method identified a potential point source that was undetected by `wavdetect`, at 33.345, -5.56. There is no optical or radio counterpart for this X-ray source, and we do not include this source when calculating the cluster contamination; however its flux is $0.02 \pm 0.02 \times 10^{-14}$ erg s^{-1} cm^{-2} and if we were to include it the AGN contamination fraction would rise from 1.2 to 1.4. In either case it appears likely that the XXL detection is a misclassified AGN or pair of AGN and not a

genuine extended source.

XLSSC 203 / ObsID 17304 / C2 / FC/PC - This cluster fell 2.9′ off-axis in an archived observation, where the PSF is 4.59″ compared with 0.83″ on-axis. A point source was detected close to the *XMM* X-ray peak, and upon visual inspection of the image it is clear that this is genuinely a point source (and not extended emission). The flux of this point source is about half of the XXL cluster flux, but the fluxes agree within the measurement errors, so this cluster could be partially or fully contaminated.

XLSSC 634 / ObsID 11741 / C2 / CC - This cluster fell 1.4′ off-axis in an archived observation, where the PSF is 1.75″ compared with 0.83″ on-axis. A source was detected at the *XMM* X-ray peak but due to the larger PSF it is not clear whether this is a genuine point source or a detection of extended emission. We do not find any radio or optical counterparts to this source, but conservatively treat it as point source emission for the analysis. However, if we were to treat it as ICM emission then the AGN contamination fraction would drop from 0.10 to 0.05.

3XLSS J022005.5-050826 / ObsID 13374 / C2 / FC - For this cluster, the XXL F_{60} value (see Table 2.2) has a large error, and the total flux from the 4 point sources detected in the 60″ cluster region is consistent with a partially contaminated cluster and also consistent with F_{60} coming solely from AGN emission. When we mask all point sources and measure the *Chandra* cluster flux (see Section 2.6.2), we find the cluster flux to be zero, with a low upper limit, and thus we conclude that most likely there is no cluster emission from 3XLSS J022005.5-050826, and it is multiple AGN misclassified as extended ICM emission.

XLSSC 046 / ObsID 18259 / C3 / CC - This is a genuine cluster (Bremer et al., 2006), with an overdensity of optical and IR galaxies, but is compact in the optical and X-ray, leading to its re-classification from a C2 in a previous pipeline version (Willis et al., 2013) to a C3 with the current XXL pipeline. We did not detect any point sources in the 60″ cluster region with our *Chandra* data.

3XLSS J022351.3-041841 / ObsID 6390 / C3 / FC/PC/CC - This cluster fell 3.7′ off-axis in an archived observation, where the PSF is 6.80″ compared with 0.83″ on-axis. The centre of the cluster falls mostly on-chip, but part of the cluster emission falls off-chip. No point sources were detected in the available cluster region, so an upper limit was computed following the normal method.

3XLSS J022812.3-043836 / ObsID 18261 / C3 / FC/PC - `wavdetect` detects a point source previously detected by XXL within 60” of the *XMM* X-ray peak, and for this point source the

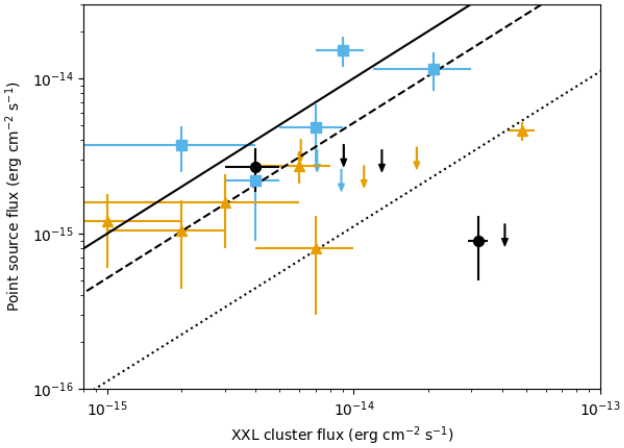

Figure 2.4: We show the total *Chandra* flux for point sources within 60" of the *XMM* X-ray peak versus the *XMM* cluster flux. C1 clusters are black circles, C2s are yellow triangles and C3s are blue squares. Arrows indicate clusters that only have a 1σ upper limit for their point source flux (column 6 in Table 2.2) - the tip of the arrow denotes the upper limit. The solid straight line is a line of equality showing locus of 100% AGN contamination and the dashed and dotted lines are lines of equality showing the the locus of 50% and 10% AGN contamination, respectively. 1σ errors are shown.

position of the centre of the ellipse enclosing the source region as detected by `wavdetect` is slightly offset from the peak pixel position when visually inspecting the image. We therefore computed the source flux at the position of the peak pixel rather than the `wavdetect` source position. When masking the point sources for the cluster flux calculation we increased the point source mask size by $1.5''$ to ensure all of the point source emission was masked. The point source flux is reported in Table 2.3, but the point source is not included in the AGN contamination fraction as it was previously detected by XXL.

3XLSS J022554.3-045059 / ObsID 18264 / C3 / FC - `wavdetect` detects three point sources within 60" of the *XMM* X-ray peak. For one of the point sources, the position of the centre of the ellipse enclosing the source region as detected by `wavdetect` is slightly off from the peak pixel position when visually inspecting the image. We treated this as for 3XLSS J022812.3-043836.

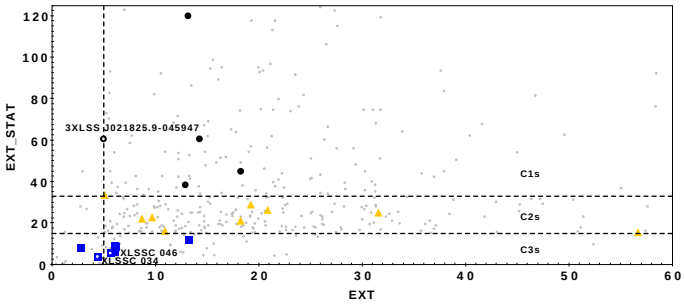

Figure 2.5: We show the EXT - EXT_STAT parameter space for the C1s, C2s and C3s in our sample (larger black circles, yellow triangles and blue squares respectively). We also show a representative sample of C1, C2 and C3 XXL clusters at $0<z<1$ for illustration (smaller grey circles). The C1/C2/C3 boundaries are explained in Section 2.3. The three larger circles/squares with the hollow centres are those with labels on the plot.

XXLID	Class XXL	z	F_{60} (10^{-14} erg s^{-1} cm^{-2})	No. of point sources	*Chandra* point source flux (10^{-14} erg s^{-1} cm^{-2})	AGN contamination fraction	Final assessment	*Chandra* cluster flux (10^{-14} erg s^{-1} cm^{-2})
XLSSC 072*	C1	1.00	4.1±0.4	1	<0.08	<0.02	CC	$3.41^{+0.85}_{-0.82}$
XLSSC 029*	C1	1.05	3.2±0.3	2	0.09±0.04	0.03	CC	$3.63^{+0.30}_{-0.29}$
XLSSC 005*	C1	1.06	0.9±0.2	0	<0.26	<0.29	CC	$1.19^{+0.69}_{-0.67}$
3XLSS J021825.9-045947	C1	1.13	0.4±0.1	1	0.27±0.09	0.67†	FC	$0.32^{+0.25}_{-0.23}$
XLSSC 122*	C1	1.99	1.3±0.3	0	<0.24	<0.18	CC	$1.98^{+0.79}_{-0.77}$
XLSSC 048*	C2	1.01	1.1±0.3	0	<0.19	<0.17	CC	$0.85^{+0.67}_{-0.63}$
XLSSC 073*	C2	1.03	0.7±0.3	1	0.08±0.05	0.11	CC	$0.46^{+0.41}_{-0.37}$
3XLSS J022755.7-043119	C2	1.05	0.3±0.3	2	0.16±0.05	0.53	FC	<0.37
3XLSS J021320.3-053411	C2	1.08	$0.1^{+0.2}_{-0.1}$	1	0.12±0.06	1.2	FC	$0.35^{+0.32}_{-0.13}$
XLSSC 203	C2	1.08	0.2±0.1	1	0.10±0.06	0.50	PC	$0.54^{+0.29}_{-0.29}$
XLSSC 634	C2	1.08	4.8±0.6	3	0.46±0.09	0.10†	CC	$5.62^{+0.35}_{-0.35}$
3XLSS J021325.0-042000	C2	1.20	1.5±0.5	0	<0.25	<0.17	CC	$1.75^{+0.87}_{-0.83}$
3XLSS J022005.5-050826*	C2	1.65	0.6±0.2	4	0.27±0.08	0.45	FC	<0.09
3XLSS J022418.4-043956*	C2	1.67	0.6±0.2	0	<0.28	<0.47	CC	<0.47
XLSSC 034	C3	1.04	2.1±0.9	1	0.15±0.13	0.07	CC	$2.52^{+0.88}_{-0.83}$
3XLSS J022059.0-043922*	C3	1.11	0.9±0.3	1	1.52±0.34	1.7	FC	$0.22^{+0.65}_{-0.22}$
XLSSC 046*	C3	1.22	0.7±0.2	0	<0.24	<0.34	CC	$0.99^{+0.49}_{-0.47}$
3XLSS J022351.3-041841	C3	1.27	0.9±0.2	0	<0.18	<0.20†	FC/PC/CC	$0.17^{+0.36}_{-0.17}$
3XLSS J021700.4-034746*	C3	1.54	0.7±0.2	2	0.48±0.20	0.69	FC	<0.45
3XLSS J022812.3-043836*	C3	1.67	0.4±0.1	1	0.22±0.13	0.55	FC/PC	$0.54^{+0.51}_{-0.48}$
3XLSS J022554.3-045059*	C3	2.24	0.2±0.2	2	0.37±0.12	1.9	FC	$0.07^{+0.47}_{-0.07}$

Table 2.2: Summary of point source detection and cluster contamination from the *Chandra* data. The *Chandra* cluster flux measurement is also shown. Column 4 is the XXL cluster flux. Column 5 gives the number of point sources detected by `wavdetect` within a 60″ radius region around the *XMM* X-ray peak that were not previously detected by XXL. Column 6 gives the total flux of all of the point sources detected by `wavdetect` within a 60″ region around the *XMM* X-ray peak that weren't detected by XXL,with the 1σ lower and upper limits are given as error. All fluxes are in the 0.5 - 2 keV energy band. Column 7 gives the fraction of F_{60} resolved into point sources by Chandra, as described in Section 2.4.1. Column 8 gives our assessment of the cluster. Column 9 is the cluster flux as calculated from *Chandra* data after point source removal (described in Section 2.4.2) with 1σ errors. XXLIDs marked with a $\star$ appear in Willis et al. (2013) and are therefore part of the XMM-LSS subset of clusters. AGN contamination fractions marked with a † have possible additional contamination from potential point sources that did not meet our detection threshold (except for XLSSC 634 which has a potentially lower AGN contamination fraction than stated in this table), and contamination values that include these sources are given in Section 2.5. Individual point source fluxes and positions are given in Table 2.3 in Appendix A

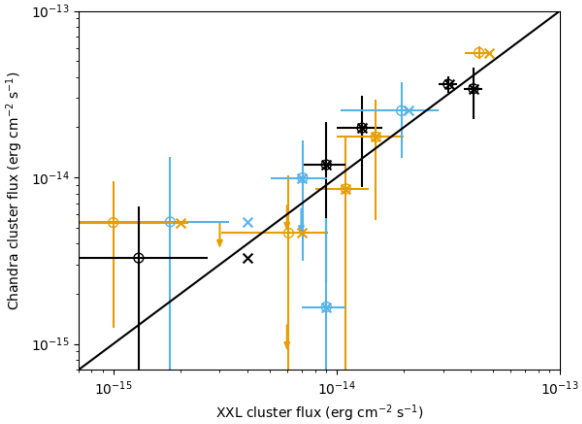

Figure 2.6: We show the *Chandra* cluster flux versus the XXL cluster flux, F_{60}. C1 clusters are black circles/crosses/arrows, C2s are yellow circles/crosses/arrows and C3s are blue circles/crosses/arrows. The crosses are F_{60} as listed in Table 2.2 column 4 (i.e. the original flux, not excluding the point sources detected by Chandra). The circles are the F_{60} minus the flux from any point source detected in the *Chandra* data that was not previously resolved by XXL data (listed in Table 2.2 column 6). The solid line is a line of equality. The arrows indicate upper limits on the *Chandra* cluster flux - the tip of the arrow denotes the upper limit and are plotted against the point source corrected XXL flux. 3XLSS J022059.0-043922 and 3XLSS J022554.3-045059 are not shown on the plot as the *Chandra* point source flux is greater than F_{60}.

2.6 Discussion

2.6.1 Cluster Contaminations

We report the point source detections, fluxes and cluster contaminations in Table 2.2. Individual point source flux measurements for each cluster can be found in Table 2.3. We plot the point source flux against the cluster flux to show the contamination levels in Figure 2.4.

Our results provide an important validation of the performance of the XXL cluster detection pipeline in classifying distant clusters. Four out of five of the C1 clusters are genuine uncontaminated clusters. Only the C1 3XLSS J021825.9-045947 is contaminated by AGN to a significant level (67% contamination, or 90% if we include the second undetected point source as discussed in Section 2.5). The C1 class is expected to be free from strongly contaminated clusters or misclassified AGN (though there is a chance to have a high confidence cluster that does have a significant fraction of AGN emission), but in this case the source was precisely at the threshold value in extension required for classification as a cluster. This is illustrated in Figure 2.5 which

shows the clusters and cluster candidates in the EXT - EXT_STAT parameter space. Furthermore, this cluster was detected 6′ off-axis in the *XMM* observation making extent measurements more challenging due to the increased asymmetry of the PSF. This appears to be a rare case of a false-positive C1 cluster at the classification threshold.

The C2 class shows a higher level of contamination than the C1 class, as expected - five clusters have no significant point source contamination (we include XLSSC 634 here, as, despite having five point sources detected in the 60″ cluster region, three of which were not detected by XXL, their contribution to F_{60} is very low) and the other four (3XLSS J022005.5-050826, XLSSC 203, 3XLSS J022755.7-043119 and 3XLSS J021320.3-053411) are either a blend of cluster and AGN emission or misclassified AGN. Our *Chandra* cluster flux measurement suggests that 3XLSS J022005.5-050826 is not a genuine cluster, as the 1σ upper limit for the cluster flux is low (see column 9 of Table 2.2). The results from our C2 clusters are consistent with the $< 50\%$ contamination that is expected in the C2 sample, and demonstrate that the XXL detection pipeline is capable of detecting extended sources even in the presence of relatively bright point sources.

Looking at the 14 C1 and C2 clusters together, nine have either no newly resolved point sources, or have new *Chandra*-detected sources that do not contribute significantly to the ICM flux (i.e. $> 15\%$). A further cluster, XLSSC 203, is more strongly contaminated (at the 50% level) but the *Chandra* measurement of the ICM flux from this system supports the conclusion that it comprises a blend of ICM and point source flux. The clusters form a useful sample that can be targeted for deeper follow up observations to probe ICM properties at $z > 1$ with good limits on the systematics from point source contamination. The legacy value of this should not be underestimated - there is no approved mission that will replace *Chandra*'s imaging capabilities (ATHENA will have a larger PSF, and Lynx, which will have a similar spatial resolution, is currently unapproved).

We can compare the updated pipeline (XXL Paper XX) directly to that used by Willis et al. (2013). If we define a "clean" cluster as having an AGN contamination fraction less than 0.15 for cases where `wavdetect` detects a point source within 60″ of the *XMM* X-ray peak, or a cluster that has no point sources detected by `wavdetect` in this region, we can see that the updated pipeline is more conservative. There is an improvement for the C2 class with the updated pipeline, giving us a more robust sample with 5/9 C2s clean, compared with 2/7 using the Willis et al. (2013) classes.

The 7 C3 candidates were optically selected and associated with XXL sources that do not meet the criteria for the C1 or C2 classes. As would be expected, this sample is less pure than the other classes, but two of the C3s are unambiguous high-z clusters, on the basis of low contamination fractions, supporting optical data and robust ICM detections in XXL and *Chandra* data. XLSSC 034 has a low level of contamination, and XLSSC 046 is a genuine cluster that was studied in detail by Bremer et al. (2006).

These C3 clusters do not have a well-defined selection function, but still present interesting targets for further study. Additional such clusters could be recovered by studying the optical/IR data for sources in the same EXT - EXT_STAT parameter space (see Section 2.3) as XLSSC 046. The location of XLSSC 034 and XLSSC 046 in the EXT - EXT_STAT parameter space is shown in Figure 2.5.

We note that the existence of clusters like the C3s that fail to meet the main survey selection criteria, and the presence of AGN contamination in the C1/C2 sample, does not represent a problem for the XXL selection function. The results of these snapshots validate the current modelling of the survey selection function, and provide useful additional input for its further refinement and testing by hydrodynamical simulations.

Galametz et al. (2009) used *Chandra* to identify X-ray AGN in optical/infrared-selected galaxy clusters. If we apply the same selection to the AGN detected in our *Chandra* observations, we would not detect any AGN in the inner 0.25 Mpc of our C1 and C2 clusters in the redshift range $1 < z < 1.5$. This is consistent with the results from Galametz et al. (2009), since based on their detection rate, we would expect ~ 1 AGN to be detected in the C1 and C2 cluster sample. Our results show that the effect of selecting clusters in the X-ray band does not strongly bias our sample towards clusters containing X-ray bright AGN. We note that the Galametz et al. (2009) study only included AGN that were cluster members, and our work includes all AGN (both cluster members and AGN projected onto the image of the cluster); additionally the clusters in Galametz et al. (2009) are likely of different masses to the clusters in our sample, and so the strength of the comparison made is limited.

A potentially important issue that has not yet been addressed is that of the variability of AGN. The *XMM* data used in the XMM-XXL survey were mostly taken years before the *Chandra* follow-up (this is true for at least the non-archival data that are the majority of our data). The typical variability in flux of AGN on this timescale is $\sim$50% (Maughan & Reiprich, 2019). Therefore, any cluster found to have a low (or undetectable) level of AGN contamination is

unlikely to have been $\gtrsim 30\%$ contaminated at the epoch of the XXL observation (or indeed at the epoch of any future, deeper observations). However, we note that any point source detected in the XXL observations is on average more likely than not to be at a higher than average luminosity, as selecting on X-ray flux will preferentially select sources at maximum flux that then fade.

2.6.2 ICM Fluxes

The cluster fluxes calculated from our *Chandra* data are shown in Table 2.2 and are compared with the XXL fluxes in Figure 2.6. For four of the clusters (XLSSC 072, XLSSC 029, XLSSC 634, XLSSC 034) the 3σ lower limit on the flux is greater than zero. The rest of the C1 and C2 clusters have 1σ lower limit greater than zero, except for 3XLSS J022755.7-043119, 3XLSS J022005.5-050826 and 3XLSS J022418.4-043956. These three clusters have upper limits that are consistent with the XXL flux (accounting for the unresolved AGN in the F_{60} measurement). In summary, after accounting for unresolved AGN in the XXL measurements and the measurement uncertainties, all of the cluster fluxes calculated from our *Chandra* data are consistent with those from XXL.

In some cases the *Chandra* cluster flux is non-zero, even when we believe there is only AGN emission and no cluster emission (3XLSS J021825.9-045947, 3XLSS J022059.0-043922, 3XLSS J021320.3-053411). In these cases, the *Chandra* ICM fluxes are not significantly different from zero and we interpret the signals as noise fluctuations rather than ICM detections.

2.7 Conclusion

We have analysed *Chandra* data for 21 clusters and cluster candidates that appear in the XMM-XXL survey catalogue in order to determine the extent of any contamination by unresolved point sources. Our main results are as follows:

- In the 14 C1 and C2 clusters which form a complete sample with a defined selection function, we find that the majority have little or no contamination of their ICM fluxes by AGN. One C1 source appears to be an AGN that was misclassified as extended, but this source was detected at the extension parameter threshold, so represents a rare interloper rather than any broad problem in the classification scheme. Three or four of the nine C2 clusters are either AGN that were misclassified as extended sources, or else have ICM emission that is strongly contaminated by AGN emission. Overall these results agree well with the calibration of the

XXL selection function and serve to validate its description of these distant cluster samples. We remind the reader that these conclusions were derived for distant clusters where the angular size of a cluster might be a similar size to the *XMM* PSF; therefore, our conclusions should not be extrapolated to the lower redshift XXL clusters.

- With this *Chandra* follow-up, we have defined a complete sample of ten $z > 1$ clusters (those marked CC in column 8 Table 2.2 and XLSSC 203) for further study. This comprises all secure C1 and C2 clusters that show evidence for X-ray emission originating from the ICM (in addition to any contaminating AGN if they are detected).
- Of the seven C3 optically selected cluster candidates with X-ray counterparts that did not meet the C1 or C2 selection criteria, we consider two (XLSSC 034 and XLSSC 046) to be genuine clusters with low levels of AGN contamination. A third, 3XLSS J0222351.3-041841 may also be a genuine cluster with low contamination, but this is unclear due to the cluster region being only partially covered by *Chandra*. The remaining four sources are either AGN or clusters with high levels of AGN contamination.
- We measured the ICM flux with *Chandra*, recording upper limits in three cases. For all clusters, the *Chandra* ICM flux was consistent with that measured by *XMM* once the *XMM* flux was corrected for unresolved point sources.
- The number of AGN per cluster for this X-ray selected sample was found to be lower, but consistent with, that of clusters selected in the optical/infra-red, indicating the X-ray selection with the XXL pipeline does not lead to a bias towards clusters with associated X-ray bright AGN.

We have demonstrated the utility of *Chandra* snapshots to test for AGN in or projected onto clusters detected in surveys with poorer resolution, for example the upcoming *eROSITA* survey, which has a HEW of $28''$ average over the entire field of view (Merloni et al., 2012; Pillepich et al., 2012). *Chandra* snapshots can be used to decontaminate *eROSITA* high-z candidate clusters using methods similar to those presented in this paper.

2.8 Acknowledgements

XXL is an international project based around an *XMM* Very Large Programme surveying two 25 deg^2 extragalactic fields at a depth of $\sim$ 6 x 10^{-15} erg cm^{-2} s^{-1} in the 0.5- 2 keV band

for point-like sources. The XXL website is http://irfu.cea.fr/xxl. Multi-band information and spectroscopic follow-up of the X-ray sources are obtained through a number of survey programmes, summarised at http://xxlmultiwave.pbworks.com. We thank Adam Mantz and Mauro Sereno for their useful comments on this work. FP and MERC acknowledge support by the German Aerospace Agency (DLR) with funds from the Ministry of Economy and Technology (BMWi) through grant 50 OR 1514. The Saclay group acknowledges long-term support from the Centre National d'Etudes Spatiales (CNES). This paper has also made use of observations obtained with the Canada-France-Hawaii Telescope (CFHT) which is operated by the National Research Council (NRC) of Canada, the Institut National des Sciences de l'Univers of the Centre National de la Recherche Scientifique(CNRS) of France, and the University of Hawaii. This paper has also made use of the NASA/IPAC Extragalactic Database (NED) which is operated by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration. This paper has also made use of TOPCAT (Taylor, 2005).

2.9 Appendix: Point Source Positions and Individual Fluxes

XXLID	Class XXL	z	RA (J2000)	Dec. (J2000)	Flux (10^{-14} erg s^{-1} cm^{-2})	Resolved by *XMM*	Separation from X-ray peak ($''$)
XLSSC 072*	C1	1.00	33.852	-3.726	<0.08	No	8
XLSSC 029*	C1	1.05	36.002	-4.225	$0.04^{+0.03}_{-0.02}$	No	52
			36.012	-4.229	$0.05^{+0.03}_{-0.02}$	No	23
XLSSC 005*	C1	1.06	-	-	-	-	-
3XLSS J021825.9-045947	C1	1.13	34.609	-4.996	$0.27^{+0.09}_{-0.08}$	No	5
XLSSC 122*	C1	1.99	-	-	-	-	-
XLSSC 048*	C2	1.01	-	-	-	-	-
XLSSC 073*	C2	1.03	33.749	-3.515	$0.08^{+0.08}_{-0.05}$	No	37
			33.737	-3.519	$0.37^{+0.14}_{-0.12}$	3XLSS J021456.8-033108	53
3XLSS J022755.7-043119	C2	1.05	36.972	-4.516	$0.05^{+0.05}_{-0.03}$	No	10
			36.984	-4.521	$0.11^{+0.06}_{-0.05}$	No	43
			36.994	-4.520	$0.61^{+0.14}_{-0.12}$	3XLSS J022758.7-043110	44
3XLSS J021320.3-053411	C2	1.08	33.331	-3.571	$0.12^{+0.06}_{-0.05}$	No	11
XLSSC 203	C2	1.08	34.429	-4.988	$0.10^{+0.07}_{-0.05}$	No	4
XLSSC 634	C2	1.08	355.692	-54.185	$0.20^{+0.06}_{-0.06}$	No	4
			355.704	-54.185	$0.22^{+0.06}_{-0.05}$	No	29
			355.683	-54.177	$1.30^{+0.14}_{-0.14}$	3XLSS J234244.2-541033	34
			355.687	-54.175	$0.09^{+0.05}_{-0.03}$	3XLSS J234244.2-541033	36
			355.712	-54.176	$0.04^{+0.02}_{-0.02}$	No	54
3XLSS J021325.0-042000	C2	1.20	-	-	-	-	-
3XLSS J022005.5-050826*	C2	1.65	35.022	-5.140	$0.15^{+0.05}_{-0.03}$	No	5
			35.021	-5.139	$0.01^{+0.02}_{-0.01}$	No	10
			35.030	-5.137	$0.07^{+0.05}_{-0.04}$	No	29
			35.014	-5.134	$0.04^{+0.03}_{-0.02}$	No	41
3XLSS J022418.4-043956*	C2	1.67	-	-	-	-	-
XLSSC 034	C3	1.04	35.372	-4.093	$1.00^{+0.32}_{-0.26}$	3XLSS J022129.1-040534	22
			35.375	-4.111	$0.15^{+0.16}_{-0.10}$	No	45
3XLSS J022059.0-043922*	C3	1.11	35.247	-4.656	$1.52^{+0.37}_{-0.32}$	No	5
XLSSC 046*	C3	1.22	-	-	-	-	-
3XLSS J022351.3-041841	C3	1.27	-	-	-	-	-
3XLSS J021700.4-034746*	C3	1.54	34.253	-3.795	$0.41^{+0.20}_{-0.15}$	No	8
			34.258	-3.784	$0.07^{+0.11}_{-0.06}$	No	50
3XLSS J022812.3-043836*	C3	1.67	37.051	-4.651	$0.10^{+0.04}_{-0.04}$	3XLSS J022812.2-043906	25
			37.045	-4.648	$0.22^{+0.15}_{-0.11}$	No	26
3XLSS J022554.3-045059*	C3	2.24	36.477	-4.851	$0.13^{+0.08}_{-0.06}$	No	5
			36.472	-4.846	$0.24^{+0.10}_{-0.08}$	No	20
			36.471	-4.837	$0.89^{+0.19}_{-0.16}$	3XLSS J022552.8-045013	50

Table 2.3: Summary of the fluxes for all point sources within $60''$ of the *XMM* X-ray peak. Column 6 is the individual point source flux as calculated from the *Chandra* data with 1σ errors. All fluxes are in the 0.5 - 2 keV energy band. XXLIDs marked with a $\star$ appear in Willis et al. (2013) and are therefore part of the XMM-LSS subset of clusters. Column 7 states whether the *Chandra* detected point source was previously resolved by XXL and thus excluded from the F_{60} measurements; for cases where the point source was resolved by XXL, its name as in XXL Paper XXVII is provided. In the case of XLSSC 634, two sources were blended into one by the *XMM* PSF, reported as one point source by XXL, and were masked from the F_{60} calculation.

3

Mass determination using X-ray data

In this chapter we derive the hydrostatic mass equation used in the estimation of X-ray cluster masses, and outline the two methods that we use in Chapter 4 to obtain mass estimates of the clusters in our sample using X-ray data.

3.1 Derivation of Hydrostatic mass equation

To estimate the cluster mass using X-ray data, it is necessary to make two key assumptions about the cluster. The first assumption is that the gas in the cluster is in hydrostatic equilibrium with the gravitational potential, such that the inward gravitational force is balanced by the outward thermal pressure of the gas - this includes the assumption that all outward pressure is thermal, and not from other sources. The second assumption is that the cluster is spherically symmetric (see Limousin et al., 2013 for a discussion of this assumption). Under the first assumption, Euler's equation for an ideal fluid can be applied. Euler's equation for an ideal fluid with a velocity **v**, density $\rho_{\rm gas}$ and pressure $P_{\rm gas}$ in a gravitational potential ϕ (see e.g. Suto et al., 2013) is

$$\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla)\mathbf{v} = -\frac{1}{\rho_{\rm gas}} \nabla P_{\rm gas} - \nabla \phi \tag{3.1}$$

By setting the gas velocity to zero and applying the second assumption - that the cluster is spherically symmetric, equation 3.1 becomes

$$\frac{1}{\rho_{\rm gas}}\frac{dP_{\rm gas}}{dr} = -\frac{d\phi}{dr} = -\frac{GM(r)}{r^2} \tag{3.2}$$

where *G* is the gravitational constant, and *M(r)* is the total mass of the cluster within radius r. $P_{\rm gas}$, the pressure of the gas, can be written as $P_{\rm gas} = \rho_{\rm gas} k T_{\rm gas}/\mu m_{\rm p}$ using the ideal gas law, where $m_{\rm p}$ is the atomic mass unit (a.m.u.), μ is the mean molecular weight in a.m.u. of the gas (we use a value of 0.611) and $T_{\rm gas}$ is the gas temperature.

Writing equation 3.2 as

$$M(r) = -\frac{r^2 k}{G\mu m_{\rm p}\rho_{\rm gas}(r)}\left(\rho_{\rm gas}(r)\frac{\partial T_{\rm gas}(r)}{\partial r} + T_{\rm gas}(r)\frac{\partial \rho_{\rm gas}(r)}{\partial r}\right) \tag{3.3}$$

we obtain a formula for the mass of a cluster within a radius r. Manipulating equation 3.3 to obtain

$$M(r) = -\frac{rkT_{\rm gas}(r)}{G\mu m_{\rm p}}\left(\frac{r}{\partial r}\frac{\partial T_{\rm gas}(r)}{T_{\rm gas}(r)} + \frac{r}{\partial r}\frac{\partial \rho_{\rm gas}(r)}{\rho_{\rm gas}(r)}\right) \tag{3.4}$$

and then using the fact that $d\ \ln x$ = dx/x (with x as $T_{\rm gas}(r)$, $\rho_{\rm gas}(r)$ and r) we can arrive at an alternate form of the equation that is often used in literature:

$$M(r) = -\frac{rkT_{\rm gas}(r)}{G\mu m_p}\left(\frac{\partial\ \ln\ T_{\rm gas}(r)}{\partial\ \ln\ r} + \frac{\partial\ \ln\ \rho_{\rm gas}(r)}{\partial\ \ln\ r}\right) \tag{3.5}$$

We can use this equation to give us the mass of the cluster within a certain radius, the total mass of the cluster, or we can use it to get a mass profile of the cluster. When defining the mass within a certain radius, it is customary to write $M_{\triangle}$, which is the mass of the cluster within $r_{\triangle}$, where $\triangle$ denotes the mean overdensity of the cluster within the radius $r_{\triangle}$ with respect to the critical density of the Universe (see equation 1.10) at that cluster's redshift. For all quantities given in this thesis, we use $\triangle$ = 500 unless stated otherwise.

In order to use equation 3.5 to obtain the cluster mass within a radius, r, one approach is to obtain the 3D gas density and temperature profiles $T_{\rm gas}(r)$ and $\rho_{\rm gas}(r)$. This can done by projecting 3D models of the gas density and temperature profiles and fitting them to the observed projected profiles i.e. the observed temperature profile and emission measure profile (obtained from the observed surface brightness profile of a cluster, described in §4.3.9); we present this method in detail in §3.2. One benefit of this approach is that computing the gradient of the profiles at r, which is necessary to find the total mass at r (see equation 3.5), is trivial, given

that the functional form of the 3D profiles is known. An alternative approach is to de-project the emission measure profile to obtain the 3D gas density profile which, alongside an assumed mass profile, can be used to obtain a 3D temperature profile (see equation 3.5), which in turn can be projected and fitted to the observed temperature profile to constrain the assumed mass profile. This method is presented in §3.3.

We note that throughout this section we refer to the 3D density and temperature profiles, which are the 'true' density and temperature profiles of the 3D cluster, as opposed to the profiles measured from the projection of the 3D cluster on to the 2D sky. The 3D profiles in a cluster do not depend on three dimensions (i.e. they do not depend on the x, y and z position in the cluster); due to the assumption of spherical symmetry of the cluster (stated above) this is not necessary, and they can be written as depending on just one dimension (see equations 3.8 and 3.11).

3.2 Forward fitting mass method

In this section we detail the forward fitting method which can be used to obtain an estimate of the mass of a cluster using the emission measure profile, and the observed temperature profile from an X-ray observation of a cluster. In §4.3.8 and 4.3.9 we detail how an emission measure profile and temperature profile can be obtained from an X-ray observation, as well as how the projected abundance profile is obtained.

Figure 3.1 shows a flowchart outlining the main steps in the forward fitting mass method (we note that it is a simplified version and leaves out the detail in §3.2.3), and is a useful reference for the following section. In short, the forward fitting method involves defining parametrised models for the 3D gas density and temperature profile of a cluster, and projecting these to be fit to the observed emission measure profile and observed temperature profile respectively.

We refer to the mass method that we detail in this section as the forward-fitting mass method, as it is termed as such in Ettori et al. (2013).

3.2.1 Temperature profile and projection

From an X-ray observation of a cluster, a projected temperature profile can be measured, however for equation 3.5 the 3D (i.e. de-projected) temperature profile is needed. One way to obtain the 3D temperature profile is by using a parametric model for it and then projecting this model

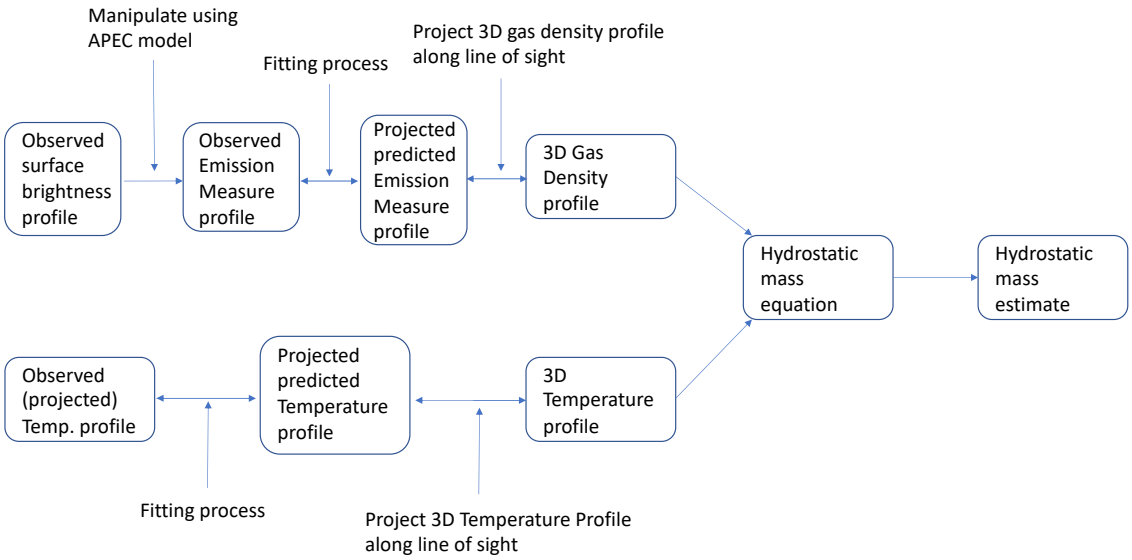

Figure 3.1: A flowchart showing the main steps in the forward fitting mass method, described in §3.2.

along the line of sight and subsequently performing a fit to the observed (and by definition projected) temperature profile. One model that has been used often before is presented in detail in Vikhlinin et al. (2006), and we summarise it here.

The model used to describe the 3D distribution of the temperature (Vikhlinin et al., 2006) is sufficient to model the temperature decline (if present) in the central core region of the cluster, in addition to the rest of the cluster temperature profile outside the central cooling region of a cluster. The part outside the central cooling region is modelled by

$$t(r) = -\frac{(r/r_{\rm t})^{-a}}{[1+(r/r_{\rm t})^{b}]^{c/b}} \tag{3.6}$$

where r is the radius and all other symbols are free parameters: a is the slope at small radii, c is the slope at large radii, and b is the width of the transition between these slopes. This transition occurs at a radius $r_{\rm t}$. The decrease in temperature found in the central cooling region of some clusters can be modelled as

$$t_{\rm cool}(r) = -\frac{(x+T_{\rm min}/T_0)}{x+1}, \qquad x = \left(\frac{r}{r_{\rm cool}}\right)^{a_{\rm cool}} \tag{3.7}$$

where r is the radius and all other symbols are free parameters: T_0 is the normalisation, $T_{\rm min}$ is

the minimum temperature, $a_{\rm cool}$ is the slope of the cool region which extends to a radius $r_{\rm cool}$, which is the radius of the central cooling region.

The 3D temperature profile model is a product of equations 3.6 and 3.7

$$T_{3D}(r) = T_0\, t_{\rm cool}(r) t(r) \tag{3.8}$$

This 3D model is then projected along the line of sight so that it can be fitted to the observed temperature profile. The projection of the 3D temperature model requires appropriate weighting of each temperature component along the line of sight. The method that we use for this projection is described in Vikhlinin (2006), and computes the temperature that would be measured when fitting a single temperature thermal plasma model of the cluster emission to the true cluster emission, that will in fact contain emission from components with different temperatures. A projection algorithm from Mazzotta et al. (2004) is also often used in the literature, but as their method just models the bremsstrahlung emission from the ICM, their projection algorithm only works above T $\gtrsim$ 3 keV. Vikhlinin (2006) improve on this algorithm by including line emission and as such the projection algorithm can be used at low temperatures too (above T $\gtrsim$ 0.5 keV).

We note that for the projection of the 3D temperature profile, the projection algorithm requires knowledge of the 3D abundance profile of the cluster; we do not use a 3D abundance profile but instead we use the observed (projected) abundance profile obtained from the cluster spectral fitting, which should be a close approximation to the 3D profile. In addition, the projection algorithm requires knowledge of the 3D ICM density profile, which is modelled simultaneously with the temperature profile.

3.2.2 Gas density profile and projection

The other quantity that needs to be determined from a cluster observation in order to be used in equation 3.5 is the 3D gas density profile. It is not possible to measure the 3D gas density profile directly, but it is possible to measure the surface brightness profile which is related to the gas density. Through manipulation of the surface brightness profile (described in §4.3.9) we can obtain the observed projected emission measure profile. We can then project the 3D gas density model along the line of sight such that it can be fitted to the emission measure profile.

One common gas density model is the β-model (Cavaliere & Fusco-Femiano, 1976):

$$\rho_{\rm g}(r) = \rho_{\rm g,0} \left[1 + \left(\frac{r}{r_{\rm c}} \right)^2 \right]^{-3\beta/2} \tag{3.9}$$

where r is the radius, $r_{\rm c}$ is the core radius, and $\rho_{\rm g}$ is the gas density. β is the ratio between the kinetic energy of the cluster galaxies and the thermal energy of the gas:

$$\beta = \frac{\sigma_{\rm v}^2}{kT_{\rm gas}/\mu m_{\rm p}} \tag{3.10}$$

where $\sigma_{\rm v}$ is the velocity dispersion of the member galaxies, $\mu_{\rm p}$ is the mean molecular weight per proton (we use 1.407) and μ is the mean molecular weight of the gas (we use 0.611). When the β-model was first presented, the model was sufficient to model the X-ray data that was obtained from cluster observations of the time. However, with the discovery of cool-core clusters (see §1.6.2), it was found that the gas density increased sharply in the centre of some clusters, so more complex models are now used to model gas density profiles from present-day cluster observations.

One example of a more complicated gas density model that is often used, and includes additional components compared to the standard β-model, is presented in detail in Vikhlinin et al. (2006). We summarise it here:

$$n_{\rm p}n_{\rm e}(r) = n_0^2 \frac{(r/r_{\rm c})^{-\alpha}}{(1+r^2/r_{\rm c}^2)^{3\beta-\alpha/2}} \frac{1}{(1+r^\gamma/r_{\rm s})^{\epsilon/\gamma}} + \frac{n_{02}^2}{(1+r^2/r_{\rm c2}^2)^{3\beta_2}} \tag{3.11}$$

with r being the cluster radius and the nine free parameters able to model both the central and outer parts of the density profile sufficiently. We fix $\gamma = 3$ as in Vikhlinin et al. (2006). $n_{\rm e}$ and $n_{\rm p}$ are the electron and proton density respectively, and r is the distance from the centre of the cluster. The first term in the equation is a modification of the β-model that allows for a power-law type cusp, with slope α at small radii, which is expected in the centres of clusters that are dynamically relaxed. The second term models the change of the slope by ϵ near the radius $r_{\rm s}$, and the width of this transition region is controlled by γ. The final term is another β-model component, which increases the freedom of the overall model near the cluster centre. We note that in this model, $n_{\rm p}n_{\rm e}$ is not the gas density as such, as is given in the equation for the β-model; however they can easily be converted from one form into the other using the values for $\mu_{\rm p}$ and $\mu_{\rm e}$ given at the very start of this chapter (using $\rho_g = (\mu_{\rm p}\mu_{\rm e})^{1/2} m_{\rm p} (n_{\rm p}n_{\rm e})^{1/2}$).

We can project the 3D gas density model profile along the line of sight to obtain what we would measure (at a certain radius of the emission integral profile) were a cluster with a given 3D gas density profile in projection on the sky. The setup that we use to do this is shown in Figure 3.2. Through the projection of the 3D gas density model along the line of sight, we can calculate the emission measure at x, EM

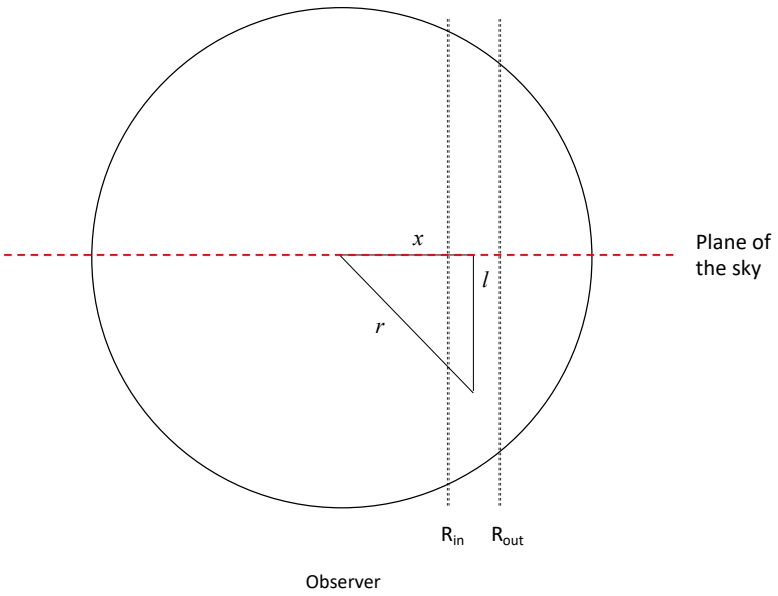

Figure 3.2: A sketch of a cluster from a top-down view. The circle represents the cluster. The observer is at a large distance from the cluster, and views the cluster face-on. The black dotted vertical lines show the annular bin. The red dashed horizontal line shows the plane of the sky. x, l and r are used in §3.2.

$$EM(x) = \int_{l=0}^{\infty} 2 n_{\rm p} n_{\rm e}(x) dl \tag{3.12}$$

where $n_{\rm p} n_{\rm e}(x)$ is the 3D gas density at the projected distance on the plane of the sky, x, l is the distance along the line of sight from the cluster mid-plane, and r is the 3D radius (shown in Figure 3.2). The factor of two is included due to the fact that we need to integrate from the cluster mid plane to both positive and negative infinity along the line of sight, and due to assumed spherical symmetry of the cluster, this is possible by evaluating from l = 0 to infinity and then doubling the integral. In practice, when evaluating this integral, instead of integrating to ∞, a maximum integration radius, r_{max} is used. r_{max} is chosen to be sufficiently large such that the cluster emission at that radius is negligible (10 Mpc is a reasonable value, and is what we use in Chapter 4).

By using Pythagoras' Theorem, we find

$$l = \sqrt{r^2 - x^2} \tag{3.13}$$

which we can differentiate to obtain

$$dl = \frac{r}{\sqrt{r^2 - x^2}} dr \tag{3.14}$$

which can be used to convert the integral in equation 3.12 to an integral along r:

$$EM(x) = \int_{r=x}^{\infty} \frac{2 n_{\rm p} n_{\rm e}(x) r}{\sqrt{r^2 - x^2}} dr \tag{3.15}$$

We now have the emission measure at a given projected radius, x, for an assumed 3D gas density profile $n_{\rm p} n_{\rm e}(x)$, which can be compared to the data point at that projected radius from the observed emission measure profile in the fitting process (see §3.2.6) to obtain the best fit 3D gas density profile for the cluster in question.

3.2.3 Self-consistent temperature and gas density profile fitting

The method to obtain the 3D temperature and gas density profiles shown in Figure 3.1, and described in §3.2.1 and 3.2.2 is not self-consistent, and is just used to show the main steps in the method in a simplified manner. It does not show a self-consistent method, as to calculate the correct conversion of the surface brightness profile to the emission measure profile requires a 3D temperature profile, and to project the 3D temperature profile requires knowing the 3D gas density profile. In many previous works (e.g. Vikhlinin et al., 2006; Giles et al., 2017), a non self-consistent method is used, where the 3D gas density model is first projected to be fitted to the emission measure profile, with the conversion between the surface brightness profile and emission measure obtained using an assumed 3D temperature profile, as the dependence on temperature is quite weak. Once the best fit 3D gas density profile is obtained, the projection of the 3D temperature model can be performed using this 3D gas density profile so it can subsequently be fitted to the observed projected temperature profile.

An improvement on this method is to fit the projected 3D gas density to the emission measure profile and the 3D temperature profile to the observed projected temperature profile *simultaneously*. This is what we call the self-consistent method. The temperature projection also requires knowledge of an abundance profile, and for this we use the projected abundance profile that we can measure as in §4.3. Ideally, the 3D abundance profile should be used, but the projected abundance profile is a good approximation to this, and an improvement on many previous works that assume a flat $Z = 0.3$ abundance profile throughout the cluster.

3.2.4 Gas mass of a cluster

In addition to calculating the hydrostatic mass using the 3D gas density and temperature profiles, we can also use the 3D gas density profile to obtain a gas mass, $m_{\rm gas}$, for the cluster in question using:

$$m_{\rm gas} = 4\pi m_{\rm p}\sqrt{\mu_{\rm e}\mu_{\rm p}}\int_0^r \sqrt{n_{\rm e}n_{\rm p}}r^2 dr \tag{3.16}$$

where $m_{\rm p}$ is the mass of a proton, and $\mu_{\rm e}$ and $\mu_{\rm p}$ are the mean molecular weights per electron and proton respectively, for which we use the values 1.173 and 1.407.

3.2.5 The fitting process / Markov Chain Monte Carlo

The fitting process uses Markov Chain Monte Carlo (MCMC), which requires knowledge of Bayes' rule, which I introduce briefly here. Bayes' rule, despite being termed as such, is less a rule than an outcome of the conditional probability equation

$$P(A,B) = P(A|B)P(B) = P(B,A) = P(B|A)P(A) \tag{3.17}$$

where the joint probability of events A and B occurring, $P(A,B)$, is equivalent to the product of event A occurring given that event B has already occurred, $P(A|B)$, with the probability of B occurring, $P(B)$. We can rearrange equation 3.17 to give Bayes' rule

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \tag{3.18}$$

If we want to use Bayes' rule for model fitting we could write it as

$$P(\theta|D) = \frac{P(D|\theta)P(\theta)}{P(D)} \tag{3.19}$$

where D is the data to be fitted and θ represents the model parameters. $P(\theta|D)$ is often referred to as the posterior probability distribution, $P(D|\theta)$ the likelihood, $P(\theta)$ the prior, and $P(D)$ the evidence. Often, for model fitting, we do not care about the evidence, which is a normalising constant for the right hand side, and can write

$$P(\theta|D) \propto P(D|\theta)P(\theta) \tag{3.20}$$

which is a more practical way of using Bayes' rule for model fitting as the evidence can often be very computationally expensive to compute.

By using Bayes' rule in any model fitting that we do, we can incorporate any prior knowledge that we have in the prior term, with the strength or 'tightness' of the prior representing the degree of belief we have in our prior knowledge. Model fitting requires obtaining the posterior probability distribution, which gives us the probability of the model parameters given the data. One ideal method for this model fitting process is MCMC.

MCMC algorithms use Markov chains to sample from a probability distribution. One MCMC algorithm is called the Metropolis Algorithm, and is a random walk that uses an acceptance/rejection criterion to converge on the high probability part of the parameter space (the target distribution). The algorithm works as follows, where θ represents the parameters in the model and x represents the data to which we are fitting the model:

- Choose initial values for θ, θ^0. This can be done randomly, or by choosing likely values for the parameter values.
- For step i = 1, 2, 3 etc. in the random walk:
  - Sample a proposed next step θ^{proposed} from a distribution centred on θ^{i-1}
  - Calculate the ratio r = $\frac{p(\theta^{\text{proposed}}|x)}{p(\theta^{i-1}|x}$)
  - Set $\theta^i = \theta^{\text{proposed}}$ with probability min(r, 1) ; θ^{i-1} otherwise.

Clearly, this algorithm is designed to converge on the target distribution, as a higher probability proposed step is always accepted, and a lower probability proposed step is only sometimes accepted, with the acceptance probability depending on the ratio of the probability of the proposed step to the probability of the current position of the walker. Once the Markov chain has converged on the target distribution, it is possible to obtain a sample of the target distribution by observing the chain after this convergence period (often referred to as the 'burn-in').

However, Metropolis MCMC is the simplest form of MCMC, and will often fail in high dimensional parameter spaces. In these cases, more complex MCMC algorithms are used, with optimizations that help them explore the parameter space more efficiently, in order to converge on the target distribution. Once converged on the target distribution, these algorithms will explore the target distribution effectively. Popular algorithms are the Metropolis-Hastings algorithm and Hamiltonian Monte Carlo.

Optimized MCMC algorithms are therefore particularly well suited to our particular fitting

process (i.e. fitting the projected 3D gas density profile to the observed emission measure profile, and the projected 3D temperature profile to the observed projected temperature profile), especially in the self-consistent fitting method (see §3.2.3) where both models are fit simultaneously, as the number of parameters to be constrained is large (17 in total, with eight from the 3D temperature model and nine from the 3D gas density model detailed in §3.2.1 and 3.2.2).

3.2.6 Using `emcee` for model fitting and estimating uncertainties

To fit our projected 3D density and temperature models to the observed temperature and observed emission measure profiles respectively, we use the MCMC tool `emcee` (Foreman-Mackey et al., 2013), which is a widely used Python implementation of the affine-invariant MCMC method proposed by Goodman & Weare (2010). An 'affine-invariant' algorithm performs equally well under all linear transformations, which means that the step in the MCMC algorithm where a new step is proposed can be computed much more quickly, leading to a shorter required burn-in. `emcee` also uses multiple 'walkers' to explore the parameter space, with one benefit being that the user is able to explore the parameter space more quickly (as `emcee` is parallelized), and another being that the proposal of the new position for each step is more stable.

Using `emcee` involves defining a likelihood function ($P(D|\theta)$ in equation 3.20) describing the probability of the current model parameters given the data. The likelihood function that we use is as follows:

$$L(x|\theta) = \sum \ln(P(T_{xi}|T_{\theta i})) + \sum \ln(P(EM_{xi}|EM_{\theta i})) \tag{3.21}$$

where $L(x|\theta)$ is the natural-log of the likelihood of the data x, given the parameters θ. It is a sum of the natural-log probabilities at each data point for the temperature (T) and emission measure (EM), of observing each data point, given the model parameters, θ. `emcee` aims to maximize the likelihood function while sampling from the probability distribution of our model parameters, and priors on these parameters may also be specified.

`emcee` provides us with chains which store information about the steps of each walker in the parameter space of the 3D temperature and density models. The probability distributions of any quantities that are derived from those models (such as hydrostatic mass or gas mass) are then straightforwardly computed by sampling random values from the chains; from these probability distributions, we can compute the mean or median value, and associated uncertainties.

Density profile (Equation 3.11)								
n_0	r_c	r_s	α	β	ϵ	n_{02}	r_{c2}	β_2
cm^{-3}	kpc	kpc				cm^{-3}	kpc	
$(10^{-4}, 10^{-1})$	(5, 800)	(100, 4000)	(0, 3)	(0.3, 1.5)	(0, 5)	$(0, 10^{-1})$	(1, 70)	(0.1, 5)
Temperature profile (Equation 3.8)								
T_0	r_{cool}	a_{cool}	T_{min}	r_t	a	b	c	
keV	kpc		keV	kpc				
(0.5, 18)	(10, 500)	(0, 3)	(0.1, 6 or T_0)	(100, 500)	(-0.5, 0.5)	(0, 5)	(0, 1)	

Table 3.1: Summary of priors for the 3D gas density and 3D temperature models used in the forward fitting method and described in§3.2.1 and 3.2.2. All are flat uniform priors with lower and upper bounds in parentheses.

When using `emcee` to fit the projected 3D gas density and temperature models to the data, flat priors are applied to all parameters (see Table 3.1). To initialize the starting points for our walkers in the parameter space, we use maximum likelihood estimation (MLE), to find an optimal starting point, and the walkers are initialized in a Gaussian ball around this optimal starting point. For the `emcee` runs, we used 10,000 steps and 100 walkers, and by visually inspecting the trace plots of all the walkers in the ensemble after the `emcee` run had completed, as well as the average model fits to the observed data, we could see if the `emcee` run had converged. In most cases, the walkers had converged, but for those that hadn't we ran `emcee` again, with the same number of walkers and steps, but this time starting from the parameter choice that gave the highest value for the likelihood in the previous run.

3.3 Backward fitting mass method

In this section we detail the backward fitting mass method. In short, the backward fitting method involves de-projecting the integrated emission measure profile to obtain the 3D gas density which is used in conjunction with an assumed mass profile in equation 3.5 to obtain a predicted 3D temperature profile which is then projected and fitted to the observed temperature profile to constrain the assumed mass profile. In Figure 3.3 we show a flowchart outlining the main steps in the backward fitting mass method, which is a useful reference for the following section.

We refer to the mass method that we detail in this section as the backward-fitting mass method, as it is termed as such in Ettori et al. (2013); however we note that the method used here is not exactly the same as that detailed in Ettori et al. (2013) - for example, instead of chi-squared fitting we use a MCMC method.

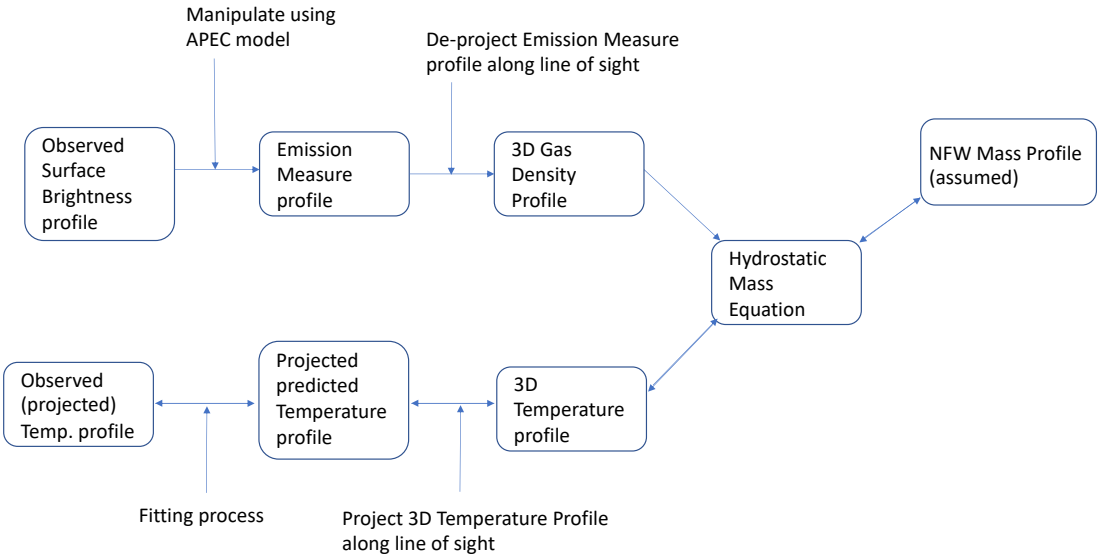

Figure 3.3: A flowchart showing the main steps in the backward fitting mass method, described in §3.3.

In contrast to the forward-fitting mass method, this method does not assume any functional form for the temperature or gas density model, so there is no shape imposed on the gas density and temperature profiles. This is an important difference, as in the forward-fitting method, the priors chosen for the parametrised gas density and temperature profiles can affect the outcome of the fitting; this is discussed further in Mantz & Allen (2011). However, in the backward fitting mass method, a functional form is used for the mass profile of the cluster. Obtaining mass estimates from the X-ray data using both mass methods is useful, as the methods are inherently different, and thus by comparing the masses obtained by both methods, we can be confident that our X-ray masses are accurate if our X-ray cluster masses agree well (we do this in §4.8.4).

3.3.1 Onion Peeling Technique

For the backward-fitting mass method, an analogy of the first step of the method, the de-projection of the integrated emission measure profile, is that we are peeling an onion: i.e. we are peeling away the layers of emission of the projection of the 3D cluster emission onto the 2D plane of the sky, which is how we see the cluster as the observer. We start by peeling the outermost layer of the integrated emission measure profile, and continue inwards, which is how

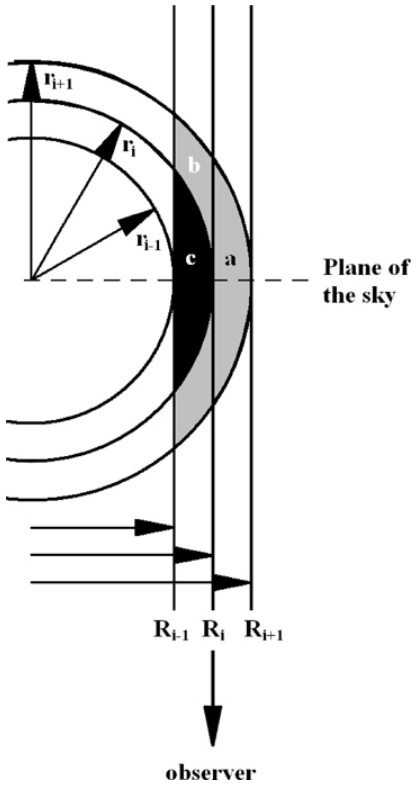

Figure 3.4: Sketch of a top-down view of a cluster, showing the onion-skin peeling setup for the geometric de-projection of the integrated emission measure profile (see §3.3). Taken from Ameglio et al. (2007).

we obtain our 3D gas density profile.

An illustration of the setup that we use to implement the onion-peeling method is shown in Figure 3.4. The cluster is divided into concentric spherical shells, which have the same radii as the annuli, or rings, of the cluster image in projection onto the 2D plane of the sky.

Referring to Figure 3.4, we start the onion peeling process by focusing on the outermost annulus and spherical shell. In this annulus we see that the only contribution to the outermost annulus in the cluster image is from the outermost spherical shell, and is defined by the region a. By calculating the volume of the region a that is bounded by the inner and outer radii of the spherical shell, $r_{\rm i}$ and r_{i+1} respectively, and the inner and outer radii of the annulus, $R_{\rm i}$ and $R_{\rm i+1}$ respectively, we can obtain the gas density in that spherical shell, using the fact that the emission measure integral EI, an observed quantity, is equal to $\int n_{\rm p} n_{\rm e} dV$ (i.e. we divide the emission

measure integral in that annulus by the volume of the spherical shell that intersects with that annulus).

We can then focus on the second outermost annulus, that has contributions to the integrated emission measure profile from region b and region c. The emission from region b comes from the same spherical shell as the emission from region a, i.e. it is within the same 3D radii, $r_{\rm i}$ and r_{i+1}, as region a, and so we already have the gas density in this spherical shell. From this we can find the contribution from region b by integrating the gas density in that spherical shell by the volume of region b, or alternatively scaling the emission measure integral contribution from a, as measured in the outermost annulus, by the ratio of the volume of region b divided by the volume of region a. We can now subtract the contribution to the integrated emission measure profile in the second outermost annulus from region b from the emission measure in that annulus which will give us the contribution from region c, and by dividing the remaining emission measure integral by the volume of region c we can find the gas density in the second outermost spherical shell. This process continues, as though we are peeling an onion, until the innermost annulus, providing us with a 3D gas density profile.

When we implement the onion peeling technique, we assume that there is no cluster emission outside the outer radius of the observed integrated emission measure profile. However, in reality, there will be some, and this 'extra' emission from outside the outermost annular bin (and outermost spherical shell) can be modelled as an extra contribution to the emission measure integral in each annular bin (see e.g. Ameglio et al., 2007).

3.3.2 Obtaining the Volume Matrix

The onion-peeling technique, as outlined in 3.3.1, can be implemented by using the following equation (see e.g. Ettori et al., 2002):

$$n_{\rm p} n_{\rm e} = 0.834 n_{\rm e}^2 = \left[(\mathbf{V}^T)^{-1} \# (EI) \right] \tag{3.22}$$

where $\mathbf{V}$ is the volume matrix (described below), EI is a vector containing information about the emission measure integral at different projected radii, # represents the matrix product, T superscript represents the transpose operation and -1 superscript represents the matrix inverse operation. With some simple manipulation and using $n_{\rm p}$ = 0.834 $n_{\rm e}$ (where $n_{\rm p}$ and $n_{\rm e}$ are the proton and electron density respectively), equation 3.22 becomes

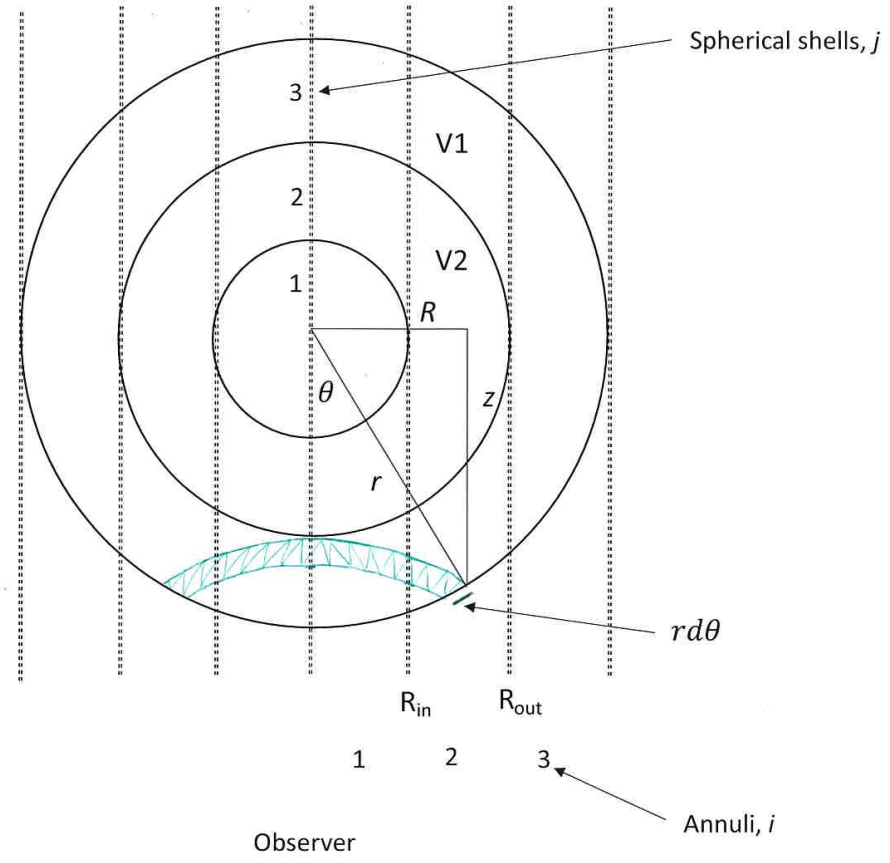

Figure 3.5: Sketch of the top-down view of a cluster to show the setup used to for the geometric de-projection of the integrated emission measure profile as detailed in §3.3 from equation 3.24 to 3.31. r, R, z and θ are defined in §3.3. The green shaded region represents dA, in equation 3.24. The spherical shells, j, and profile annuli, i, are also shown. V1 and V2 are equivalent to b and c respectively in Figure 3.4.

$$n_{\rm e} = \left[(\mathbf{V}^T)^{-1}\#(EI/0.834)\right]^{1/2} \tag{3.23}$$

To be able to use equation 3.23, we need to calculate the (upper triangular) volume matrix which contains the volume elements, $V_{\rm ij}$, for the intersection of the i^{th} annuli and j^{th} spherical shell for the cluster (see Figure 3.5). Obviously, for $j < i$, the intersection is 0. Otherwise it is non-zero.

We can see from Figure 3.5 that R is the projected radius on the sky, so also can be thought of as the annuli widths; r refers to the radii of the shells and is the 3D radius in the sky; z is the distance along the line of sight.

From Figure 3.5 we can see that the area, dA (shaded in green), of a strip on the 3rd spherical shell of radius r in the 2nd annulus would be $2\pi R r d\theta$ where r is the radius to the edge of the sphere and $rd\theta$ is the width of that small strip we are considering. We can use the following simple manipulation of

$$dA = 2\pi R r d\theta \tag{3.24}$$

by using the derivative of

$$R = r\sin\theta \tag{3.25}$$

which gives

$$d\theta = \frac{dR}{r\cos\theta} \tag{3.26}$$

and as

$$\cos\theta = \frac{z}{r} = \frac{(r^2 - R^2)^{0.5}}{r} \tag{3.27}$$

we can obtain

$$d\theta = (r^2 - R^2)^{-0.5} dR \tag{3.28}$$

We can then substitute equation 3.28 into equation 3.24 to give us

$$dA = \frac{2\pi R r}{(r^2 - R^2)^{0.5}} dR \tag{3.29}$$

which can be integrated to obtain

$$A_{ij} = \int_{R_{in}}^{R_{out}} \frac{2\pi R r}{(r^2 - R^2)^{0.5}} dR \tag{3.30}$$

Then, quite simply to get from the area of the j^{th} shell, A_{ij}, within the i^{th} annulus, to a volume element we integrate along *dr* for the inner radius of the shell to the outer radius of the shell

$$V_{ij} = \int_{r_{in}}^{r_{out}} \int_{R_{in}}^{R_{out}} \frac{2\pi R r}{(r^2 - R^2)^{0.5}} \, dR \, dr \tag{3.31}$$

which for the example that we were looking at as in Figure 3.5 corresponds to V1 or using the V_{ij} notation, *i*=2 and *j*=3. We remind the reader once more that equation 3.31 is only evaluated for $j \geq i$, otherwise the volume element is equal to zero.

Now we have created the volume matrix, we use it in equation 3.23 to obtain $n_{\rm e}$.

3.3.3 Constraining the mass profile

Once we have obtained $n_{\rm e}$, the hydrostatic mass equation described in §3.1 (equation 3.3) can be inverted to give

$$kT(r) = -\frac{1}{n_{\rm e}(r)} \left[G\mu m_{\rm p} \int_{r_{out}}^{r} \frac{n_{\rm e} M(<r)}{r^2} dr - P(r_{out}) \right] \tag{3.32}$$

where the symbols are described in §3.1, apart from P, which is the electron pressure. $r_{\rm out}$ is the outermost radius from which the deprojection is performed (the outer radius of the observed integrated emission measure profile in our case), and $P(r_{\rm out})$ is the electron pressure at that radius. The value of $P(r_{\rm out})$ is small compared to the value of the other term in the square brackets in equation 3.32 for most values of r, so its value doesn't have a large impact at most radii; the value of $P(r_{\rm out})$ only has a significant effect close to $r_{\rm out}$, which for all clusters is significantly above R_{500}, and also significantly above the outermost radii of the temperature profile (which is the strictest constraint on the reliability of the extent of the mass profile that we derive). An estimate of $P(r_{\rm out})$ can be obtained by using the temperature measurement in the outermost point of the temperature profile and the ideal gas law, $P = Nk_{\rm b}T = \rho_{\rm gas} kT_{\rm gas}/\mu m_{\rm p}$, where $k_{\rm b}$ is Boltzmann's constant, and N is the number of gas molecules. We obtain N from $n_{\rm e}$, and the $n_{\rm e}$ profile itself is smoothed in order to suppress noise when being used to obtain the estimate of $P_{\rm out}$.

From equation 3.32, we can see that if we assume a functional form for a mass profile (we choose to use the NFW profile, described below), and set the parameters for that mass profile, we will be able to obtain $T(r)$, the 3D temperature profile of the cluster. We can then project this 3D temperature profile (using the method in Vikhlinin, 2006, and described briefly in §3.2) to obtain the predicted observed projected temperature profile (given the mass profile in question and the value of $P_{\rm out}$).

The NFW profile (Navarro et al., 1997) is a well tested, and widely used model for mass profiles of galaxy clusters. It is parametrised by just two parameters, $r_{\rm s}$, a characteristic radius, and c, the concentration parameter. The NFW profile, $M(r)$, is defined as follows:

$$M(r) = 4\pi r_{\rm s}^3 \rho_{\rm s} f(x),$$

where

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

and

$$f(x) = \ln(1+x) - \frac{x}{1+x}$$

where $x = r/r_{\rm s}$ and $\rho_{\rm c}$ is the critical density, and Δ is the chosen overdensity, chosen such that $r_{\Delta} = r_{\rm s}\ c$. We choose to use $\Delta = 500$, so as to be able to easily obtain R_{500}.

As mentioned, we want to vary the parameters of the NFW model, to obtain different predicted temperature profiles, $T(r)$, which are then projected and fitted to the observed temperature profile of the cluster in question, in order to constrain the parameters of the NFW model. For this fitting process, we use `emcee` (described in §3.2.6), with the likelihood defined as $L(x|r_{\rm s}, c, P_{\rm out})$ where x is the data, more specifically, the surface brightness profile (or integrated emission measure profile) and the observed temperature profile. We apply (loose) flat priors for $r_{\rm s}$ and c, allowing $r_{\rm s}$ to vary between 10kpc and the outer radius of the outer temperature bin of our observed temperature profile, and allowing c to vary between 0.25 and 40. We also let $P_{\rm out}$ vary around the 'measured' value (which we describe above), up to five times above and five times below this value, using a flat prior to implement this.

We note that for the backward fitting method we convert the surface brightness profile into an integrated emission measure profile using the observed temperature profile. Ideally it should

be the 3D temperature profile that we used, however the observed temperature profile will be a good approximation to this, and in any case, the conversion is not sensitive to temperature. Although this is not completely self-consistent, in contrast to our approach in the forward-fitting method, we cannot achieve self-consistency as the backward-fitting method uses the deprojected emission measure integral as input. We use the gas density profile obtained in the de-projection of the integrated emission measure profile to project the predicted 3D temperature. Finally, as in the forward fitting mass method, we use the projected abundance profile that we can measure as in §4.3 for the projection of the predicted 3D temperature profile.

3.3.4 Gas Mass

We can also obtain the gas mass using this backward fitting method, using equation 3.16. $n_{\rm e}$ can easily be converted into $n_{\rm p}$ $n_{\rm e}$ using $n_{\rm p}$ = 0.834 $n_{\rm e}$ to subsequently be used in equation 3.16.

4

X-ray/Caustic Mass Comparison

4.1 Introduction

Due to their gargantuan size, galaxy clusters are tracers of large scale structure, and thus can be used as cosmological probes. The exponential drop in the number density of clusters at high masses means that the mass function of clusters can, in principle, be used to place precise constraints on cosmological parameters. However, the accuracy of those constraints depends crucially on the accuracy with which cluster masses can be measured, due to the exponentially decreasing cluster mass function. In fact, an error of $\sim 10\%$ on the cluster mass can lead to a factor of two difference in the space density [1]. The measurement of galaxy cluster masses is challenging, as most of the cluster mass is in the form of dark matter. Mass estimation techniques thus probe the masses of clusters indirectly via the effect of the gravitational potential of the cluster on its intracluster medium (ICM), its member galaxies, or images of background galaxies.

The most widely used and successful methods for estimating the total mass of galaxy clusters can be broadly split into three categories: hydrostatic masses based on X-ray observations of the ICM, weak gravitational lensing (WL) and galaxy dynamics. The X-ray hydrostatic mass estimation method assumes hydrostatic equilibrium, with the thermal pressure of the ICM preventing its gravitational collapse. WL masses use large samples of background galaxies that are

[1] for this approximation we used `hmfcalc` (Murray et al., 2013)

lensed by the cluster in question, and by finding the average distortion of the background galaxies, the cluster mass can be inferred. Methods using galaxy dynamics include caustic masses, which use the line of sight velocities of a large sample of member galaxies of a given cluster as a function of their projected distance from the centre of the cluster to trace the escape velocity of the cluster, from which a cluster mass can be calculated.

Optical data were used to give the first mass estimates of galaxy clusters. Zwicky (1933) measured the radial velocities of a number of cluster galaxies in the nearby Coma cluster, and by computing the radial velocity dispersion, used the virial theorem to estimate a mass for the cluster. In the late 1960s and 1970s, the first X-ray satellites were launched, and it was in the 1990s that the first samples of hundreds of clusters were produced (Ebeling et al., 1998, 2001; Böhringer et al., 2000, 2001), from which X-ray hydrostatic masses could be calculated (e.g. Reiprich & Böhringer, 2002). The launch of *XMM-Newton* and *Chandra* in 1999 has led to even more precise X-ray masses being measured for large samples of clusters (e.g. Martino et al., 2014). The first work using WL to map the dark matter distribution, undertaken by Tyson et al. (1990), was followed by a number of other papers (e.g. Blandford et al., 1991; Miralda-Escude, 1991; Kaiser, 1992), and in the last 3 decades it has become increasingly common to use WL as a mass estimation technique (see Bartelmann & Maturi, 2017 for a review).

Historically, X-ray hydrostatic masses have given the most precise masses, and hence have been used for calibrations of cluster masses for cosmology. Currently, there is significant tension between the cosmological constraints from the *Planck* cluster number counts (Planck Collaboration et al., 2016b) and the *Planck* cosmic microwave background (CMB) results (Planck Collaboration et al., 2016a). This could be due to a negative bias in the X-ray hydrostatic mass method (often referred to as the hydrostatic bias). The cluster masses from *Planck* were derived from a Sunyaev-Zeldovich mass proxy calibrated with X-ray hydrostatic masses, and thus the tension between the *Planck* results could be resolved if this bias is large ($\sim 40\%$). This bias has been explored significantly in recent years (e.g. Nagai et al., 2007; Lau et al., 2009; Nelson et al., 2014; Eckert et al., 2019), however no real agreement has been reached regarding its magnitude. However, recent work (Douspis et al., 2019) has shown that this tension could be reduced by using a lower value of the reionisation optical depth in the *Planck* analysis, compatible with lower values of the hydrostatic bias.

The hydrostatic bias arises from the assumption in the hydrostatic mass method that the only outwards pressure in the cluster is the thermal pressure of the ICM. However, if any other sources

of pressure (e.g. turbulence, bulk motions of the ICM or cosmic rays) are acting in addition to this thermal pressure measured with X-ray observations, this leads to the X-ray hydrostatic mass method underestimating the true cluster mass (Nagai et al., 2007; Lau et al., 2009; Nelson et al., 2014).

The level of hydrostatic bias is expected to be higher in unrelaxed clusters compared with relaxed clusters, as unrelaxed systems are expected to have more non-thermal pressure support from bulk motions of the cluster gas. However, all clusters, regardless of dynamical state, will experience some amount of non-thermal pressure support, as gas and substructures are always infalling onto the cluster as they undergo constant growth, leading to non-thermal pressure due to the residual motion of the ICM and the turbulence created. Constraining turbulence in the ICM requires high resolution X-ray spectrometry, and the only mission thus far capable of this has been *Hitomi*. Before the mission failed, *Hitomi* observed the Perseus cluster, finding that in the core of the Perseus cluster the pressure support from turbulence was 4% of the thermodynamic pressure (Hitomi Collaboration et al., 2016), which on its own would not lead to a large value of hydrostatic bias (though we note that the Perseus cluster is a dynamically relaxed cluster and, of course, not representative of the cluster population as a whole).

The hydrostatic bias may also be a function of radius, as the infalling gas may lead to stronger gas motions and therefore significant non-thermal pressure support at the outskirts of clusters. Bonamente et al. (2013) and Fusco-Femiano & Lapi (2018) both find evidence for significant non-thermal pressure at the cluster outskirts (for A1835 and A2142 respectively), and work using hydrodynamical simulations suggests that the non-thermal pressure due to the bulk gas motion increases with radius (Nagai et al., 2007; Lau et al., 2009; Vazza et al., 2009, 2018; Battaglia et al., 2012; Rasia et al., 2012). Specifically, Nagai et al. (2007) and Rasia et al. (2012) both find that at low ($< 0.2\ R_{500}$) and high radius ($> R_{500}$) the bias increases. Nagai et al. (2007) find the bias to increase from $<10\%$ at R_{500} to $\sim 30\%$ at twice this radius. Studies that have used observational data agree with this general trend (Siegel et al., 2018; Eckert et al., 2019).

Hydrostatic bias may also be a function of mass, with larger values of this bias potentially expected in lower mass systems, as these systems have a smaller gravitational potential, so the effect of any non-thermal pressure sources present will be greater than in higher mass systems. However, some work has found the opposite to be true: using simulations, Henson et al. (2017) find that the hydrostatic bias increases with mass (though these results depends on how cluster temperatures are calculated). Henson et al. (2017) also find that the hydrostatic bias shows a

strong dependence on the sphericity of the cluster, with a larger negative bias for less spherical clusters: this is likely due to sphericity being correlated with the dynamical state of the cluster.

To get a grasp on the biases of the cluster mass estimation methods, comparisons between methods themselves are undertaken, and simulated data is also used to compare the known mass from the simulation to the mass that is recovered if observational methods are applied to synthetic observations of the simulated cluster. Comparisons of the X-ray hydrostatic masses with WL masses can yield insights into the magnitude of the hydrostatic bias, as WL masses are not sensitive to the state of the ICM (in contrast to X-ray hydrostatic masses). However, WL masses are sensitive to mass along the line of sight to the observed cluster, which can lead to scatter of $\sim$ 20 - 30% and positive bias of up to $\sim$ 20% (Hoekstra, 2001, 2003; Hoekstra et al., 2011; Becker & Kravtsov, 2011), which needs to be understood and taken into account in these comparisons. Some comparisons between WL masses and X-ray hydrostatic masses suggest that X-ray hydrostatic masses underestimate the true mass by $\sim$ 20 - 30% (von der Linden et al., 2014; Donahue et al., 2014; Sereno et al., 2015; Hoekstra et al., 2015). Work using hydrodynamical simulations also suggest a significant hydrostatic bias of $\sim$10 - 30% (Rasia et al., 2006, 2012; Nagai et al., 2007; Lau et al., 2009; Nelson et al., 2014). However Gruen et al. (2014), Israel et al. (2014), Applegate et al. (2016) and Smith et al. (2016) find no significant evidence for hydrostatic bias when comparing X-ray hydrostatic masses to WL masses.

One factor that may contribute to the different estimates of the hydrostatic bias is that there exist uncertainties in the absolute calibration of *Chandra* and *XMM-Newton*, meaning these instruments are giving biased temperature measurements, resulting in biased X-ray hydrostatic masses (e.g. Mahdavi et al., 2013; Rozo et al., 2014; Schellenberger et al., 2015). However, Martino et al. (2014) find that, for a sample of 50 clusters with both *XMM-Newton* and *Chandra* data, their X-ray hydrostatic masses calculated with the data from each observatory agree well.

Comparisons between X-ray and WL masses from observations also support the idea that disturbed clusters will have a larger hydrostatic bias than relaxed clusters (Zhang et al., 2010; Mahdavi et al., 2013). Biffi et al. (2016) also find a difference in hydrostatic bias between relaxed and disturbed clusters using simulations (though find no difference between cool core and non cool core systems).

Comparisons between X-ray and caustic masses are much rarer. Caustic masses are an ideal mass estimator with which to compare X-ray hydrostatic masses, as they are not sensitive to the state of the ICM (similarly to WL masses), and any biases in the caustic method are well

understood from simulations (Serra et al., 2011). One drawback of using caustic masses for comparison is that they have a large scatter of $\sim 30\%$ (Serra et al., 2011), which comes predominantly from projected nearby structure affecting the projected distance-line of sight velocity diagrams used for the caustic mass method and from the assumption that the cluster is spherical. Thus, for a meaningful comparison, a large cluster sample is needed (several tens of clusters). Maughan et al. (2016), Andreon et al. (2017) and Foëx et al. (2017) use observational data to infer caustic masses to be $\sim$20% lower, $\sim$15% larger and $\sim$30% larger than the hydrostatic masses respectively. Armitage et al. (2019) use simulation data to compare caustic and X-ray hydrostatic masses and find a similar mass ratio to Foëx et al. (2017).

Simulations suggest caustic masses should be biased high by 10 - 20% (Serra et al., 2011), with the exact value depending on radius, which when combined with the predicted hydrostatic bias should lead to a ratio of hydrostatic to caustic mass of $\sim$ 0.6 - 0.8.

In this chapter, we use a sample of galaxy clusters from the Hectospec Cluster Survey (Rines et al., 2013). By comparing the X-ray hydrostatic and caustic mass profiles for the clusters in our sample, we investigate the ratio of these two mass methods over each cluster's radial range and as a function of dynamical state. Our work expands on the work done in Maughan et al. (2016) by using a larger sample of 44 clusters compared to the 16 used in Maughan et al. (2016), and using an improved mass estimation method.

4.2 Cluster Sample

We have performed *Chandra* follow-up of all of the clusters in the flux-limited Hectospec Cluster Survey (HeCS; Rines et al., 2013), a spectroscopic survey of X-ray selected clusters, for which there were no pre-existing *Chandra* observations. The HeCS sample was constructed by matching clusters selected using the *ROSAT* All-Sky Survey (RASS; Voges et al., 1999) with the imaging footprint of the SDSS Data Release 6 (Adelman-McCarthy et al., 2008). The SDSS data were used to select candidate cluster member galaxies for spectroscopic follow-up, which was performed with MMT/Hectospec (Fabricant et al., 2005).

The flux limit applied to the HeCS sample in order to create the flux-limited HeCS sample was 5×10^{-12} erg s^{-1} cm^{-2}, and excluded A750, A2187, A2396, A2631, and A2645 from the original HeCS sample. We then exclude three further clusters (A689 Giles et al., 2012, A1366,

Cluster	RA deg	DEC deg	z	ObsID	clean time (ks)	Ngal
Zw1478	119.9190	53.9990	0.1027	18248	19.7	82
A655	126.3610	47.1320	0.1271	15159	8 .0	315
A697	130.7362	36.3625	0.2812	4217	19.5	185
MS0906	137.2832	10.9925	0.1767	924	29.8	101
A773	139.4624	51.7248	0.2173	5006,533,3588	15.0,10.0,9.1	173
A795	141.0240	14.1680	0.1374	11734⋆	29.4	179
Zw2701	148.1980	51.8910	0.2160	12903⋆,3195⋆	93.2,22.2	93
A963	154.2600	39.0484	0.2041	903⋆	34.2	211
A980	155.6275	50.1017	0.1555	15105	13.9	222
Zw3146	155.9117	04.1865	0.2894	909,9371	42.5,32.7	106
A990	155.9120	49.1450	0.1416	15114	9.9	91
Zw3179	156.4840	12.6910	0.1422	13375	8.9	69
A1033	157.9320	35.0580	0.1220	15614,15084	32.6,29.4	191
A1068	160.1870	39.9510	0.1386	1652⋆	25.8	129
A1132	164.6160	56.7820	0.1351	19770,13376	19.7,8.9	160
A1201	168.2287	13.4448	0.1671	9616	47.4	165
A1204	168.3324	17.5937	0.1706	2205	23.6	92
A1235	170.8040	19.6160	0.1030	18247	18.1	131
A1246	170.9912	21.4903	0.1921	11770	5.0	226
A1302	173.3070	66.3990	0.1152	18245	18.8	162
A1413	178.8260	23.4080	0.1412	5003	75.0	116
A1423	179.3420	33.6320	0.2142	11724,538	25.7,9.9	230
A1437	180.1040	03.3490	0.1333	15306,15188	9.9,9.4	194
A1553	187.6959	10.5606	0.1668	12254	8.6	171
A1682	196.7278	46.5560	0.2272	11725	19.9	151
A1689	197.8750	- 1.3353	0.1842	6930,7289,5004	76.1,75.1,19.9	210
A1763	203.8257	40.9970	0.2312	3591	19.6	237
A1835	210.2595	02.8801	0.2506	6880,6881,7370	117.9,36.0,39.5	219
A1918	216.3420	63.1830	0.1388	18249	20.5	80
A1914	216.5068	37.8271	0.1660	20026,18252,20023,20025,20024	28.9,26.7,25.4,21.7,17.2	255
A1930	218.1200	31.6330	0.1308	11733⋆	34.5	76
A1978	222.7750	14.6110	0.1459	18250	19.8	63
A2009	225.0850	21.3620	0.1522	10438	19.9	195
RXJ1504	226.0321	- 2.8050	0.2168	17670,17197,17669,4935	50.7,29.8,28.7,11.8	120
A2034	227.5450	33.5060	0.1132	12886,12885	91.3,81.2	182
A2050	229.0680	00.0890	0.1191	18251	14.9	106
A2069	231.0410	29.9210	0.1139	4965	38.8	441
A2111	234.9337	34.4156	0.2291	11726, 544	20.9,10.3	208
A2219	250.0892	46.7058	0.2257	14356,14431,14355,14451	49.4,38.6,29.7,19.8	461
Zw8197	259.5480	56.6710	0.1132	18246	18.8	76
A2259	260.0370	27.6702	0.1605	3245	10.0	165
RXJ1720	260.0370	26.6350	0.1604	4361,1453	13.8,7.8	376
A2261	260.6129	32.1338	0.2242	5007	24.3	209
RXJ2129	322.4186	00.0973	0.2339	9370,552	29.6,10.0	325

Table 4.1: Summary of sample. Column 1 gives the cluster name, column 2 and 3 give the RA and DEC of the cluster respectively, column 4 gives the spectroscopic redshift of the cluster, column 5 gives the *Chandra* ObsID, column 6 gives the length of the observation after lightcurve cleaning, column 7 gives the number of galaxies classed as member galaxies of each cluster in the caustic method, as defined in Rines et al. (2013) (equivalent to the number of galaxies within the caustics). Starred ObsIDs are ACIS-S otherwise all are ACIS-I

and A2055) from the flux-limited sample as they are clearly dominated by AGN upon inspection of the X-ray observations (*Chandra* for A689 and A1366, and *XMM-Newton* for A2055). More specifically, A1366 and A2055 have BLLacs in the brightest cluster galaxy (BCG), and A689 has a BLLac very close to the BCG. We also excluded four clusters due to flaring in the available observations (A267, A667, A1361, and A1902), one cluster as it was part of a double cluster (A1758), and one cluster as it was off-chip in the *Chandra* observation available (A646 - the observation was of a radio galaxy in the cluster, but around $7'$ from the cluster core). All clusters are in the redshift range $0.1 < z < 0.3$. We call our sample the CHeCS (*Chandra* observations of Hectospec Cluster Survey clusters) sample. The final sample consists of 44 clusters, and a pilot study of 16 clusters was presented in Maughan et al. (2016). Our sample is summarised in Table 4.1.

4.3 Initial X-ray Data Analysis

In this section, I detail the analysis process used for our sample of clusters. All of the clusters in our sample were analysed with the CIAO[2] 4.10 software package and CALDB[3] version 4.8.1 (Fruscione et al., 2006).

4.3.1 Reprocessing

The `chandra_repro` tool was used to reprocess the level 1 events files (the unprocessed, raw *Chandra* data) to produce level 2 events files which are to be used for the rest of the data analysis (after the lightcurve cleaning in §4.3.2), following the standard data reduction threads[4] [5]. The events files are named as such as they contain an 'event' for each time charge is deposited and subsequently recorded by the CCD. Each event has the following information: the energy of

[2]See `http://cxc.harvard.edu/ciao`

[3]See `http://cxc.harvard.edu/caldb`

[4]See `http://cxc.harvard.edu/ciao/threads/index.html`

[5]We note the following: ACIS observations can be taken in FAINT or VFAINT mode, and the cleaning of the level 1 events file is dependent on this mode; in FAINT mode a 3×3 pixel event island is recorded which is centred around the pixel where the event occurs, and in VFAINT mode a 5×5 pixel island is recorded. In VFAINT mode, the charge in the 16 edge pixels of the event island can be used to differentiate between good and bad events; this is not the case for the FAINT mode, where the 8 edge pixels are too close the event pixel itself. As the blank sky backgrounds that we use in our analysis (see §4.3.3) have already been reprocessed, and the blank sky backgrounds were taken only in FAINT mode if the observation was taken before 01/12/2000 (and processed as such), for all observations prior to that date, we apply FAINT mode cleaning to our events files, so as to be consistent with the cleaning of the blank sky backgrounds, even if the observation itself was taken in VFAINT mode. The existence of different observing modes is necessary due to telemetry restrictions.

the detected photon (assuming it was a photon that deposited the charge, though it could also be from other sources, such as soft protons), the time at which the photon was detected, the position on the CCD where the photon was detected, and a grade of the photon which describes the quality of the event (dependent on factors such as how much charge leaked into neighbouring pixels).

4.3.2 Lightcurve Cleaning

X-ray observations are often affected by periods of backround flaring, due to an increase in the solar wind, which releases high energy particles (soft protons), which can deposit charge directly in the CCDs, or interact with the telescope itself, releasing photons which are subsequently detected by the CCD. To identify these periods of background flaring, we extracted lightcurves in the 0.3 - 12 keV band for front illuminated chips, in the 2.5 - 6 keV range for the S1 back illuminated chip, and in the 2.5 - 7 keV range for the S3 back illuminated chip[6] (see §1.9). We then identified periods of background flares in the lightcurves as either having count rates with a $> 20\%$ deviation from the mean rate, or a 3 σ deviation from the mean rate; any periods fulfilling either criterion were removed (this was done using the `deflare` tool). In addition, the cleaned lightcurves were visually inspected to identify any periods where there might be residual flaring (see §4.3.11 for details) that was subtle enough not to be picked up by the automated cleaning step.

4.3.3 Background Subtraction

The X-ray background consists of multiple components. One component is termed the cosmic X-ray background (CXB), which originates predominantly from unresolved extragalactic sources. Another component is the soft Galactic foreground, which is relevant at lower X-ray energies ($\sim$1 keV), and is believed to be predominantly from bubbles of warm gas in our galaxy that are left over from supernova explosions. There is also a contribution from the instrument itself, originating from the interaction of high energy particles with the CCD (directly depositing charge) and with the telescope itself (producing, for example, fluorescent X-rays which then strike the CCD).

In many of our observations most of the CCDs' area is filled by cluster emission, and so cannot be used as a representative sample of the background. For this reason, we make use of

[6] see `http://cxc.harvard.edu/contrib/maxim/acisbg/COOKBOOK`

blank sky background files, which are provided in the CALDB. These blank sky background files consist of combined observations of *Chandra* pointings at areas of the sky above the Galactic plane with no large extended sources in the field of view, and with all other sources masked out. For each observation, the blank sky background files were adjusted so as to match the parameters of the cluster observation in question. Specifically, the blank sky background files needed to contain the same CCDs as the cluster observation, and were normalized such that they matched the exposure time of the cluster observation. These blank sky background files, when subtracted from the observations' events files, account for a large part of the CXB.

We also account for the background from the interaction of the CCD and the telescope with high energy particles (sometimes referred to as the Quiescent Particle Background). The *Chandra* effective area in the 9.5 - 12 keV band is virtually zero, and so we can use the entire area of all CCDs to renormalise the emission in this area. We compare the emission (in the 9.5 - 12 keV band) from the observation in question with that from the blank sky background file. We then renormalise the blank sky background in this energy band such that it matches the observed count rate from the observation.

For any analysis that uses image data, which includes the creation of the surface brightness profiles and the defining of the binning of the temperature profiles (see §4.3.8 and 4.3.9), the blank sky background files are renormalised by matching the count rate in the imaging energy band (0.7 - 2 keV, see §4.3.4) from the blank sky background to that from a cluster free region (see §4.3.5 for detail on how this is found) for the observation in question. The blank sky background is then subtracted from the source profile data.

For the spectral analysis, the observation in question and the blank sky background files contain different contributions from the soft Galactic foreground (especially true at energies below 1 keV), as the soft Galactic foreground has a spatial variation. To account for this, we extract spectra in a cluster free region, and after subtracting the background spectrum from this, we fit the residuals (as in Vikhlinin et al., 2006) in the 0.3 - 3 keV band with an APEC model (Smith et al., 2001 details the APEC model), using ATOMDB version 3.0.9 and relative abundances fixed to the solar ratios of Asplund et al. (2009). The absorbing column was set to the NHTOT value, which is the Galactic value plus a molecular hydrogen column density component (Willingale et al., 2013), and we used Sherpa (Freeman et al., 2001) for our cluster background spectral fitting, using the *C*-statistic, and grouping our spectra to have at least 5 counts per bin. In most cases, the residual spectra are fitted well by a 0.18 keV temperature

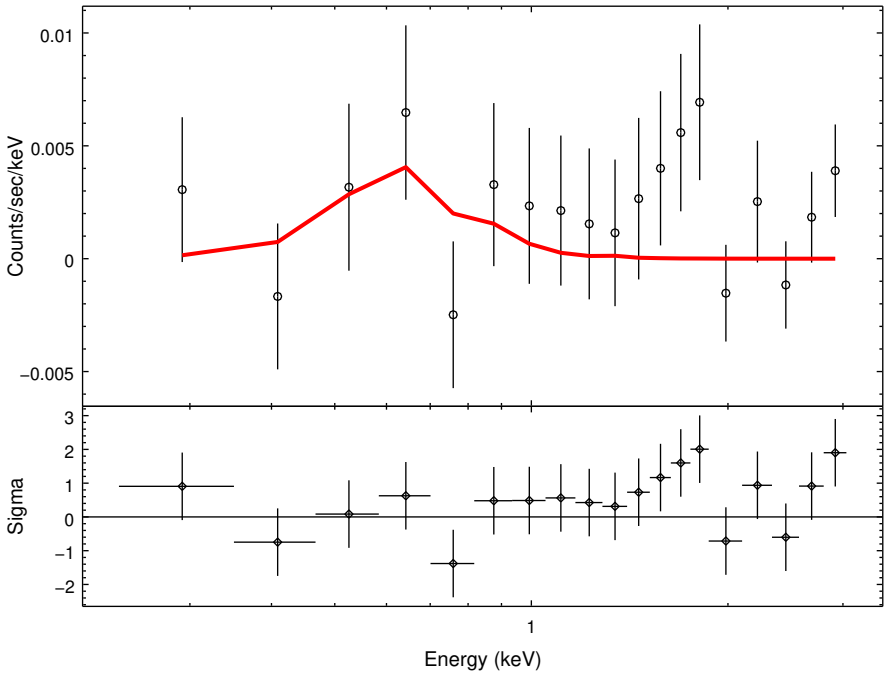

Figure 4.1: This plot shows the spectral model with a temperature of 0.18 keV (red) that fits well to the residuals between the blank sky background and the observation's particular background (black).

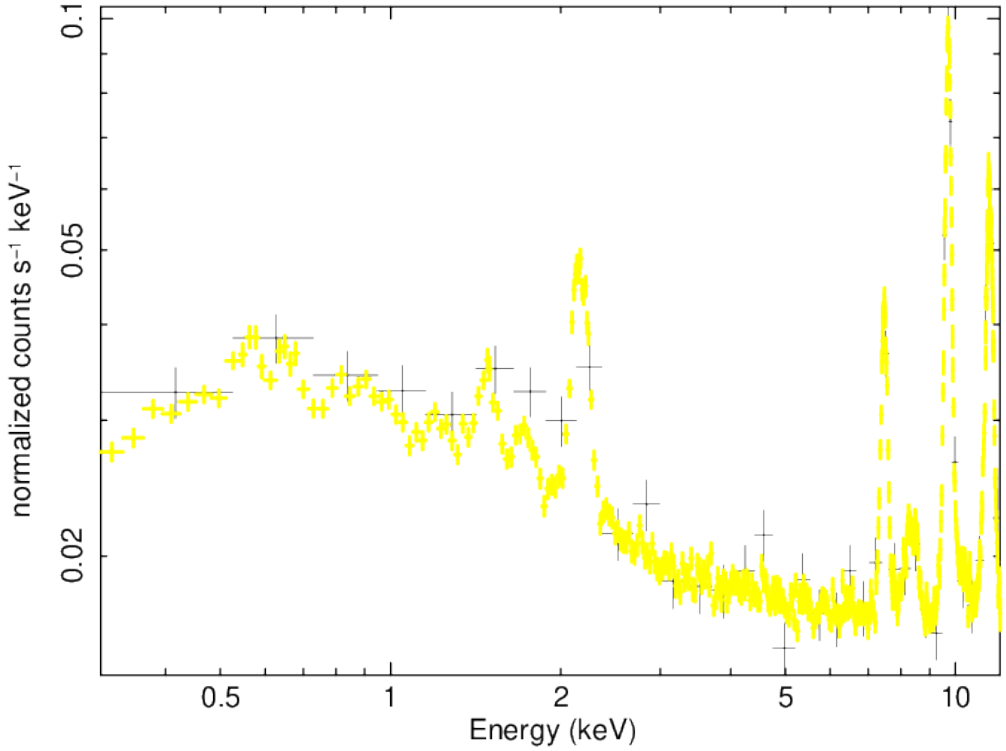

Figure 4.2: This plot shows the spectrum in the blank-sky (yellow) and target observation (black) in a source free region. In this case they agree well.

APEC model (see Figure 4.1); however, for observations where this is not the case, we vary the temperature of the APEC model to find the best fitting temperature (these exceptions are detailed in §4.3.11). This 'extra' model is included in the spectral fitting as an additional component to the cluster emission. One diagnostic that we use to check that the amended blank sky background file is indeed a good approximation to the local background in our observation, is to extract a spectrum from a source-free region from our observational data, and compare it to the spectrum from the blank sky background file. An example of this diagnostic is shown in Figure 4.2.

4.3.4 Image Preparation and Contaminating Source Detection

X-ray images are affected by instrumental artefacts, mainly due to vignetting (the decrease in effective area as a function of distance from the aimpoint), chip gaps, and bad pixels. We generate an exposure map to correct for these issues that gives the effective area as a function of position.

We create images from the events files in the 0.7 - 2.0 keV band, which are used for source detection. To remove point sources in the image, we used the wavelet decomposition detection algorithm (described in Vikhlinin et al., 1998) to detect point sources in the 0.3 - 7 keV band and extended sources in the 0.7 - 2 keV band, and excluded all contaminating sources from all further analysis. We also visually inspected all images with their corresponding source detections to verify the process had run successfully. There were often cases where the cool core of a cluster was mistaken for a point source and excluded, and this had to be amended. Additionally, there were cases where we had to manually exclude extra sources: we give further detail in §4.3.11. These additional contaminating sources were also removed from all further analysis.

4.3.5 Finding the cluster centroid and extent of the cluster emission

We also used the image (in the 0.7 - 2 keV band) to find the centroid of the cluster emission, using the coordinates of the cluster target as our starting point, and refining the cluster centroid five times within a circle of radius $150''$, and five more in a radius of $50''$.

To find the extent of the cluster emission, we define the detection radius of the cluster where no cluster emission is detected at $> 0.5\ \sigma$. This low detection threshold ensures that there is no significant cluster emission detected outside the cluster extent radius. An iterative process is used to find the detection radius: i) the detection radius is determined, ii) the background is renormalised by a factor determined by the ratio of the source to background count rates (excluding the cluster emission within this radius) iii) a new detection radius is calculated. This

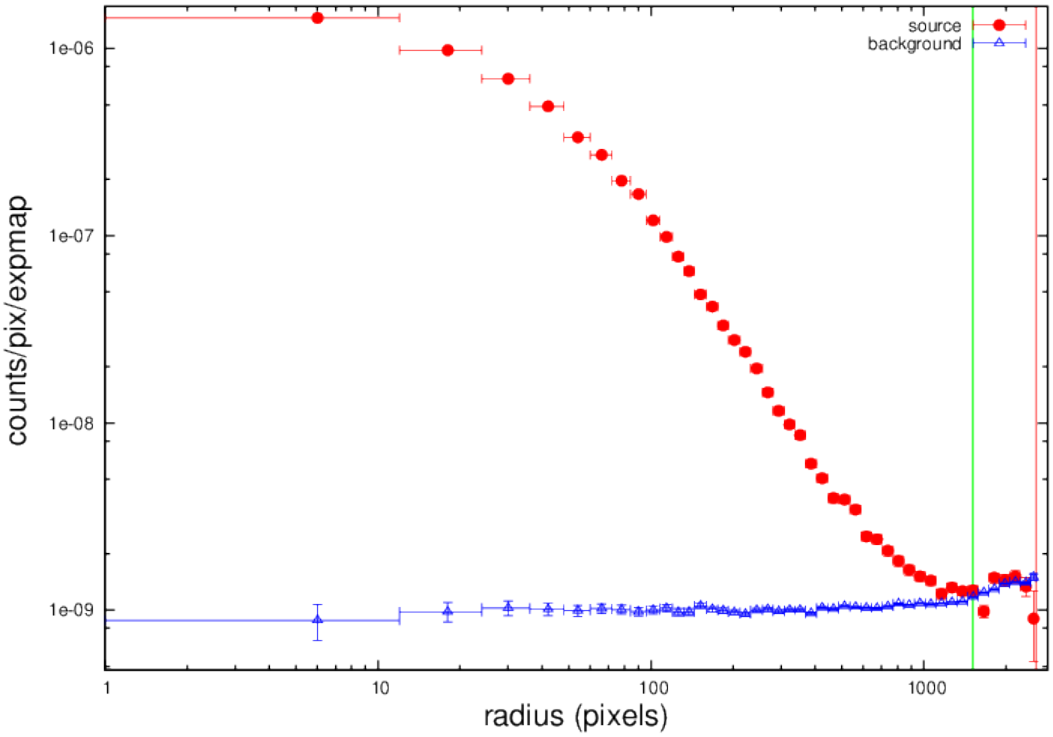

Figure 4.3: This plot shows how we determine the extent of the cluster emission: i.e. it is where the source meets the background. The red line shows the detection limit of the cluster, and the green line shows the outer radius of the measured surface brightness profile.

process continues until the cluster radius converges to within 1%. An example of the result of this process is show in Figure 4.3.

4.3.6 Determining R_{500}

A cluster's R_{500} is defined as the radius within which the mean density of the cluster is equal to 500 times the critical density of the Universe at the cluster's redshift. It is useful to measure cluster quantities within this radius as it means we can compare clusters in a consistent way, regardless of their redshift and mass.

We determine the cluster X-ray R_{500} from the hydrostatic mass profile, once obtained (for which details are given in Chapter 3). We mention how the R_{500} is determined here, as we calculate various cluster properties within this radius: the luminosity and the temperature, as well as the cluster mass and the gas mass. We also determine a separate R_{500} from the caustic mass profile, though only use the caustic mass within the caustic R_{500} in §4.8.2.

4.3.7 Combining Observations

For a number of clusters in our sample, we have multiple observations of the same cluster (see Table 4.1), and want to combine these observations so as to obtain a higher total exposure time for the cluster. We do this by combining the images, and then proceed with the next steps in the data analysis as normal, with the exception of the spectral analysis. For this, spectra are taken from individual observations (masking sources detected from when both observations were combined in a single image, however) and fitted simultaneously due to different response files (ARF and RMF, described in §1.3.1 c)) for each observation due to the different dates of each observation. We also use different blank-sky background files for the same reason (if appropriate).

4.3.8 Temperature and Abundance Radial Profiles

To extract the temperature profile (to which we later fit our projected 3D temperature model), spectra were extracted from annuli defined such that they had a minimum signal-to-noise ratio (SNR) of 30, and each bin was at least 30% bigger than the adjacent inner bin. For the final bin, if it has a SNR > 15, then we include it in our temperature profile and use it for all further analysis (note that the last bin is added regardless of whether it meets the other criterion of being

at least 30% bigger than the previous bin). Adding this final bin extends the outer radius of the measured temperature profile, such that the 3D temperature profile can be constrained to a larger radius during the fitting process. In some cases, the temperature was not well constrained in the outermost bin due to the relatively low SNR, and we thus discarded outer bins where the temperature was $>$ 15 keV, or the errors on the temperature were $>$ 50% (A697, A773, A1758, A1914, A2034), or if including the last bin led to an un-physical sudden increase in the best fit temperature profile (ZW8197, A1978, A1204, ZW3146).

For the cluster spectral fitting, fits to the cluster spectra were performed in the 0.6 - 9 keV band using a PHABS $\times$ APEC model (Smith et al., 2001), using ATOMDB version 3.0.9 and relative abundances fixed to the solar ratios of Asplund et al. (2009). The absorbing column was set to the NHTOT value (Willingale et al., 2013), and we used XSPEC for our cluster spectral fitting, using the *C*-statistic, and grouping our spectra to have at least 5 counts per bin. We note that any additional contribution from the CXB that is not included in the standard blank sky background files (which we find in §4.3.3) are also included in this fitting as an 'extra' component.

From our cluster spectral fitting we obtained an abundance profile. The radius out to which we could measure the abundance profile was limited by the radial extent of the measured temperature profile. Outside this outer radius we assumed an abundance of 0.3 $Z_{\odot}$. This is in line with observations that abundance profiles flatten at large cluster radii (Leccardi & Molendi, 2008). Leccardi & Molendi (2008) found that the abundance profile for galaxy clusters flattens at $\sim$ 0.3 $Z_{\odot}$ at 0.2 R_{180} (in fact they report that it flattens at 0.2 $Z_{\odot}$ using the metallicity tables of Anders & Grevesse, 1989, but converting to the metallicity tables of Asplund et al., 2005 used in our analysis, scaling by a factor of 1/0.6 is necessary); Rasmussen & Ponman (2007) and Urban et al. (2017) found similar results. As this flattening of the abundance profile occurs significantly before even our lowest outer radius value, this assumption is valid, and we do, in general, see a flattening of the abundance profile at larger radii for our cluster sample.

4.3.9 Surface Brightness Profile and Emission Measure Profile

To extract the surface brightness profile (to which we later fit our projected 3D ICM density model), we used bins (annuli) defined such that each had net counts of at least 50, and each bin was at least 5% bigger than the adjacent inner bin. The surface brightness profile is measured in the 0.7 - 2 keV band.

We can convert the surface brightness profile into an integrated emission measure profile, by some simple manipulation. This step is necessary, as it is actually the emission measure profile to which we fit our 3D density model (see §3.2). For each annular bin in the surface brightness profile, we can convert the net counts to an integrated emission measure. First, we extract a spectrum from each annular bin. An absorbed APEC model, using absorption set to the NHTOT value (Willingale et al., 2013) and the metal abundance set to the observed (projected) abundance profile obtained from the cluster spectral fitting, is used to simulate a high SNR spectrum. The data in each of the surface brightness profile's annular bins were not sufficient to measure a temperature, so the temperature of the simulated APEC model in each annular bin is set to the temperature given by the 3D temperature model profile at that radius. We then generate a Redistribution Matrix File, RMF, and an Ancilliary Response File, ARF, (see §1.3.1 c)) from each annular bin in the surface brightness profile, and convolve the simulated spectrum with these files. By setting the normalization of the APEC model to 1, XSPEC will return a predicted count rate (given the observation's ARF and RMF) and by rearranging the equation for the APEC model's normalization[7] we can obtain the emission measure integral for each bin (using the count rate that we derived from the surface brightness profile).

The APEC model's normalization is defined as follows:

$$\text{normalization} = \frac{10^{-14}}{4\pi(D_{\rm A}(1+z))^2}\int n_{\rm e}n_{\rm p}dV \tag{4.1}$$

where $D_{\rm A}$ is the angular diameter distance to the cluster, and is defined in equation 1.19.

By setting the normalisation in equation 4.1 to 1 and obtaining the predicted count rate for this, CR_1, and using the count rate that is derived from the surface brightness profile, $CR_{\rm obs}$, we can obtain the emission measure integrated over the annular bin, the integrated emission measure (or emission measure integral), $\int n_{\rm p}n_{\rm e}dV$:

$$\int n_{\rm e}n_{\rm p}dV = \frac{4\pi(D_{\rm A}(1+z))^2}{10^{-14}} \times \frac{CR_1}{CR_{\rm obs}} \tag{4.2}$$

where the $\frac{CR_1}{CR_{\rm obs}}$ term is equivalent to the normalisation required for the observed count rate.

To convert the integrated emission measure, $\int n_{\rm p}n_{\rm e}dV$, to the emission measure, $\int n_{\rm p}n_{\rm e}dl$, we can simply divide the integrated emission measure by the area of the appropriate annular bin. We note that throughout this thesis we always refer to the quantity $\int n_{\rm p}n_{\rm e}dV$ as the integrated

[7] https://heasarc.gsfc.nasa.gov/xanadu/xspec/manual/XSmodelApec.html

emission measure or the emission measure integral, and refer to the quantity $\int n_{\rm p} n_{\rm e} dl$ as the emission measure, though there are various different names for these quantities in the literature.

4.3.10 Determining the dynamical state of a cluster

To determine if a cluster was both dynamically relaxed and hosted a cool core (RCC), we used the following indicators: the central cooling time and the cuspiness of the density profile were used to probe for the presence of a CC, and the centroid shift was used to probe the dynamical state of the cluster. Giles et al. (2017) define and justify these indicators, but we summarize them here.

The central cooling time is defined (in Sarazin, 1988) as

$$t_{cool} = 8.5 \times 10^{10} \mathrm{yr} \left(\frac{n_{\rm p}}{10^{-3} \mathrm{cm}^{-3}} \right)^{-1} \left(\frac{kT_{\rm CCT}}{8.62 \mathrm{keV}} \right)^{1/2} \tag{4.3}$$

where $n_{\rm p} = \sqrt{(1.17 n_{\rm p} n_{\rm e})}/1.17$, and $n_{\rm p} n_{\rm e}$ is from the best fitting 3D gas density models found from the `emcee` fitting (see §3.2). $kT_{\rm CCT}$ is the temperature measured by extracting a spectrum within $0.048 R_{500}$ (the same radius as used in Hudson et al., 2010; we note that this is within a radius much larger than the cooling radius found in most cool cores).

The cuspiness of the density profile refers to whether the density profile was centrally peaked.

The centroid shift, w, is defined as the standard deviation of the distance between the centroid of the cluster and the X-ray peak (as in Poole et al., 2006). The centroid shift is measured by taking the variance of the distance between the centroid and the X-ray peak, as measured within a series of apertures from 0.05 - 1 R_{500} (centred on the X-ray peak), in steps of 0.05 R_{500}. Errors for the measurement of the centroid shift were obtained using a Monte-Carlo method. We made 100 randomised realisations of the source and background images, drawing pixel values from a Poisson distribution whose centre was the observed photon counts in that pixel. These randomisations were then analysed using the same method that was applied to the real images giving us a distribution of centroid shift values, from which errors could be calculated.

For a cluster to be a RCC cluster, the central cooling time must be < 7.7 Gyr, and the logarithmic slope at 0.04 R_{500} must be > 0.7 (both indicators following the cut suggested in Hudson et al., 2010), and the centroid shift must be < 0.009, following the cut used in Giles et al. (2017). Our definition of what constitutes a RCC cluster is a conservative one: for our

sample of 44 clusters, 10 are classified as RCC. This is a similar fraction to the fraction found in other studies (e.g. Bauer et al., 2005; Chen et al., 2007) We present our classifications in Table 4.2.

4.3.11 Notes on individual clusters

In this section we detail where we differed from the analysis methods presented in §4.3, or any particular points of interest for each cluster in our sample.

When extended sources other than the cluster were detected in the images, we attempted to determine if they were physically associated with the cluster or were unrelated projected sources. We did this by querying the NASA Extragalactic Database[8] to see if the source has a known counterpart.

- A773 - for ObsID 5006, due to flaring we only use the first 15 ks of the observation.
- MS0906 - At 9:08:58.02, 11:01:58.36 (RA, DEC = 137.24, 11.03), there is a large region of diffuse extended emission which we mask with a 230″ radius circle.
- ZW2701 - At 9:53:05.85, 51:49:16.39 (148.27, 51.82) there is a region of extended emission that we mask with a 150″ radius circle.
- A963 - for ObsID 903, a temperature of 0.28 keV was used when fitting the APEC model to the soft background residuals.
- ZW3146 - for ObsID 9371, due to flaring we only use the first 34 ks of the observation.
- A1423 - for ObsID 538, due to flaring we only use the first 11 ks of the observation.
- A1553 - for ObsID 12254, due to flaring we only use the first 10 ks of the observation.
- A1682 - At 13:06:59.904, 46:31:40.65 (196.75, 46.53) there is a cluster galaxy (Morrison et al., 2003) and at 13:07:13.47, 46:29:02.31 (196.81, 46.48) there is another small region of extended emission. Both sources are masked by radius 40″ circles.
- A1763 - At 13:34:52.8, 40:57:21.6 (203.72, 40.96) we use a 90″ radius circle to mask a region of extended emission, which is likely to be associated with a previously known X-ray source (Evans et al., 2010).
- A1930 - At 14:32:42.72, 31:33:50.4 (218.18, 31.56) there is a region of extended emission that we mask with a circle of radius 160″. The extended emission is the cluster RM J143242.6+313407.1 (Rozo et al., 2015) at redshift $z = 0.137$, and is likely part of the same dark matter halo as A1930 (z = 0.1308).

[8] https://ned.ipac.caltech.edu/

Cluster	Cuspiness	tcool (Gyr)	w $10^{-3}R_{500}$	Status
ZW1478	$0.44^{+0.14}_{-0.08}$	$18.4^{+0.91}_{-0.70}$	12.7±5.53	NRCC
A655	$0.47^{+0.10}_{-0.12}$	$20.9^{+1.84}_{-1.98}$	20.9±8.18	NRCC
A697	$0.21^{+0.07}_{-0.03}$	$10.8^{+0.57}_{-0.62}$	4.97±4.25	NRCC
MS0906	$0.35^{+0.03}_{-0.02}$	$9.48^{+0.31}_{-0.32}$	58.3±2.45	NRCC
A773	$0.24^{+0.05}_{-0.02}$	$9.22^{+0.77}_{-0.69}$	5.82±2.24	NRCC
A795	$0.85^{+0.03}_{-0.02}$	$8.86^{+0.16}_{-0.14}$	8.94±0.98	NRCC
ZW2701	$0.85^{+0.02}_{-0.01}$	$2.01^{+0.02}_{-0.03}$	5.11±0.46	RCC
A963	$0.63^{+0.02}_{-0.02}$	$6.93^{+0.17}_{-0.14}$	2.1±1.27	NRCC
A980	$0.5^{+0.06}_{-0.08}$	$14.4^{+0.66}_{-0.67}$	10.5±1.35	NRCC
ZW3146	$1^{+0.01}_{-0.01}$	$1^{+0.01}_{-0.01}$	6.87±0.32	RCC
A990	$0.36^{+0.08}_{-0.08}$	$15.4^{+0.93}_{-0.91}$	8.42±3.38	NRCC
ZW3179	$1^{+0.02}_{-0.02}$	$9.99^{+0.41}_{-0.28}$	3.56±0.80	NRCC
A1033	$0.38^{+0.03}_{-0.02}$	$11^{+0.38}_{-0.41}$	22.1±2.64	NRCC
A1068	$1.1^{+0.02}_{-0.03}$	$5.71^{+0.09}_{-0.08}$	4.38±0.35	RCC
A1132	$0.2^{+0.08}_{-0.05}$	$16.6^{+1.05}_{-1.16}$	3.25±1.79	NRCC
A1201	$0.87^{+0.03}_{-0.04}$	$12^{+0.25}_{-0.38}$	17.3±2.09	NRCC
A1204	$1.2^{+0.02}_{-0.02}$	$3.33^{+0.04}_{-0.04}$	4.65±0.91	RCC
A1235	$0.77^{+0.04}_{-0.04}$	$28^{+1.74}_{-1.30}$	2.4±0.80	NRCC
A1246	$0.36^{+0.14}_{-0.11}$	$13.5^{+1.89}_{-1.56}$	5.91±6.21	NRCC
A1302	$0.65^{+0.05}_{-0.06}$	$16.2^{+0.68}_{-0.54}$	4.85±1.80	NRCC
A1413	$0.78^{+0.02}_{-0.02}$	$11.9^{+0.13}_{-0.17}$	1.15±0.41	NRCC
A1423	$0.83^{+0.04}_{-0.03}$	$3.72^{+0.14}_{-0.16}$	13±1.30	NRCC
A1437	$0.062^{+0.08}_{-0.01}$	$30.2^{+6.45}_{-5.40}$	17.4±9.66	NRCC
A1553	$0.11^{+0.16}_{-0.02}$	$17.4^{+1.44}_{-1.69}$	17.5±9.47	NRCC
A1682	$0.49^{+0.07}_{-0.08}$	$14.9^{+1.09}_{-0.90}$	37.5±6.19	NRCC
A1689	$0.9^{+0.01}_{-0.03}$	$3.62^{+0.05}_{-0.04}$	2.19±0.67	RCC
A1763	$0.22^{+0.10}_{-0.03}$	$12.3^{+0.66}_{-0.74}$	16.3±4.77	NRCC
A1835	$1.2^{+0.00}_{-0.00}$	$1.2^{+0.01}_{-0.01}$	2.78±0.10	RCC
A1918	$0.84^{+0.06}_{-0.05}$	$10.1^{+0.31}_{-0.29}$	6.1±1.33	NRCC
A1914	$0.19^{+0.02}_{-0.01}$	$7.12^{+0.15}_{-0.16}$	13.8±0.74	NRCC
A1930	$1.1^{+0.01}_{-0.02}$	$11.7^{+0.26}_{-0.27}$	1.41±0.48	NRCC
A1978	$0.62^{+0.06}_{-0.04}$	$9.89^{+0.40}_{-0.33}$	20.6±5.01	NRCC
A2009	$1^{+0.03}_{-0.05}$	$8.14^{+0.25}_{-0.18}$	1.25±0.46	NRCC
RXJ1504	$1.2^{+0.00}_{-0.00}$	$0.906^{+0.00}_{-0.01}$	1.52±0.16	RCC
A2034	$0.1^{+0.01}_{-0.00}$	$19.7^{+0.46}_{-0.52}$	6.86±1.87	NRCC
A2050	$0.43^{+0.09}_{-0.09}$	$17.8^{+0.97}_{-1.01}$	15±9.39	NRCC
A2069	$0.16^{+0.05}_{-0.11}$	$41.6^{+1.68}_{-1.84}$	10.9±4.82	NRCC
A2111	$0.17^{+0.07}_{-0.02}$	$12.3^{+1.02}_{-1.10}$	29.7±10.66	NRCC
A2219	$0.28^{+0.03}_{-0.04}$	$10.8^{+0.40}_{-0.42}$	11.2±2.58	NRCC
ZW8197	$1^{+0.03}_{-0.04}$	$10.6^{+0.22}_{-0.28}$	1.43±0.55	NRCC
A2259	$0.44^{+0.07}_{-0.08}$	$11.2^{+0.50}_{-0.59}$	7.28±5.57	NRCC
RXJ1720	$1.1^{+0.02}_{-0.04}$	$2.37^{+0.04}_{-0.04}$	2.22±0.47	RCC
A2261	$0.7^{+0.02}_{-0.02}$	$5.77^{+0.12}_{-0.13}$	1.58±0.34	RCC
RXJ2129	$0.98^{+0.02}_{-0.02}$	$1.93^{+0.04}_{-0.04}$	3.22±0.44	RCC

Table 4.2: Summary of dynamical states.

- A2009 - for ObsID 10438, a temperature of 0.27 keV was used when fitting the APEC model to the soft background residuals.
- A2069 - There is a separate cluster at 15:24:25.846, +30:00:16.039 (231.11, 30.00) at redshift $z = 0.119$ (MaxBCG J231.10029+30.00604, Koester et al., 2007); it is likely part of the same dark matter halo as A2069. We mask it with a rectangle of $23'$ by $14'$ at an inclination angle of 35°. There is either a group of point sources or a filament at 15:23:38.87, +29:58:30.226 (230.91, 29.98) that we mask using an ellipse. In addition, for ObsID 4965, due to flaring we only use the first 40 ks of the observation.
- A2261 - At 17:22:12.78, 32:06:36.95 (260.55, 32.11) there is a region of diffuse extended emission that we mask with a $80''$ radius circle. This source is associated with a galaxy cluster (GMBCG J260.55436+32.11438) at $z = 0.304$ (Hao et al., 2010), so is not associated with A2261 which is at redshift $z = 0.2242$.
- RXJ1720 - for ObsID 4361, due to flaring we only use the first 15 ks of the observation.

4.4 Mass determination

The two mass methods (the forward-fitting method and backward-fitting method) that we use in this chapter are detailed in Chapter 3. For the main results in this paper we use the masses derived using the forward-fitting method, however we do compare the masses from both mass methods as a consistency check (see §4.8). From these mass methods we can also obtain the gas mass (see Chapter 3), which we use in §4.8.

For the fitting methods used for the mass determination, detail on the priors used can be found in Chapter 3. However, we note that for A1835, we relaxed the priors on the parameters r_t, a, b and c in the temperature model to five times the original upper limit in order to get a good fit to the temperature profile. A1835 is a special case (discussed further in §4.7), with an unphysical hydrostatic mass profile (Bonamente et al., 2013).

4.5 Galaxy caustic masses

HeCS uses the caustic method (described in §1.7.2 a)) to estimate masses for all of the clusters in this sample. In brief, the caustic mass method uses spectroscopic redshifts for a number of member galaxies in a cluster (on average ∼180 for our sample, see column 7 Table 4.1), to determine their line of sight velocity (relative to the centre of mass of the cluster), which when

combined with their projected distances from the cluster centre, can be used to estimate a mass profile of a cluster.

We note that we use 0.5 for the value of the filling factor, $\mathcal{F}_\beta$, which is the appropriate $\mathcal{F}_\beta$ for the algorithm (Diaferio, 1999) used to calculate the caustic masses presented in Rines et al. (2013). We also note that the uncertainties on the caustic masses are due to the uncertainties in the location of the caustics. Thus, clusters with poorly defined caustics, due to the distribution of member galaxies in the redshift diagram itself, or due to having a low number of member galaxies with measured line of sight velocities, have larger uncertainties on their cluster mass. The uncertainties are driven by the statistical precision of the measurements of the caustic location, and do not account for the $\sim$30% scatter expected between true mass and caustic mass (Serra et al., 2011), driven largely by the viewing angle of the cluster.

4.6 Modelling the mass biases

Using the hydrostatic masses as calculated from the `emcee` chains, we use a Bayesian framework (of which a graphical plot is shown in Figure 4.4) to constrain the bias and scatter between the two mass measurement techniques. The modelling framework is the same as in Maughan et al. (2016), but we present it in detail in this section. In Table 4.3 and Figure 4.6, we report the mean and standard deviation of the X-ray and caustic mass measurements in *linear space*, which is the form in which masses are often reported in other work. However, we find that the uncertainties on the mass measurements are in fact better described using lognormal errors, and so choose to model the biases associated with the caustic and X-ray hydrostatic mass measurement methods in log space. Additionally, the ratio of two quantities that are lognormally distributed follows a lognormal distribution, whereas the ratio of two quantities that are normally distributed follows a Cauchy distribution, which has various summary statistics that are undefined (such as variance), such that the resulting uncertainty on the ratio of the X-ray mass to the caustic mass, M_X/M_C, would be harder to interpret. Under the assumption that the uncertainties on the masses are lognormal, we can relate the mean (M) and standard deviation (S) of the X-ray and caustic mass measurements in linear space to the mean (μ) and standard deviation (σ) in base 10 log space as:

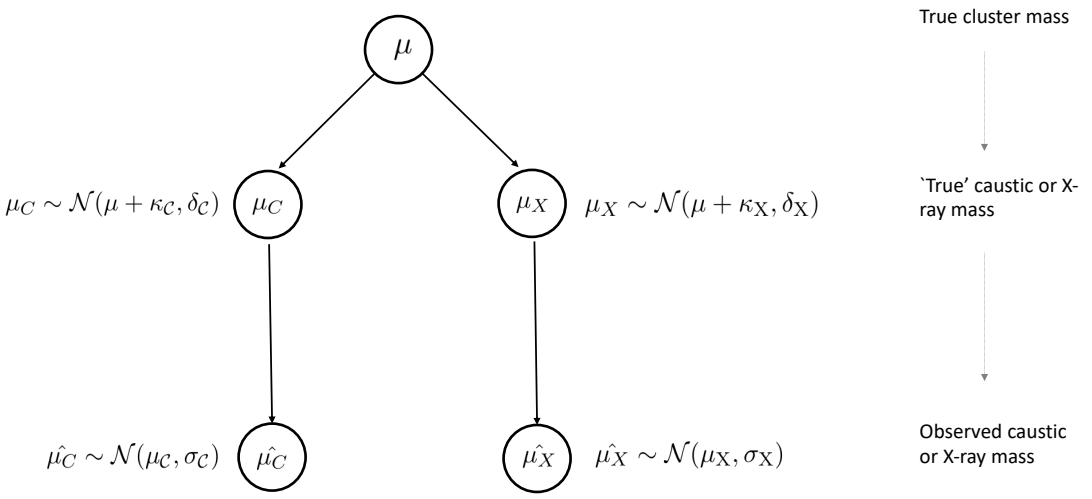

Figure 4.4: A graphical model summarising the mass modelling framework that we present in §4.6. The symbols are also described in that section; in summary, the symbol μ represents mass, κ is the bias term, δ is the scatter term and σ is the measurement error term (all in log space), and the c and x subscripts denote whether the quantity refers to that from the X-ray or caustic method. The $\hat{}$ refers to an observed quantity. $\mathcal{N}$ refers to a normal distribution, and "$\sim$" is equivalent to "is distributed as".

$$\mu = \log_{10}\left(\frac{M}{\sqrt{1+S^2/M^2}}\right) \tag{4.4}$$

$$\sigma = \sqrt{\log_{10}\left(\frac{S^2}{M^2}+1\right)}. \tag{4.5}$$

To constrain the bias and scatter between the two mass measurement techniques, we use a Bayesian framework. We assume that a given cluster has a 'true' mass, μ. A given cluster also has a 'true' caustic mass, μ_C, and 'true' hydrostatic mass, μ_X, which are related to the true mass, μ, as follows:

$$\mu_X \sim \mathcal{N}(\mu + \kappa_{\rm X}, \delta_{\rm X}) \tag{4.6}$$

$$\mu_C \sim \mathcal{N}(\mu + \kappa_{\rm C}, \delta_{\rm C}) \tag{4.7}$$

where κ_X and κ_C parametrise the bias between the true mass and the 'true' hydrostatic and caustic mass, respectively, and δ_X and δ_C parametrise the intrinsic scatter between the true mass and the 'true' hydrostatic and caustic mass, respectively. $\mathcal{N}$ refers to a normal distribution, and "$\sim$" is equivalent to "is distributed as".

These 'true' hydrostatic and caustic masses are related to the observed hydrostatic and caustic masses ($\hat{\mu_X}$ and $\hat{\mu_C}$, respectively), as follows:

$$\hat{\mu_X} \sim \mathcal{N}(\mu_{\rm X}, \sigma_{\rm X}) \tag{4.8}$$

$$\hat{\mu_C} \sim \mathcal{N}(\mu_{\rm C}, \sigma_{\rm C}) \tag{4.9}$$

where σ_X and σ_C represent the standard deviation of the lognormal error for the observed hydrostatic and caustic masses, respectively.

We apply weak priors on the model parameters. For the cluster masses in log space, μ, μ_X and μ_C, we applied a uniform prior from 12 to 17. For the bias terms in log space, κ_X and κ_C, we applied normal priors with mean 0 and a standard deviation of 1. For the intrinsic scatter terms in log space, δ_X and δ_C, we applied normal priors (truncated at zero) with mean 0.09 and a standard deviation of 2.2, corresponding to a weak prior with a mean of $\sim$20% in normal space.

Using the model described above, we used the observational data that we have for each cluster ($\hat{\mu_X}$, σ_X, $\hat{\mu_C}$, σ_C) to constrain the 'true' hydrostatic and caustic mass terms (μ_X, μ_C), the scatter terms (δ_X and δ_C) and bias terms (κ_X and κ_C) for all of the clusters in our sample. Clearly, both the scatter terms and the bias terms will be degenerate. However, the intrinsic scatter between the X-ray and caustic mass measurements

$$\delta = \sqrt{\delta_X^2 + \delta_C^2} \tag{4.10}$$

and the mean bias between the X-ray and caustic mass measurements

$$\kappa = \kappa_X - \kappa_C = \mu_X - \mu_C = \log_{10}\left(\frac{M_X}{M_C}\right) \tag{4.11}$$

can be constrained by the data. When reporting the values of κ and δ (in Tables 4.4, 4.5, and 4.6), we report the median value of δ, with errors given as the difference between the median and the 16th and 84th percentiles. For κ, we report the mean value. As $\kappa = \log_{10}(M_X/M_C)$, and κ (a quantity in log space) is normally distributed, the posterior of M_X/M_C (a quantity in linear space) is lognormally distributed. Therefore, in Tables 4.4, 4.5, and 4.6), we summarise the posterior of M_X/M_C by reporting its median and errors in the same way that we did for δ. In these tables, the κ values are given in base 10 log space, and the δ values have been converted to percentage scatters.

We use the probabilistic programming language STAN[9] to implement the model described above, specifically using the No-U-Turn Sampler (Hoffman & Gelman, 2011). We sampled the parameters in our model with four chains of 5,000 steps each. The analysis can be done at any radius, and so by repeating the analysis at increasing radii, we produced a profile of the mean bias between the two mass measurement methods.

4.7 Initial Results for the 44 Cluster Sample

We show the caustic and hydrostatic mass profiles for each cluster in Figure 4.5. For A1835, the hydrostatic mass profile decreases unphysically at around R_{500}. This is believed to be due to the

[9] http://mc-stan.org

Cluster	z	Status	R_{500} Mpc	M_X $10^{14}M_\odot$	M_C $10^{14}M_\odot$	R_{500} NFW Mpc	M_X NFW $10^{14}M_\odot$	M_X NFW $10^{14}M_\odot$	Ngal subsample	Ngal
ZW1478	0.103	NRCC	0.80 ± 0.03	1.6 ± 0.3	0.8 ± 0.0	0.8±0.1	1.7±0.6	1.7±0.4	low	82
A0655	0.127	NRCC	1.08 ± 0.04	4.1 ± 0.8	3.9 ± 0.2	0.85±0.14	2.1±1.1	2.5±1.1	high	315
A0697	0.281	NRCC	1.47 ± 0.05	11.9 ± 1.8	5.9 ± 2.8	1.46±0.14	10.6±3.0	10.7±2.0	mid	185
MS0906	0.177	NRCC	1.04 ± 0.02	3.7 ± 0.3	1.8 ± 0.2	1.08±0.06	4.3±0.7	4.2±0.5	low	101
A0773	0.217	NRCC	1.32 ± 0.03	8.0 ± 0.9	9.8 ± 0.1	1.37±0.11	8.7±2.0	8.4±1.3	mid	173
A0795	0.137	NRCC	1.09 ± 0.04	4.2 ± 0.7	3.5 ± 0.1	0.97±0.06	3.0±0.6	3.4±0.5	mid	179
ZW2701	0.216	RCC	1.07 ± 0.03	4.2 ± 0.5	2.4 ± 0.7	1.1±0.04	4.4±0.5	4.4±0.3	low	93
A0963	0.204	NRCC	1.06 ± 0.02	4.0 ± 0.3	4.1 ± 0.0	1.22±0.04	6.1±0.7	5.4±0.4	high	211
A0980	0.155	NRCC	1.36 ± 0.07	8.3 ± 2.2	5.7 ± 1.8	1.08±0.15	4.3±1.8	5.3±1.7	high	222
ZW3146	0.289	RCC	1.29 ± 0.03	8.0 ± 1.0	4.0 ± 1.8	1.53±0.05	12.0±1.2	10.2±0.6	low	106
A0990	0.142	NRCC	1.29 ± 0.05	6.9 ± 1.3	2.2 ± 0.8	1.34±0.12	8.2±2.2	7.8±1.3	low	91
ZW3179	0.142	NRCC	1.44 ± 0.13	10.2 ± 4.6	1.5 ± 0.1	0.96±0.1	3.0±1.0	3.8±1.1	low	69
A1033	0.122	NRCC	1.11 ± 0.02	4.3 ± 0.4	3.2 ± 0.0	1.39±0.08	9.0±1.6	7.3±0.8	mid	191
A1068	0.139	RCC	1.15 ± 0.05	5.0 ± 1.1	8.3 ± 0.7	1.02±0.04	3.6±0.4	4.0±0.3	low	129
A1132	0.135	NRCC	1.64 ± 0.08	14.1 ± 2.9	5.1 ± 0.2	1.63±0.13	14.6±3.5	14.7±2.3	mid	160
A1201	0.167	NRCC	1.15 ± 0.02	5.0 ± 0.5	3.1 ± 0.1	1.15±0.05	5.1±0.7	5.1±0.4	mid	165
A1204	0.171	RCC	0.91 ± 0.02	2.5 ± 0.3	1.4 ± 0.2	0.85±0.04	2.1±0.3	2.2±0.2	low	92
A1235	0.103	NRCC	1.05 ± 0.05	3.7 ± 0.9	2.0 ± 0.2	0.95±0.13	2.9±1.1	3.2±0.9	mid	131
A1246	0.192	NRCC	1.12 ± 0.06	4.8 ± 1.3	5.8 ± 0.1	1.04±0.19	3.8±2.2	4.0±1.7	high	226
A1302	0.115	NRCC	1.05 ± 0.03	3.6 ± 0.5	2.3 ± 0.0	0.77±0.07	1.6±0.4	1.9±0.4	mid	162
A1413	0.141	NRCC	1.33 ± 0.02	7.5 ± 0.6	6.7 ± 0.0	1.66±0.07	15.4±1.8	12.3±0.8	low	116
A1423	0.214	NRCC	1.10 ± 0.03	4.6 ± 0.5	4.1 ± 0.1	1.21±0.07	6.0±1.1	5.4±0.6	high	230
A1437	0.133	NRCC	1.19 ± 0.02	5.3 ± 0.4	9.4 ± 1.2	1.38±0.12	8.8±2.3	7.6±1.1	mid	194
A1553	0.167	NRCC	1.31 ± 0.06	7.4 ± 1.6	5.7 ± 0.0	1.23±0.17	6.3±2.6	6.6±2.0	mid	171
A1682	0.227	NRCC	1.13 ± 0.03	5.1 ± 0.7	6.7 ± 0.0	1.12±0.12	4.7±1.5	4.8±0.9	mid	151
A1689	0.184	RCC	1.47 ± 0.02	10.6 ± 0.7	10.4 ± 3.3	1.57±0.03	13.1±0.7	12.5±0.5	high	210
A1763	0.231	NRCC	1.23 ± 0.03	6.5 ± 0.8	10.9 ± 1.2	1.07±0.08	4.1±0.9	4.5±0.8	high	237
A1835	0.251	RCC	1.45 ± 0.02	10.7 ± 0.7	9.6 ± 0.7	1.49±0.01	11.2±0.3	10.9±0.2	high	219
A1918	0.139	NRCC	1.12 ± 0.05	4.6 ± 1.1	2.5 ± 0.1	1.11±0.13	4.6±1.6	4.6±1.1	low	80
A1914	0.166	NRCC	1.44 ± 0.04	9.9 ± 1.2	6.1 ± 0.2	1.62±0.07	14.4±1.8	13.0±1.1	high	255
A1930	0.131	NRCC	1.06 ± 0.05	3.8 ± 0.8	2.4 ± 0.2	0.88±0.04	2.3±0.3	2.6±0.3	low	76
A1978	0.146	NRCC	0.95 ± 0.02	2.8 ± 0.3	1.2 ± 0.3	1.05±0.11	3.9±1.2	3.5±0.7	low	63
A2009	0.152	NRCC	1.33 ± 0.05	7.7 ± 1.5	4.3 ± 0.2	1.08±0.05	4.2±0.6	4.9±0.6	mid	195
RXJ1504	0.217	RCC	1.30 ± 0.01	7.7 ± 0.4	3.1 ± 2.2	1.51±0.02	11.7±0.5	10.3±0.3	low	120
A2034	0.113	NRCC	1.31 ± 0.02	7.1 ± 0.3	5.5 ± 0.0	1.75±0.03	18.0±0.8	13.0±0.3	mid	182
A2050	0.119	NRCC	1.04 ± 0.03	3.5 ± 0.4	3.9 ± 1.1	0.98±0.06	3.2±0.6	3.3±0.5	low	106
A2069	0.114	NRCC	1.33 ± 0.04	7.5 ± 1.1	6.1 ± 0.1	1.05±0.09	3.9±1.0	4.9±1.0	high	441
A2111	0.229	NRCC	1.19 ± 0.03	5.9 ± 0.7	3.7 ± 0.4	1.18±0.07	5.6±0.9	5.6±0.6	mid	208
A2219	0.226	NRCC	1.55 ± 0.02	12.9 ± 0.6	10.1 ± 2.6	1.95±0.07	25.0±2.6	19.0±1.0	high	461
ZW8197	0.113	NRCC	0.87 ± 0.03	2.1 ± 0.4	1.9 ± 0.0	0.94±0.1	2.8±0.9	2.6±0.6	low	76
A2259	0.161	NRCC	1.16 ± 0.05	5.1 ± 1.0	4.6 ± 0.8	0.96±0.09	3.0±0.9	3.5±0.8	mid	165
RXJ1720	0.160	RCC	1.24 ± 0.03	6.3 ± 0.7	5.1 ± 0.3	1.39±0.06	9.0±1.1	8.2±0.7	high	376
A2261	0.224	RCC	1.25 ± 0.03	6.8 ± 0.7	3.3 ± 1.2	1.3±0.07	7.4±1.2	7.2±0.8	high	209
RXJ2129	0.234	RCC	1.31 ± 0.05	8.1 ± 1.5	6.1 ± 1.3	1.22±0.07	6.1±1.0	6.6±0.8	high	325

Table 4.3: Summary of the caustic and hydrostatic masses as calculated within the hydrostatic R_{500} for each cluster. All X-ray hydrostatic masses and R_{500} values are calculated using the forward-fitting mass method unless stated otherwise. Column 1 is the cluster name; column 2 is the redshift of the cluster; column 3 is the dynamical state of the cluster; column 4 is the (mean) R_{500}; column 5 is the mean hydrostatic mass at R_{500}; column 6 is the mean caustic mass at the X-ray R_{500}; column 7 is the median R_{500} calculated using the backward-fitting mass method; column 8 is the median M_{500} calculated using the backward-fitting mass method (at the median R_{500} calculated from using the backward-fitting mass method); column 9 is the mass from the backward-fitting method at the forward-fitting mass method mean R_{500} (column 4); column 10 is the Ngal subsample that each cluster is in (see §4.8.3 and §4.8.5); column 11 is the number of galaxies of each cluster within the caustics, as given in Rines et al. (2013). All errors are 1σ errors.

hydrostatic assumption becoming worse at large radii and is discussed in detail in Bonamente et al. (2013).

We calculate the caustic and hydrostatic M_{500} values from these mass profiles, using the R_{500} value as calculated from the hydrostatic mass profile (sometimes referred to as the X-ray R_{500}) for each cluster. We report these masses in Table 4.3 and they are compared in Figure 4.6 (we note that some of the errors on the caustic masses in Figure 4.6 are tiny as the errors are driven by statistical uncertainties on the measurements of the location of the caustics, and do not include the scatter between true mass and caustic mass, as mentioned in §4.5). We used both the caustic and hydrostatic M_{500} values for our main results, and note that this does introduce a covariance between the mass measurements. For this reason, we repeat the analysis using masses measured in a fixed radius of 1 Mpc and obtain fully consistent results (we note that only four out of the 44 clusters have $R_{500} < 1$ Mpc, see Table 4.3).

For our hydrostatic mass measurements, the observed temperature profiles were measured from our X-ray data for each cluster close to, or beyond, R_{500} for the majority of clusters (see Figure 4.5). For the mass profiles at radii greater than the extent of the temperature profile, we extrapolated the best fitting 3D temperature profile. As a consequence, our hydrostatic mass profiles beyond the extent of the measured temperature profiles are less robust. The median radius out to which the temperature profiles were measured is 0.95 R_{500}, and the range of radii is 0.51 - 1.62 R_{500}.

We show the observed M_X/M_C profile of each cluster (calculated as $\hat{\mu_X} - \hat{\mu_C}$) in Figure 4.7. NRCC and RCC clusters are shown in red and blue respectively. We also plot the mean M_X/M_C profile (calculated as the mean bias κ, see equation 4.11). The M_X/M_C profile is consistent with the bias not being radially dependent. At smaller radii the M_X/M_C value decreases, but this is likely due to the caustic masses being overestimated at small radii (see e.g. Figure 12 Serra et al., 2011).

We show the mean M_X/M_C profiles of NRCC and RCC clusters separately in Figure 4.8. These two profiles are consistent with each other.

We show the M_X/M_C ratio at R_{500} for all clusters in our sample in Figure 4.9. At R_{500} the caustic and hydrostatic masses generally do not agree well, with the hydrostatic mass on average $\sim 30\%$ higher than the caustic mass. Similar results are found for both the NRCC and RCC clusters.

We summarise our results in Table 4.4. We also include the mean bias κ, M_X/M_C ratio and

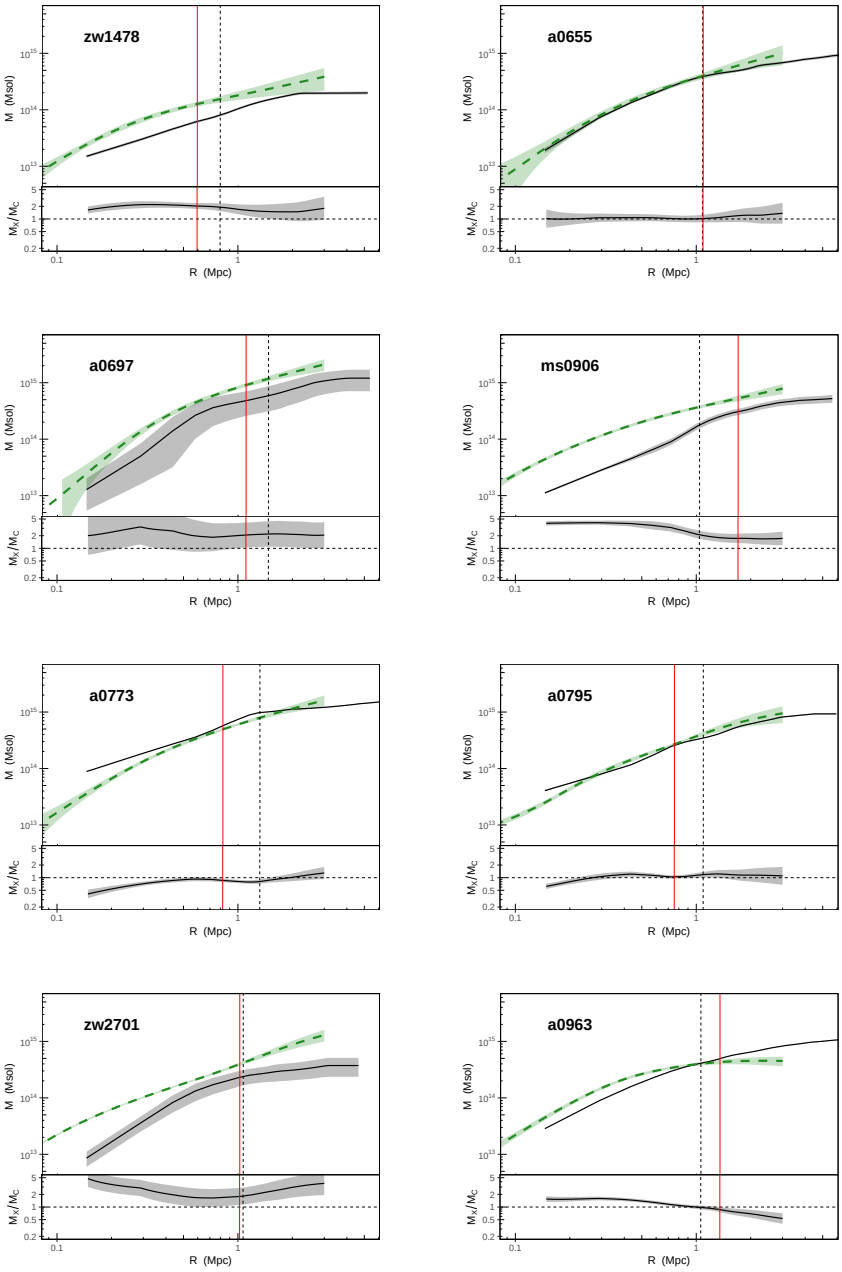

Figure 4.5: The caustic and hydrostatic mass profiles for each cluster as black solid and green dashed lines respectively in the upper panels of each plot. The ratio of the hydrostatic to caustic mass, M_X/M_C, are shown in the lower panels of each plot. 1σ uncertainties are shown by the shaded regions. The vertical black line is at the value of R_{500} as calculated from the hydrostatic mass profile; the solid red vertical line is at the outer radius of the measured temperature profile (note that hydrostatic masses beyond this radius are based on extrapolation).

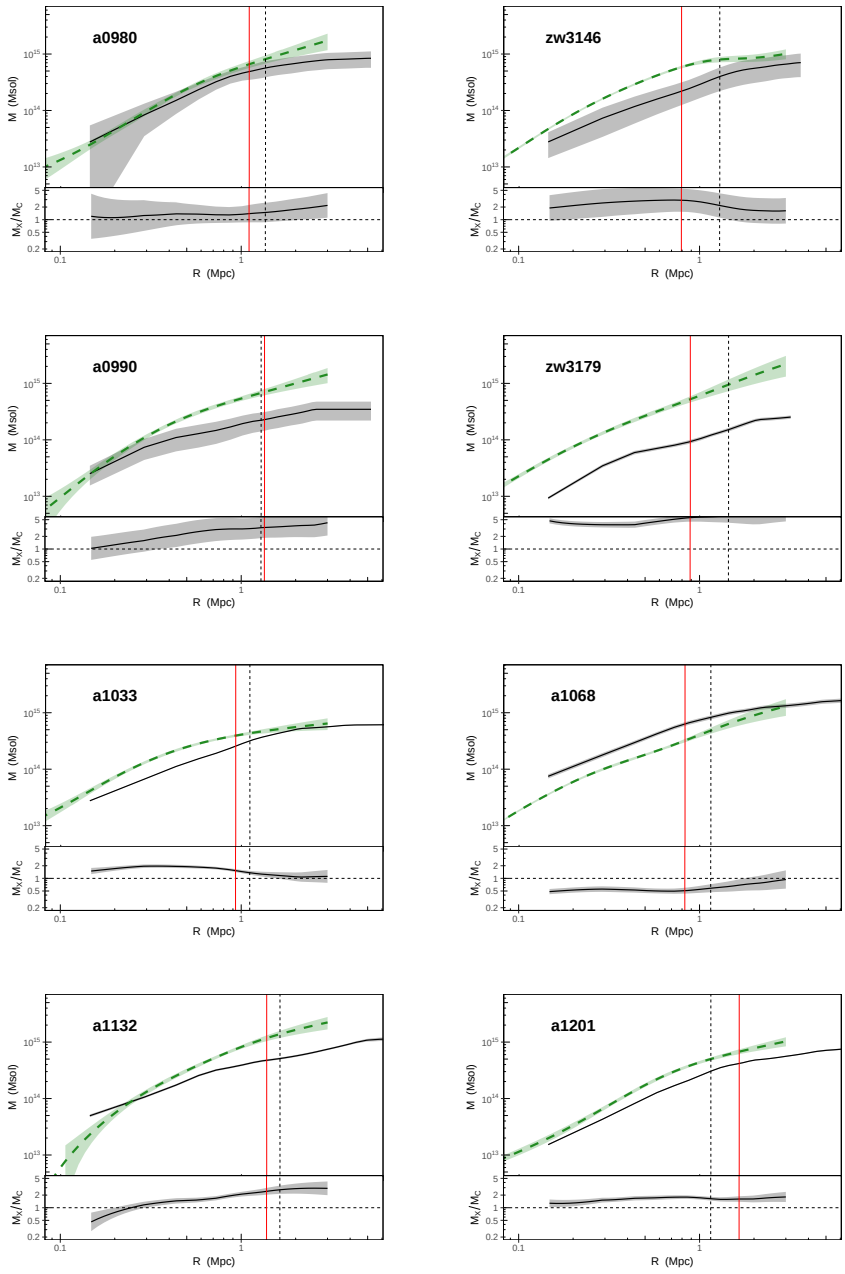

Figure 4.5: - *continued*

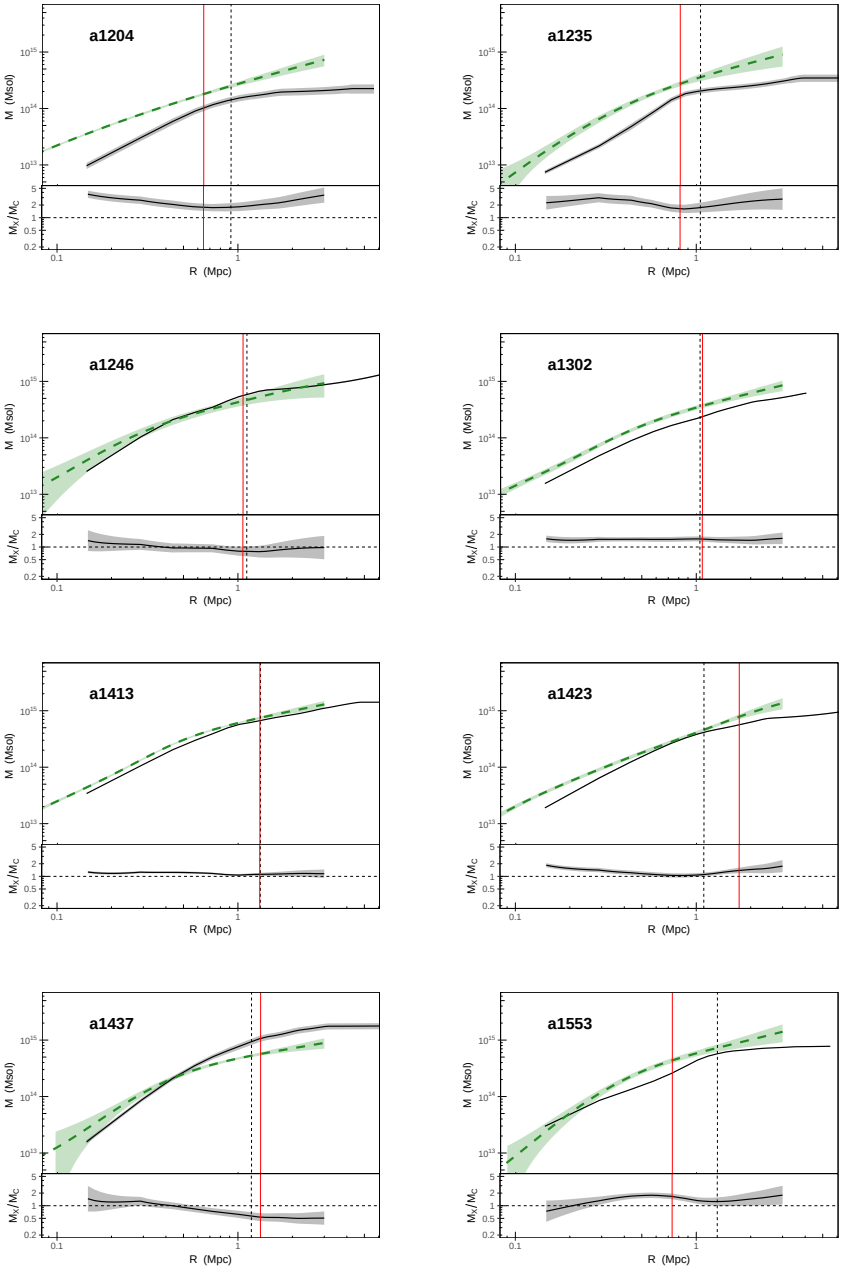

Figure 4.5: *- continued*

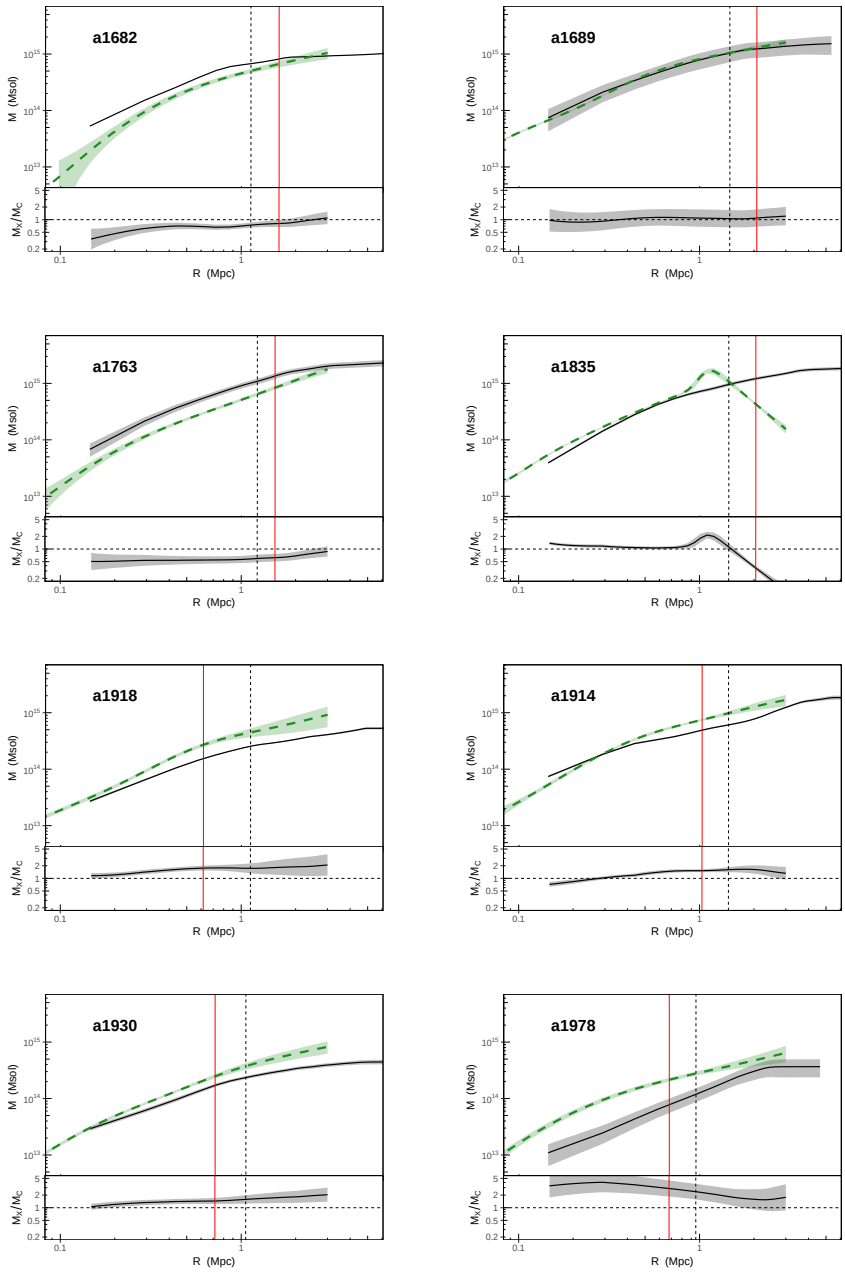

Figure 4.5: - *continued*

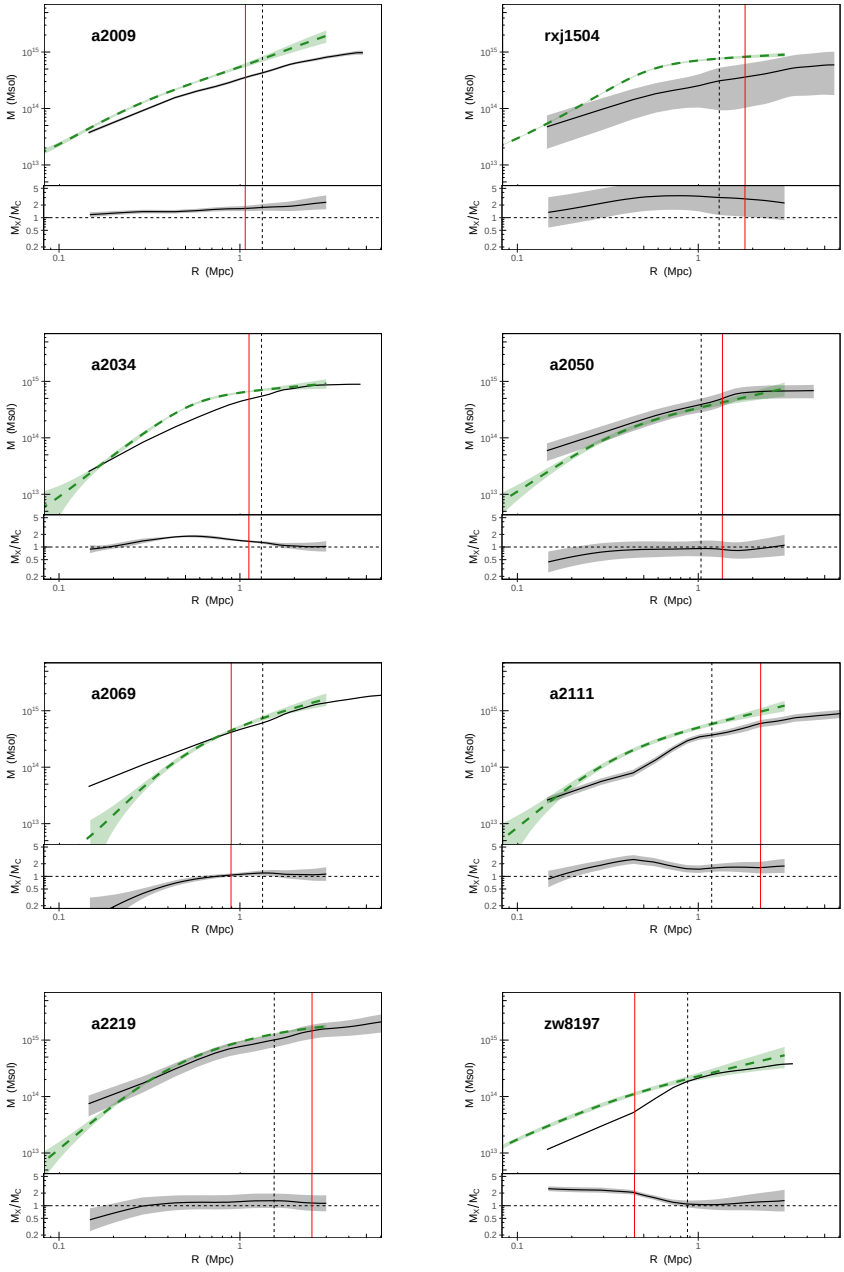

Figure 4.5: - *continued*

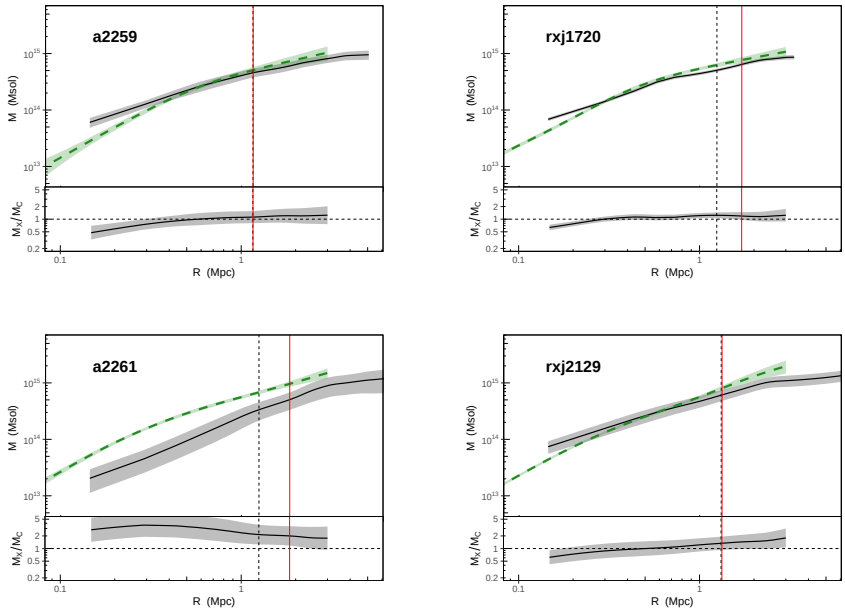

Figure 4.5: - *continued*

scatter δ for the caustic and hydrostatic masses computed at a fixed radius of 1 Mpc, and obtain fully consistent results with our main results where the masses were calculated at the hydrostatic R_{500}. This shows that the choice of aperture for the mass measurement is not important, and that our results are not significantly affected by scaling the caustic masses to the hydrostatic R_{500}.

Our results are somewhat unexpected, as the X-ray mass is expected to be underestimated compared to the true mass and the caustic mass expected to be overestimated compared to the true mass; thus we expect the M_X/M_C ratio to be less than 1, whereas we find it to be 1.33$\pm$0.10. We investigate potential reasons for this in the following section.

4.8 Checks and comparisons

In order to understand why we have an unexpectedly high M_X/M_C ratio obtained from the analysis of our sample of 44 clusters, we carried out a number of tests and comparisons to thoroughly investigate this unexpected result. In this section we present the results for the most relevant techniques that we used for this. In this section, when a line is fitted (as in §4.8.1 c), 4.8.2 and 4.8.4), all models are fitted in log space using BCES orthogonal regression (Akritas & Bershady, 1996), which allows for different errors in the x and y direction, and intrinsic scatter.

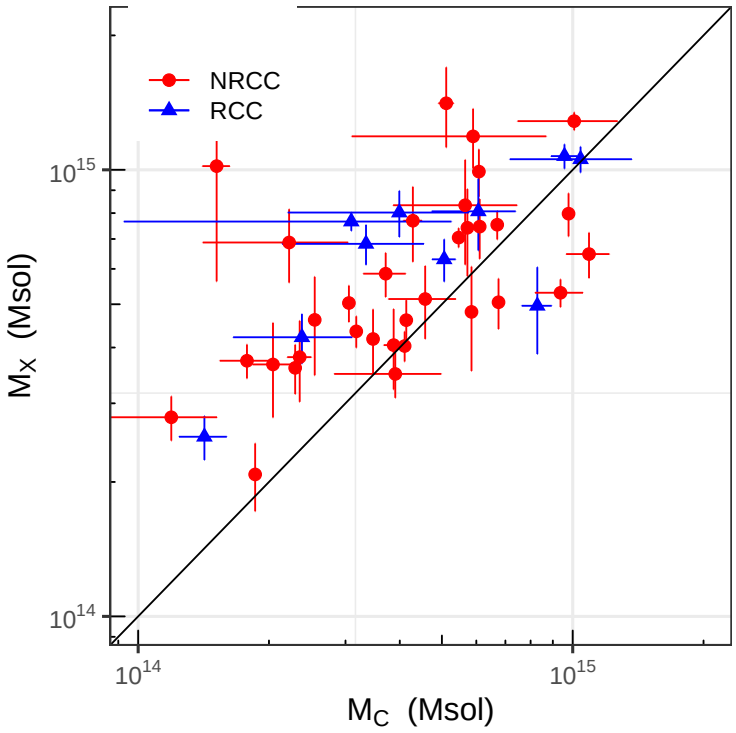

Figure 4.6: We show the hydrostatic masses versus the caustic masses for all 44 clusters in our sample, both calculated at the hydrostatic R_{500}. NRCC clusters are red circles and RCC clusters are blue triangles. The 1:1 line is also plotted as a solid black line. 1σ errors are shown.

Aperture	Subset	N_C	κ	M_X/M_C	δ (%)
R_{500}	All	44	0.123±0.031	$1.33^{+0.10}_{-0.09}$	39^{+06}_{-06}
R_{500}	RCC	10	0.126±0.102	$1.32^{+0.33}_{-0.24}$	51^{+25}_{-16}
R_{500}	NRCC	34	0.125±0.035	$1.33^{+0.11}_{-0.10}$	41^{+07}_{-06}
1 Mpc	All	44	0.122±0.031	$1.32^{+0.10}_{-0.09}$	40^{+07}_{-06}

Table 4.4: Summary of the results from our mass comparison analysis. The first column gives the aperture within which the caustic and hydrostatic masses were calculated. For this column, R_{500} refers to the hydrostatic R_{500}. The second column gives the subset for which the summary statistics are given, and the third column N_C gives the number of clusters in that subset. Columns 4, 5 and 6 give the summary statistics. The mean bias is given in column 4, the median ratio of masses is given in column 5, and the intrinsic scatter between the caustic and hydrostatic mass is given in column 6. See §4.6 for details on how these values were derived.

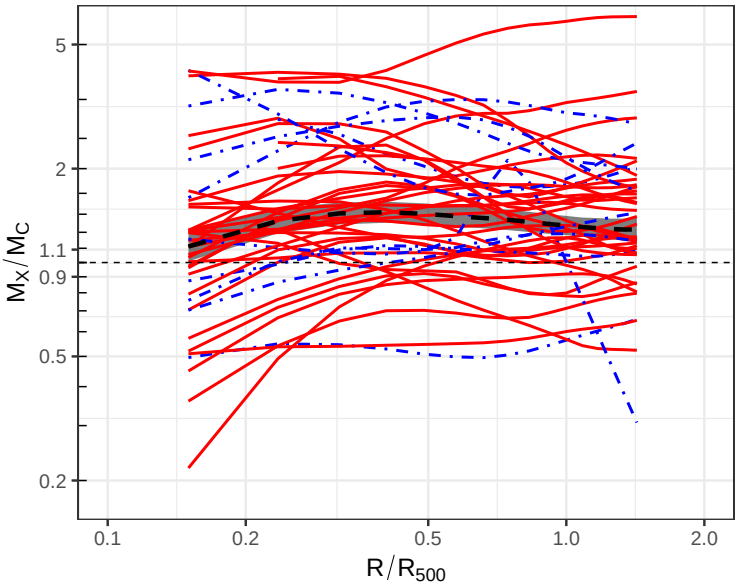

Figure 4.7: We show the M_X/M_C profiles for all 44 clusters in our sample, calculated at the hydrostatic R_{500}. NRCC clusters are solid red lines and RCC clusters are blue dot dashed lines. The 1:1 line is also plotted as a solid black line. The dashed black line shows the average M_X/M_C ratio, and the shaded region shows the 1σ uncertainty.

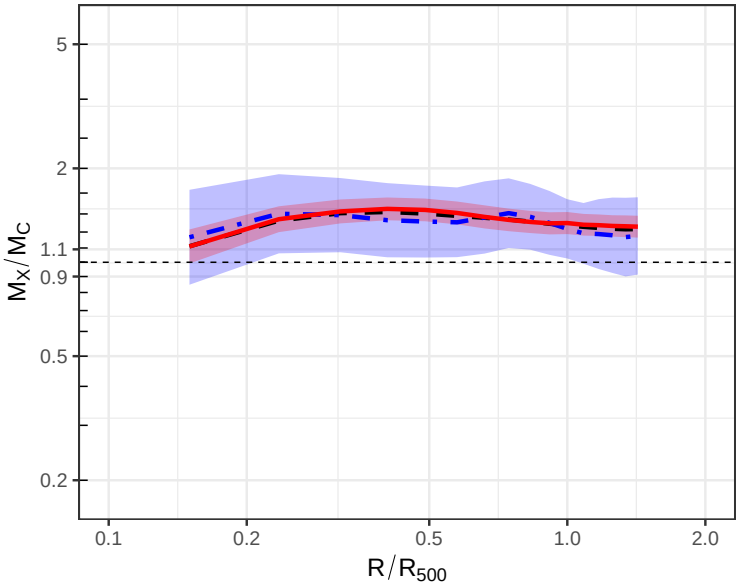

Figure 4.8: We show the average M_X/M_C profiles for all 44 clusters in our sample, for NRCC clusters as solid red lines and RCC clusters as blue dot-dashed lines. The red shaded region shows the 1σ uncertainty for the NRCC clusters, and the blue shaded region shows the same uncertainty for the RCC clusters.

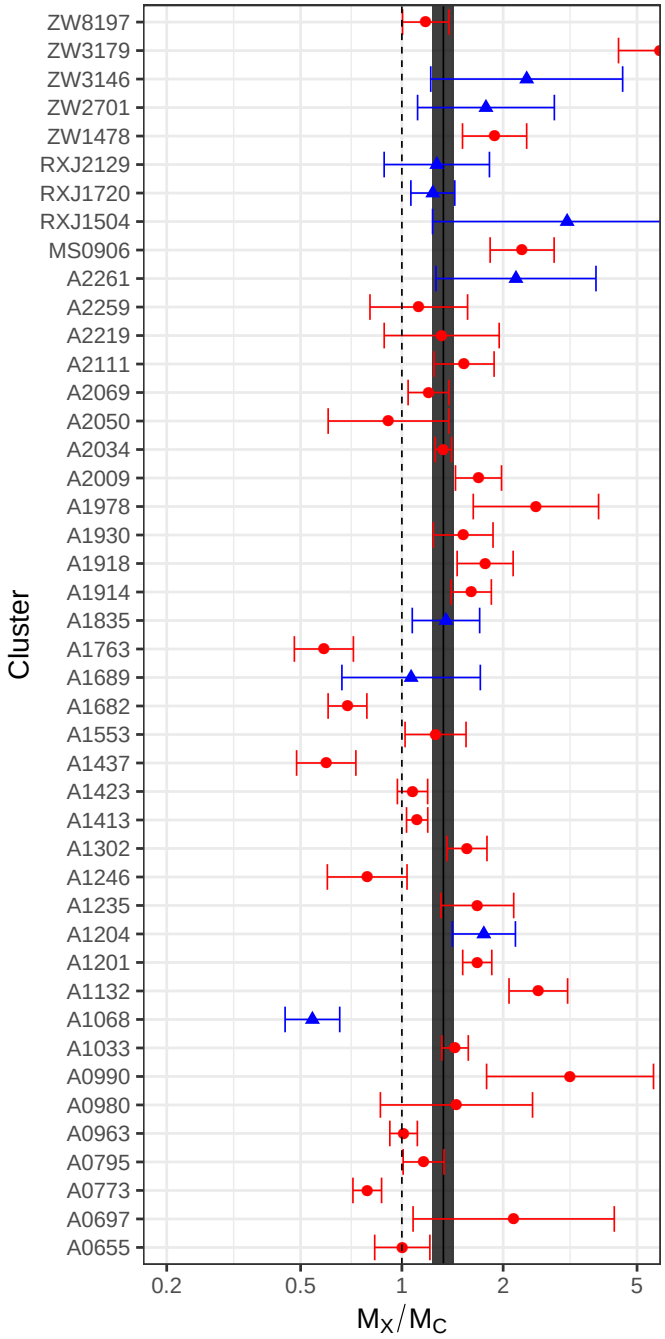

Figure 4.9: We show the M_X/M_C ratio for all 44 clusters in our sample. NRCC clusters are red circles and RCC clusters are blue triangles. The solid black line and shaded region show the average M_X/M_C ratio at the hydrostatic R_{500}. 1σ errors are shown.

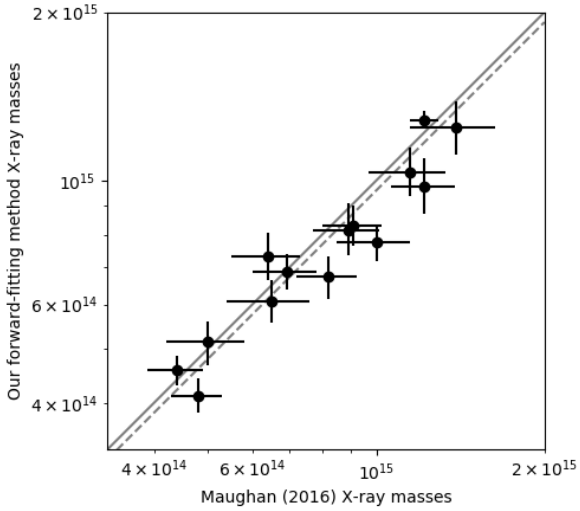

Figure 4.10: We compare our X-ray masses to the X-ray masses from Maughan et al. (2016) for the 14 clusters that we have in common. We re-calculated our masses using the Maughan et al. (2016) R_{500} values. The weighted average ratio of our masses over those from Maughan et al. (2016) is 0.96±0.04. The 1:1 line is shown as a solid grey line, and the grey dashed line shows a line with gradient 1 and intercept of 0.96 (the weighted mean ratio value).

We employ a variety of different methods to test for systematics in both the X-ray hydrostatic and caustic mass methods. In §4.8.3 we find the most compelling evidence for the reason behind the unexpectedly high M_X/M_C ratio, which is that the caustic masses are significantly underestimated for clusters with a low Ngal value (see Figure 4.25). Ngal is the number of galaxies of each cluster within the caustics defined in the caustic method, as given in Rines et al., 2013, and we show the Ngal values for each cluster in Table 4.1.

4.8.1 Comparison with other masses in the literature

One method to probe the accuracy of our X-ray cluster masses, and to investigate if there is an issue with them, is to compare them to any cluster samples with estimated masses in the literature that have a significant number of overlapping clusters with our sample. The results presented in this subsection show consistency with other X-ray analyses.

a) Comparison to Maughan et al. (2016)

One comparison we can make is for the clusters that our sample has in common with Maughan et al. (2016). As mentioned in §4.2, this work is an extension of Maughan et al. (2016), however, of the 16 clusters in Maughan et al. (2016), we have dropped two (A267 and A2631), due to reasons mentioned in §4.2. Thus, we can make a comparison of the 14 clusters in common between our samples. We plot this comparison in Figure 4.10, recomputing our masses at the R_{500} value from the Maughan et al. (2016) paper. We find a close agreement, with the weighted average ratio of our masses to the Maughan et al. (2016) masses being 0.96±0.04. The plot is suggestive of the fact that we may have systematically lower mass values than those from Maughan et al. (2016), although the mass ratio suggests it is not significant. This suggested decrease in mass between papers would lead to a lower M_X/M_C ratio in the present work when compared with Maughan et al. (2016), which is not the case (as we find a higher M_X/M_C ratio).

b) Comparison to other X-ray masses

We also compared our masses to 17 clusters in Martino et al. (2014) and 8 in Mahdavi et al. (2013), for which X-ray hydrostatic masses had been estimated using *Chandra* data. The comparison plots are shown in Figures 4.11 and 4.12 respectively. For consistency, we recomputed our masses in the R_{500} as measured in the other work. We compute a weighted mean ratio of our hydrostatic masses to their hydrostatic masses as our comparison method. For the 17 clusters in Martino et al. (2014), we obtain a weighted mean ratio of 0.98 ± 0.05 and for Mahdavi et al. (2013) we obtain 0.99 ± 0.07. This strongly suggests that we are not overestimating the hydrostatic mass in our hydrostatic mass estimation method. However, it is worth noting that the clusters in the comparison studies are predominantly high mass clusters (and therefore also clusters with likely higher Ngal values - see Figure 4.13, where it is clear that clusters with higher masses have higher Ngal values on average).

c) Comparison with Weak Lensing masses

We also compared our masses to 23 weak lensing masses from a catalogue from Sereno (2015) that was compiled by taking weak lensing mass measurements from multiple literature sources, and reporting masses at different overdensity radii, which if not reported in the original paper, were recalculated by Sereno (2015) at each overdensity radius, using the assumed mass profile

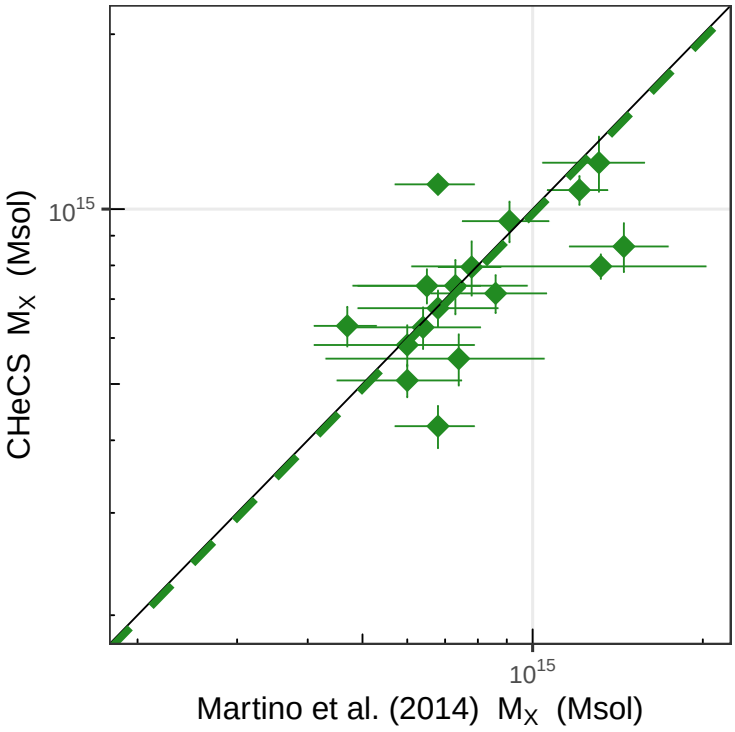

Figure 4.11: We compare our X-ray masses to the X-ray masses from Martino et al. (2014) for the 17 clusters that we have in common. We re-calculated our masses using the Martino et al. (2014) R_{500} values. The weighted mean ratio of our masses to theirs is 0.98±0.05. The 1:1 line is shown as a solid black line, and the green dashed line shows a line with gradient 1 and intercept of 0.98 (the weighted mean ratio value).

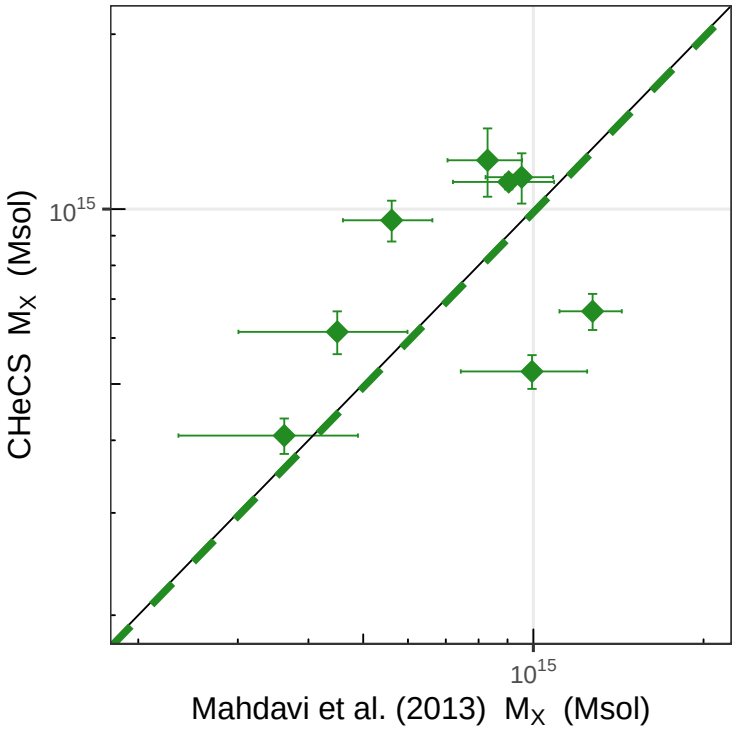

Figure 4.12: We compare our X-ray masses to the X-ray masses from Mahdavi et al. (2013) for the 8 clusters that we have in common. We re-calculated our masses using the Mahdavi et al. (2013) R_{500} values. The weighted mean ratio of our masses to theirs is 0.99 ± 0.07. The 1:1 line is shown as a solid black line, and the green dashed line shows a line with gradient 1 and intercept of 0.99 (the weighted mean ratio value).

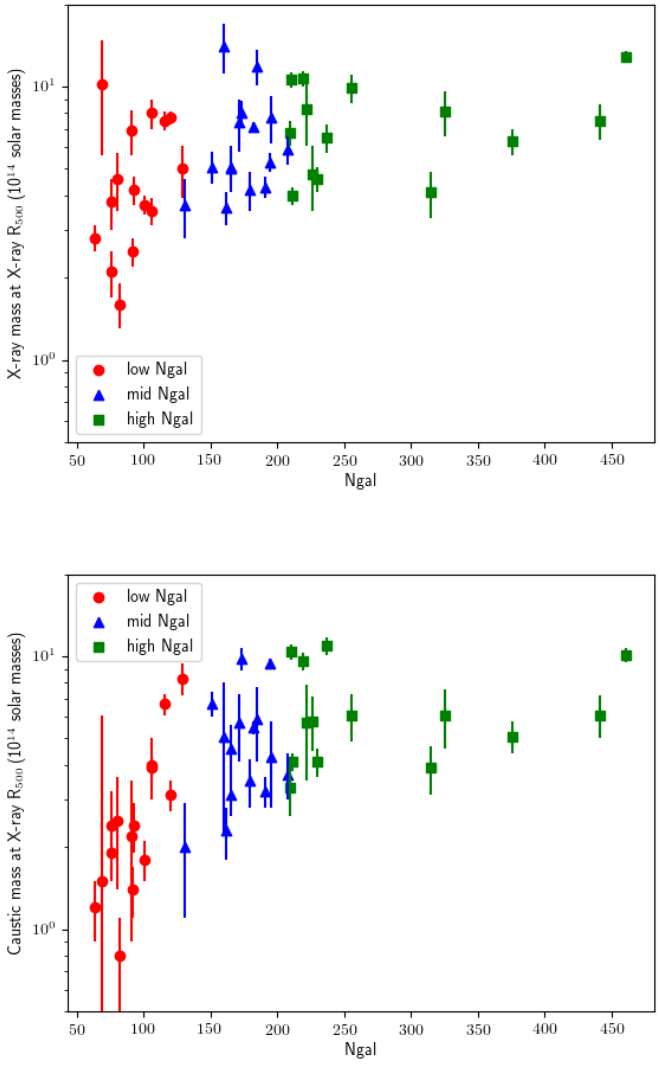

Figure 4.13: We show our X-ray (top panel) and caustic (bottom panel) masses as a function of Ngal. Both masses are calculated within the the X-ray R_{500}.

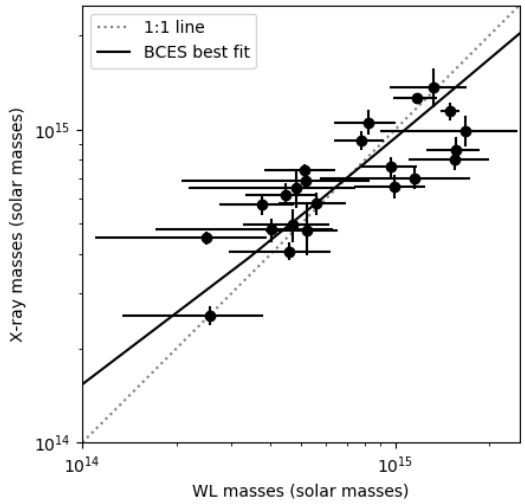

Figure 4.14: We compare our X-ray masses to the weak lensing masses from Sereno (2015) for the 23 clusters that we have in common. We re-calculated our masses using the Sereno (2015) R_{500} values. We plot a line to the data and find a gradient of 0.85±0.14, plotted as a solid black line.

from the appropriate analysis. We note that for the comparison of our X-ray masses with the weak lensing masses from Sereno (2015), we recompute our masses in the R_{500} value from Sereno (2015). We present the comparison plot in Figure 4.14. For this comparison, we fit a line to the data, as the comparison plot suggests that there may be a trend with mass. We find the best fit line to have a gradient of 0.85±0.14, almost consistent with a gradient of 1, which shows that there is no strong trend with mass (i.e. we are not overestimating low mass clusters and neither are we underestimating high mass clusters). We find the weighted average ratio to be 0.91±0.05, which does not suggest that we are over-estimating our X-ray masses.

4.8.2 Scaling relations

From our high signal-to-noise data we can easily measure global quantities such as luminosity and temperature. Additionally, we can easily measure core-excised luminosities and temperatures, which are measured within an annulus with inner radius 0.15 R_{500} and outer radius of R_{500}. Using these measured quantities, we can employ another method to probe the accuracy of our X-ray cluster masses, which is to look at X-ray scaling relations for our sample of 44 clusters, and compare our sample's results with those from literature. Using scaling relations

can help us understand mass discrepancy, as we will expect to see a certain relation between two quantities, and if we see a significant difference between our results and those from the literature, we can highlight, or gain insight into, any issues present in our X-ray analysis. We also look at scaling relations that have a quantity that is related to the caustic mass method as one of the variables, in order to investigate issues that may be present in the caustic mass method.

When comparing to literature values, we try to choose a reference analysis that has a cluster sample as similar as possible to ours, in terms of the cluster masses and redshifts. We also try to ensure that the reference analysis measurements have been obtained using X-ray data from *Chandra*, as otherwise differences that we see in our results could be due to calibration issues between observatories. The results presented in this subsection generally show consistency with other X-ray analyses.

In the following subsection we are fitting a line using BCES orthogonal regression (Akritas & Bershady, 1996) in log space to quantities that have the relation $Y = A \times (X/X_0)^B$, such that in log space the relation is $\log Y = \log A + B \log(X/X_0)$, where the gradient is B and the intercept is $\log A$ in the log space fit. For all cases below, we report the value of the gradient of the fit (B), the normalisation of the line ($\log A$), and the pivot point (X_0) whose value is chosen so as to be the same as that used in the reference analysis to facilitate comparison. We also plot our best fit line and the appropriate relation from literature as a comparison.

a) Luminosity-Temperature Relation

Here, the model that we fit is $L/E(z) = A \times (T/T_0)^B$ where the value of T_0 that we use is 6 keV. We plot the core-excised luminosity and core-excised temperature (both measured in the 0.15 - 1 R_{500} region) relation in Figure 4.15. We find the line of best fit to have a gradient of 2.71$\pm$0.17, in agreement with the gradient of 2.72$\pm$0.18 from Maughan et al. (2012), which also uses *Chandra* data, for the same relation. However, our measured intercept (converted to normal space) is 4.90$\pm$0.26 $\times$ 10^{44} compared to 6.98$\pm$0.30 $\times$ 10^{44} from Maughan et al. (2012). This offset is due partially to the fact that between the Maughan et al. (2012) analysis and our analysis, there have been significant changes in the calibration files used in the *Chandra* analysis, and also because in the Maughan et al. (2012) analysis, the chi-squared statistic is used for the cluster spectral fitting in contrast to the C-statistic that we use in our analysis (Duffy, private communication).

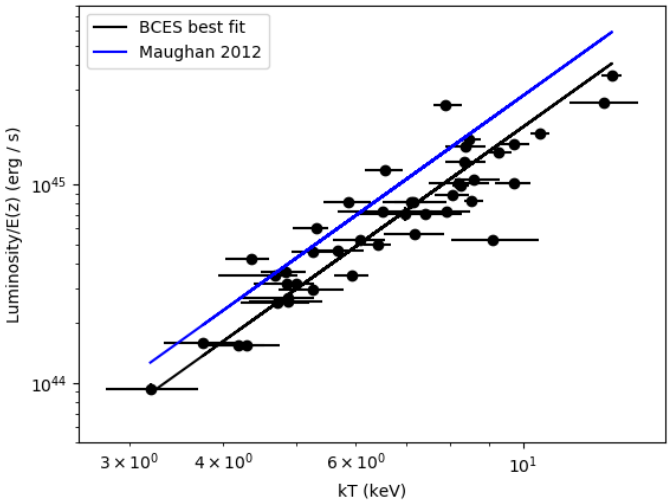

Figure 4.15: We show the core-excised L - T relation for the 44 clusters in our sample. The L and T values are measured within the X-ray 0.15 - 1 R_{500}. We plot the best fit line as a solid black line, and also plot the best fit line for the same relation from Maughan et al. (2012) with a solid blue line.

b) Mass-Temperature Relation

Here, the model that we fit is $M \times E(z) = A \times (T/T_0)^B$ where the value of T_0 that we use is 5 keV. We plot the X-ray M_{500} - temperature (M_X-T) relation in Figure 4.16, and the caustic M_{500} - temperature (M_C-T) relation in Figure 4.17. We also compare the two relations in Figure 4.18. The temperatures are measured within the 0.15 - 1 R_{500} region. The lower scatter in the X-ray plot is expected due to the lower scatter in the X-ray masses, and the use of the same X-ray data on both axes which introduces a covariance. We find the M_X-T relation to have a gradient of 1.48$\pm$0.11, and the M_C-T relation to have a gradient of 2.40$\pm$0.56. The M_X-T relation agrees well with the literature, which finds gradients of $\sim$ 1.5-1.6 (e.g. 1.58$\pm$0.11 from Vikhlinin et al., 2006, which also uses *Chandra* data). The M_X-T relation also has a measured intercept (converted to normal space) of 4.07$\pm$0.18 $\times$ 10^{14} compared to 4.12$\pm$0.16 $\times$ 10^{14} in Vikhlinin et al. (2006) . The offset between our results and those from Vikhlinin et al. (2006) is not significant. The M_C-T relation has a significantly steeper slope than the M_X-T relation, and the fact that the M_C-T relation is somewhat steeper than found in many other analyses (e.g. Hoekstra, 2007; Okabe et al., 2010; Kettula et al., 2013; Lieu et al., 2016), supports the possibility that the low temperature (and therefore likely low mass and therefore likely lower

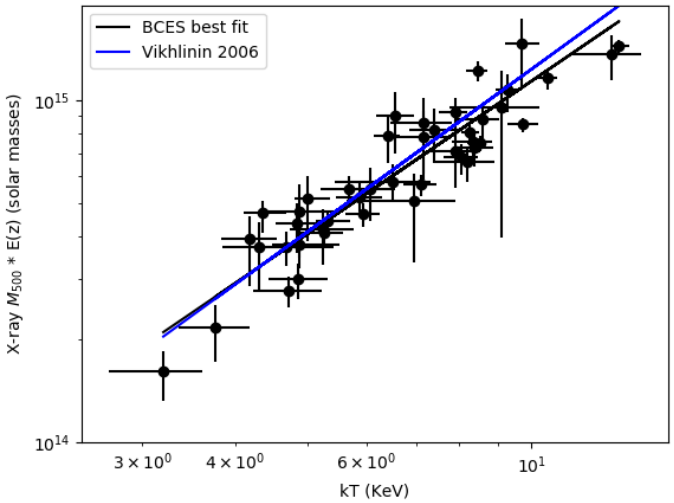

Figure 4.16: We show the M_X - T relation for the 44 clusters in our sample. The M_X values are measured within the X-ray R_{500}, and the T values are measured within the X-ray 0.15 - 1 R_{500}. We plot the best fit line as a solid black line, and also plot the best fit line for the same relation from Vikhlinin et al. (2006) with a solid blue line.

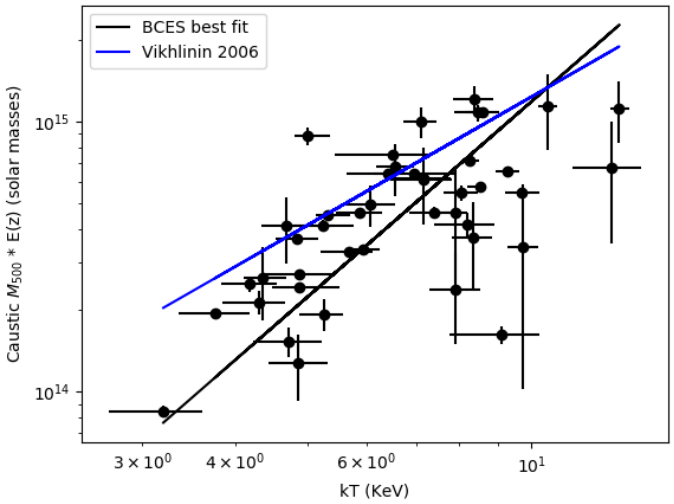

Figure 4.17: We show the M_C - T relation for the 44 clusters in our sample. The M_C values are measured within the caustic R_{500} and the T values are measured within the X-ray 0.15 - 1 R_{500}. We plot the best fit line as a solid black line, and also plot the best fit line for the same relation from Vikhlinin et al. (2006) with a solid blue line.

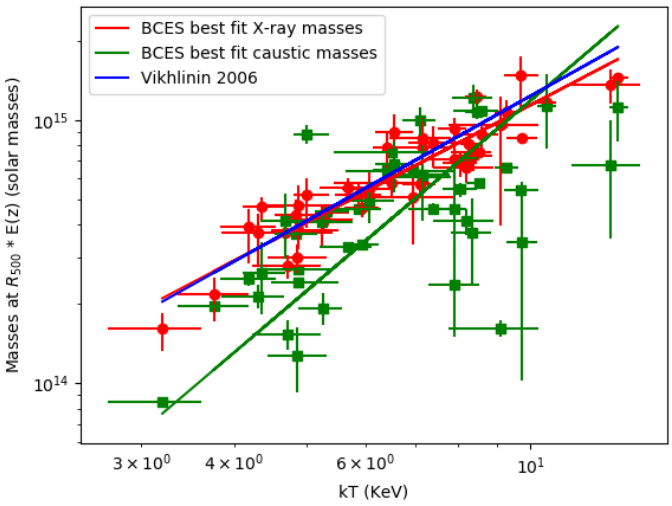

Figure 4.18: We show the M_X - T and the M_C - T relation for the 44 clusters in our sample. The M_X values are measured within the X-ray R_{500}, the M_C values are measured within the caustic R_{500}, and the T values are measured within the X-ray 0.15 - 1 R_{500}. The data points for the M_X - T relation are red circles, and the corresponding best fit line is a solid red line. The data points for the M_X - T relation are green squares, and the corresponding best fit line is a solid green line. We also plot the best fit line for the M_X - T relation from Vikhlinin et al. (2006) with a solid blue line.

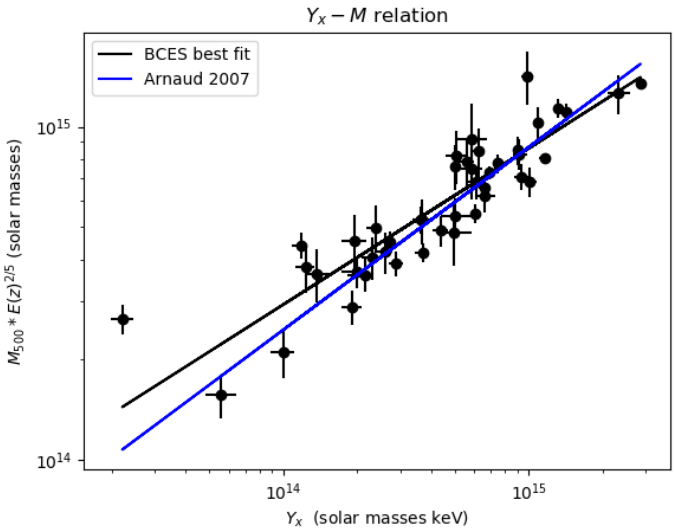

Figure 4.19: We show the M - Y_X relation for the 44 clusters in our sample, where Y_X is $m_{\rm gas} \times T$. The M and $m_{\rm gas}$ values are measured within the X-ray R_{500}, and the T values within the 0.15 - 1 R_{500} region. We plot the best fit line as a solid black line, and also plot the best fit line for the same relation from Arnaud et al. (2007) with a solid blue line.

Ngal, see Figure 4.13) clusters have their mass underestimated in the caustic method. This is the first suggestion that there may be some systematic issue with the caustic masses.

c) M-Y_X Relation

Here, the model that we fit is $M \times E(z)^{2/5} = A \times (Y_X/Y_{X,0})^B$ where $Y_{X,0}$ has the value 2 $\times$ 10^{14} $M_\odot$. We plot the M-Y_X relation (where Y_X = $m_{\rm gas} \times T$) in Figure 4.19, and find the gradient to be 0.469$\pm$0.052, which is consistent with the value of 0.548$\pm$0.027 from Arnaud et al. (2007). The M-Y_X relation has a measured intercept of 14.610$\pm$0.027, which is similar to that from Arnaud et al. (2007) of 14.556$\pm$0.015. The temperature values for the calculation of Y_X were calculated in the 0.15 - 1 R_{500} region, and $m_{\rm gas}$ was measured in the R_{500} region. We note that the results from Arnaud et al. (2007) are derived from *XMM-Newton* data, and so there may be some difference between our results due to calibration issues between *Chandra* and *XMM-Newton*.

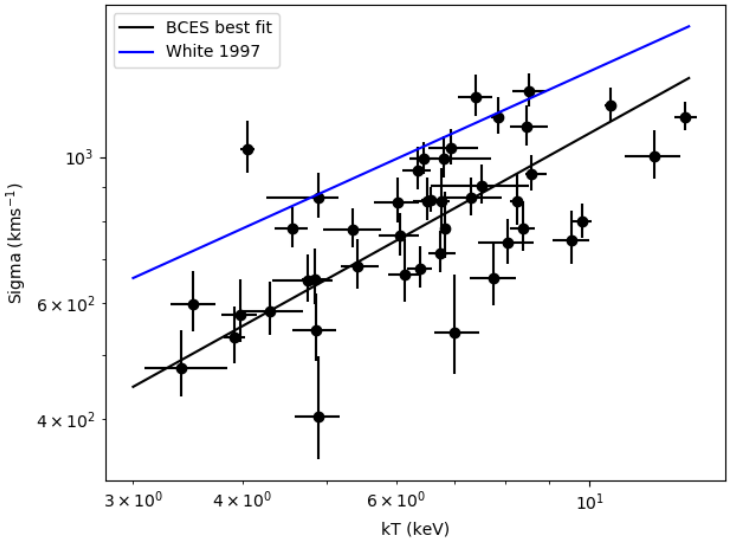

Figure 4.20: We show the σ - T relation for the 44 clusters in our sample, where σ is the velocity dispersion of the galaxies in each cluster. The T values are measured within the X-ray R_{500}. We plot the best fit line as a solid black line, and also plot the best fit line for the same relation from White et al. (1997) with a solid blue line. We note that there is a high normalisation of the sigma-T relation from the literature as the literature relations are generated by overestimated sigma values calculated from < 100 member galaxies, and is not related to the X-ray data.

d) σ-T Relation

Here, the model that we fit is $\sigma \times E(z) = A \times T^B$ (where σ refers to the statistical velocity dispersion about the mean for the member galaxies of a cluster). We investigate this velocity dispersion-temperature ($\sigma - T$) relation using the temperature within a circle of radius R_{500} that we measured in our X-ray analysis and σ from Rines et al. (2013). Figure 4.20 shows a plot of the $\sigma - T$ relation and the best fit line to our points which has a gradient of 0.74$\pm$0.14; in addition, we also plot the best fit line from the data in White et al. (1997) which has a gradient of 0.60$\pm$0.10. The gradient from our data and that from White et al. (1997) are consistent within error, however our $\sigma - T$ relation has a measured intercept of 2.30$\pm$0.11 which is significantly different to the 2.54$\pm$0.08 from White et al. (1997). This offset arises as, in general, our measured temperatures are either too high for a given σ, or alternatively, the σ values are too low for a given temperature. If the former is the case (i.e. we are overestimating the temperatures for our clusters) this would lead to a systematic overestimation of the cluster mass for the low mass clusters using the X-ray hydrostatic method. If the latter is the case, (i.e. the σ values are underestimated for a given temperature) we presume this would lead to the caustic masses being underestimated. From the $\sigma - T$ relation it is not clear where the systematic lies, but we do note that the $f_{\rm gas}$ measurement (presented in §4.8.3) is largely independent of temperature, so a systematic in the temperature measurement would not explain the anomalous gas fractions when caustic masses are used. We also note that the sigma-T relations from the literature are generated using sigma values derived from low numbers ($<$100) of members galaxies, and so are overestimated, leading to a high normalisation of the sigma-T relation from literature. We note that the results from White et al. (1997) are derived from *Einstein* data, and so there may be some difference between our results due to calibration issues between *Chandra* and *Einstein*, however, the results from White et al. (1997) are typical of other more recent results from literature (see Wilson et al., 2016).

4.8.3 $f_{\rm gas}$ comparisons

Another method to probe the accuracy of our X-ray and caustic masses is to calculate the gas fraction, f_{gas}. $f_{\rm gas}$ is calculated as the gas mass divided by the total cluster mass, and as it is well known that we can measure the mass of the ICM accurately from X-ray observations (e.g. Nagai et al., 2007), we can calculate the $f_{\rm gas}$ values with either the X-ray or caustic mass in the denominator to test the accuracy of these cluster mass measurements. We can use the $f_{\rm gas}$ values

for each cluster to test for a systematic in the mass estimation methods that is dependent on a certain quantity by splitting the clusters into sub-samples based on whatever quantity we choose to investigate.

We choose to split the cluster sample into three roughly evenly sized subsamples based on their Ngal value, to test for a systematic that is dependent on this quantity. The reason that we choose to investigate a potential systematic that is dependent on Ngal, is because the results from the $M_C - T$ relation (see Figure 4.17) suggest that there is an issue with the caustic masses for lower temperature (and therefore likely lower Ngal) clusters. Additionally, Serra et al. (2011) and Wojtak et al. (2018) both find that the caustic mass bias depends on Ngal (discussed in more detail later in this section). We plot the $f_{\rm gas}$ values at 1 Mpc for all 44 clusters with caustic mass as the denominator in Figure 4.21 as a function of Ngal, and do the same, but with the X-ray mass as the denominator, in Figure 4.22. We expect that the $f_{\rm gas}$ values will be $\sim$ 0.1 - 0.15 at this radius as this is roughly the value of the cosmic baryon fraction (e.g. Eckert et al., 2016). We see a clear trend with Ngal when the caustic mass is used as the denominator (Figure 4.21), suggesting that the clusters with low Ngal have their caustic masses significantly underestimated. This is especially surprising as the caustic mass is expected to be *overestimated* by $\sim$ 20% at this radius (Serra et al., 2011). We do not see this trend when the X-ray mass is used in the denominator (Figure 4.22).

We split the clusters into subsamples based on their Ngal value into subsamples of 15, 15 and 14 clusters for the low, mid and high Ngal subsamples respectively (see Ngal subsample column in Table 4.3; the low, mid and high mass subsamples have ranges of their Ngal values as 63 - 129, 131 - 208, and 210 - 461 respectively, and their means are 93, 174, and 281 respectively). Doing this we find the mean and standard error on the mean for the $f_{\rm gas}$ values when calculated using the caustic mass in the denominator are 0.231$\pm$0.027, 0.151$\pm$0.010, 0.148$\pm$0.015, for the low, mid and high Ngal subsamples respectively and 0.118$\pm$0.006, 0.112$\pm$0.004, 0.124$\pm$0.007, for the low, mid and high Ngal subsamples respectively for when the X-ray mass is used as the denominator. The standard error on the mean will be an underestimate as it does not include the uncertainties on the $f_{\rm gas}$ values, however, there is a clear difference in the $f_{\rm gas}$ values between the Ngal subsamples, with the lowest Ngal bin significantly exceeding the cosmic baryon fraction.

Given that the caustic method is not optimised for smaller radii, and that gas masses can be measured reliably to relatively large radii ($\sim$2/3 of our 44 clusters have $m_{\rm gas}$ profiles constrained out to 1.8 Mpc), we extended this analysis to measure $f_{\rm gas}$ at 1.8 Mpc, which equals or exceeds

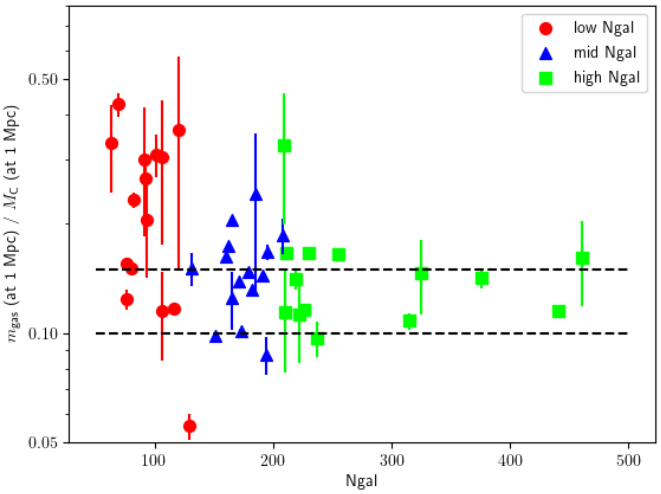

Figure 4.21: We plot the f_{gas} values at 1 Mpc as a function of Ngal using the caustic mass in the denominator for the f_{gas} value. Splitting the 44 clusters into into roughly evenly sized (15, 15, 14 members) of low, mid, high Ngal subsamples, we find the f_{gas} values to be (with standard error): 0.23±0.03, 0.15±0.01, 0.15±0.01 respectively. The black dotted lines show f_{gas} = 0.10 and 0.15, which is where it is expected that the f_{gas} values would lie.

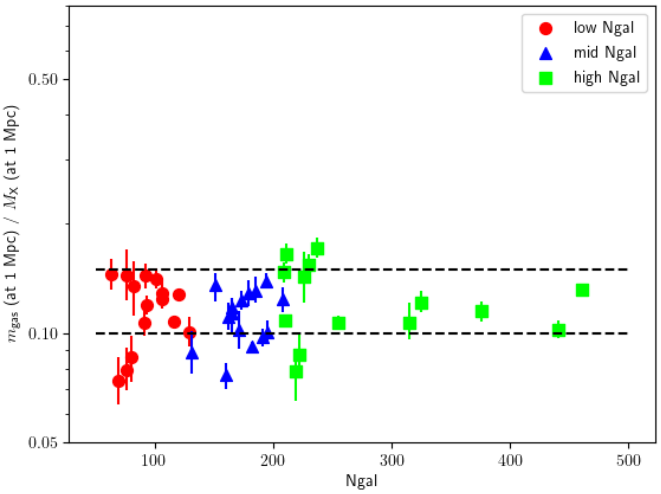

Figure 4.22: We plot the f_{gas} values at 1 Mpc as a function of Ngal using the X-ray mass in the denominator for the f_{gas} value. Splitting the 44 clusters into into roughly evenly sized (15, 15, 14 members) of low, mid, high Ngal subsamples, we find the f_{gas} values to be (with standard error): 0.12±0.01, 0.11±0.01, 0.12±0.01. The black dotted lines show f_{gas} = 0.10 and 0.15, which is where it is expected that the f_{gas} values would lie.

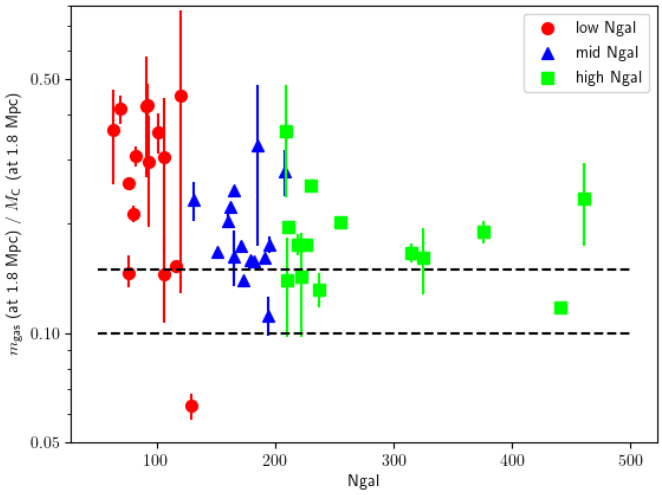

Figure 4.23: We plot the f_{gas} values at 1.8 Mpc as a function of Ngal using the caustic mass in the denominator for the f_{gas} value. Splitting the 44 clusters into into roughly evenly sized (15, 15, 14 members) of low, mid, high Ngal subsamples, we find the f_{gas} values to be (with standard error): 0.29±0.03, 0.19±0.01, 0.19±0.02. The black dotted lines show f_{gas} = 0.10 and 0.15, which is where it is expected that the f_{gas} values would lie.

r_{200} for most of the clusters (from Table 4 in Rines et al., 2013) and find a very similar trend, with the f_{gas} values for the lower Ngal mass clusters being significantly higher than expected. We plot the f_{gas} values at 1.8 Mpc for all 44 clusters with caustic mass as the denominator in Figure 4.23 as a function of Ngal, and do the same but with the X-ray mass as the denominator in Figure 4.24. We find the mean and standard error on the mean for the f_{gas} values when calculated using the caustic mass in the denominator are 0.29±0.03, 0.19±0.01, 0.19±0.02, for the low, mid and high Ngal subsamples respectively and 0.15±0.01, 0.14±0.01, 0.17±0.02, for the low, mid and high Ngal subsamples respectively for when the X-ray mass is used as the denominator. We note that in Figure 4.24, most of the X-ray hydrostatic mass profiles are extrapolated at 1.8 Mpc as the temperature profile is not measured out to this radius. This partly explains the very high f_{gas} values of two clusters in the high Ngal subsample, A1835 and A963, which are due to the decreasing and flattening of the hydrostatic mass profiles of clusters A1835 and A963 respectively (see Figure 4.5), though we note the mass profile of A1835 starts to decrease unphysically before the outer radius of the measured temperature profile (as mentioned in §4.7).

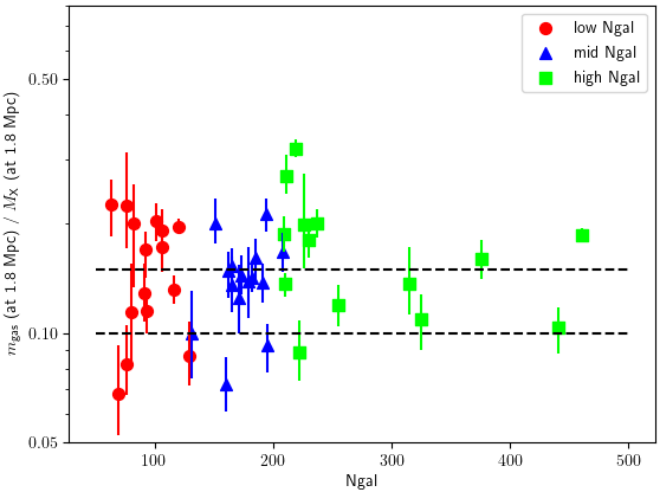

Figure 4.24: We plot the $f_{\rm gas}$ values at 1.8 Mpc as a function of Ngal using the X-ray mass in the denominator for the $f_{\rm gas}$ value. Splitting the 44 clusters into roughly evenly sized (15, 15, 14 members) of low, mid, high Ngal subsamples, we find the $f_{\rm gas}$ values to be (with standard error): 0.15±0.01, 0.14±0.01, 0.17±0.02. The black dotted lines show $f_{\rm gas}$ = 0.10 and 0.15, which is where it is expected that the $f_{\rm gas}$ values would lie.

There is some evidence that at large radii ($\sim r_{200}$), $m_{\rm gas}$ may be overestimated due to clumping of the X-ray emitting gas (see e.g. Figure 4 Simionescu et al., 2011), which could contribute to the fact that on average the $f_{\rm gas}$ values at 1.8 Mpc are higher than those at 1 Mpc. However, the impact of clumping is expected to be smaller than the discrepancy we see here, and given that the same trend is seen at 1 Mpc (where the impact of clumping is negligible), we don't consider this to be a likely explanation for the high $f_{\rm gas}$ values.

We note that the gas mass values derived using the forward-fitting method (that we have used for these $f_{\rm gas}$ tests) and those derived from the backward-fitting method are very similar (the weighted average ratio of the gas mass from the forward-fitting method divided by that from the backward-fitting method is 0.930±0.001 at 1 Mpc), and thus we get very similar results as those above if we use the gas mass as calculated from the backward-fitting method.

From these $f_{\rm gas}$ comparisons, we have a clear indication that the caustic masses are significantly underestimated for those clusters with a low Ngal value, and show the M_X/M_C ratio (with both mass estimates measured at the X-ray R_{500}) versus Ngal in Figure 4.25, which also shows the same strong trend with Ngal, suggesting that the caustic masses are significantly underestimated for clusters with low Ngal values. In fact, this is somewhat expected, as Serra et al.

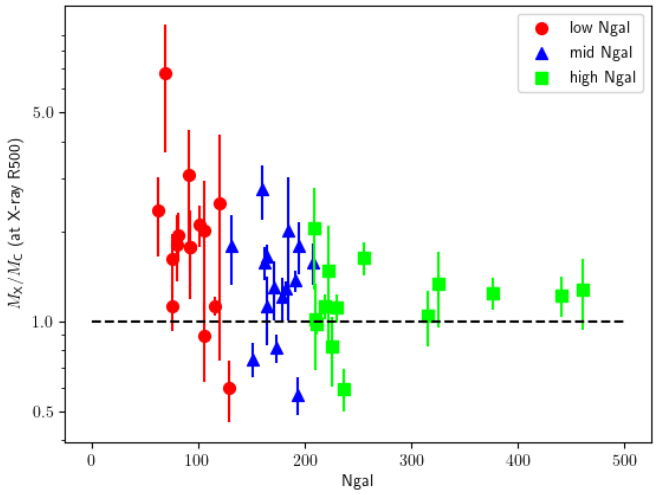

Figure 4.25: We plot the M_X/M_C ratio (using the masses calculated at the X-ray R_{500}) as a function Ngal. The black dotted line shows where M_X/M_C = 1.

(2011) find that the caustic mass method does underestimate the cluster mass when the number of galaxies with measured redshift in the central cluster region is not large enough (specifically see Figure 20 in that paper). However, despite this, the expected underestimation of cluster masses (from Serra et al., 2011) for clusters with low Ngal is not strong enough to account for the very high M_X/M_C values we are obtaining in the low Ngal subsample. This could either be due to the fact that the results from Serra et al. (2011) underestimate the effect of the value of Ngal on the bias in the caustic mass method; alternatively, there could be another effect leading to the underestimation of caustic masses, or indeed overestimation of the X-ray masses in this subsample. We note that the results from Serra et al. (2011) are based on results from simulations, and no work has yet been done using observational data to understand the trend with Ngal of the bias in the caustic mass. Additionally, Wojtak et al. (2018) find that higher incompleteness in the sampling of the galaxies for each cluster leads to underestimating the mass (specifically see Figure 1 'ESC' subplot in that paper).

4.8.4 Comparison of forward-fitting and backward-fitting masses

In order to explore another source of possible systematics in our X-ray masses, we compare the masses calculated using our forward-fitting method (described in §3.2) to those from the

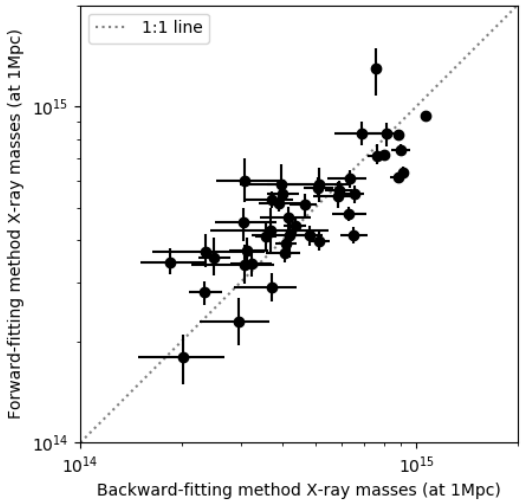

Figure 4.26: We plot the ratio of the forward-fitting masses over the backward-fitting masses at 1Mpc, and find this ratio to have a weighted average of 0.89±0.01. If we split the clusters into two evenly sized subsamples based on their backward-fitting mass value, we find a weighted average ratio of 1.15±0.04 for the low backward-fitting mass subsample and 0.85±0.01 for the weighted average ratio for the high backward-fitting mass subsample.

backward-fitting method (described in §3.3). Figure 4.26 shows the forward-fitting masses versus the backward-fitting masses at 1 Mpc. We see that they agree well, with the weighted average ratio of the forward-fitting mass over the backward-fitting being 0.89±0.01. However, the forward-fitting method tends to gives higher mass estimates for the low mass clusters compared to the backward-fitting method, and the forward-fitting method tends to give lower mass estimates than the backward-fitting method for the high mass clusters in our sample. To quantify this trend, we fit a line to the data in Figure 4.26 and find the best fit line to have a gradient of 0.89±0.11 which confirms the trend, but is also consistent with a 1:1 relationship. Splitting the sample into two subsamples based on their backward-fitting mass, we find the weighted average ratio of the forward-fitting masses divided by the backward-fitting masses for the low mass subsample to be 1.15±0.04 and 0.85±0.01 for the high mass subsample.

As lower mass clusters typically have lower Ngal values (as the more massive, larger clusters typically contain more member galaxies; also see Figure 4.13), and so the potential overestimation of the forward-fitting masses (compared to the backward-fitting masses) for the lower mass clusters that we see in Figure 4.26, could also imply that our forward-fitting X-ray method

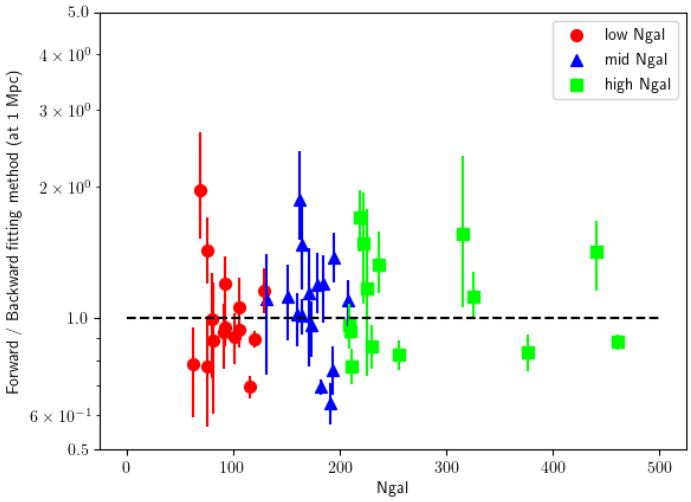

Figure 4.27: Now splitting the ratio of the forward-fitting masses over the backward-fitting masses at 1 Mpc as a function of Ngal, and splitting the 44 clusters into roughly evenly sized (15, 15, 14 members) of low, mid, high Ngal subsamples we get the weighted average ratio for the low, mid and high subsamples to be: 0.90±0.02, 0.85±0.02, 0.93±0.02. The black dotted line shows where the mass ratio is equal to 1.

overestimates the masses of clusters with lower Ngal. However, despite this, we note that it is not sufficient to *fully* explain the strong dependence of the M_X/M_C values on Ngal that we see in Figure 4.25 and could not explain the trend with Ngal for the $f_{\rm gas}$ values in Figure 4.21 (as the X-ray masses are not used in this plot). In Figure 4.27 we plot the forward-fitting mass divided by the backward-fitting mass, and we split the clusters by Ngal into three subsamples as in §4.8.3, finding the weighted average ratio of the forward-fitting mass divided by the backward-fitting mass to be 0.90±0.02, 0.85±0.02, and 0.93±0.02 for the low, mid and high Ngal subsamples respectively at 1 Mpc.

4.8.5 Final sample

We have employed a variety of tests and techniques to investigate the unexpectedly high M_X/M_C ratio that we found for our sample of 44 clusters, probing both the X-ray mass method and the caustic mass method that we used to estimate cluster masses. We find the strongest evidence for a systematic in §4.8.3, which suggests that the caustic mass method underestimates the cluster mass significantly for clusters with low Ngal values. We thus conclude that there may be a

Aperture	Subset	N_C	κ	M_X/M_C	δ (%)
R_{500}	High	15	0.050±0.043	$1.12^{+0.11}_{-0.10}$	25^{+12}_{-08}
R_{500}	Mid	15	0.109±0.053	$1.28^{+0.16}_{-0.14}$	41^{+12}_{-09}
R_{500}	Low	14	0.230±0.082	$1.69^{+0.33}_{-0.27}$	59^{+18}_{-13}

Table 4.5: Summary of the results from our mass comparison analysis when we define subsets of our clusters based on the Ngal value. The columns are the same as in Table 4.4.

systematic in the caustic masses.

For further studies using this sample (Chapter 5), we use just the 14 clusters in the high Ngal subsample (those labelled 'high' in the Ngal subsample column in Table 4.3) for the primary analysis, as the tests in this section suggest that we are 'safely' above the regime where cluster masses are underestimated by the caustic method due to their low Ngal value.

We note that for these 14 clusters, the median radius out to which the temperature profiles were measured is significantly higher than that for the 44 cluster sample. The median radius is 1.29 R_{500} (compared to 1.01 R_{500} for the 44 cluster sample) and the range is 0.67 - 1.62 R_{500}. This means that the extrapolation of the hydrostatic mass profiles, for which the best fitting 3D temperature profiles are needed, starts at a larger radius, on average.

4.9 Final Results

4.9.1 Our 44 cluster sample in subsamples split by Ngal value

We repeat our M_X/M_C profile analysis as detailed in §4.7 on the 44 clusters in our sample, this time splitting the clusters into subsamples based on their Ngal value with subsamples of 15, 15 and 14 clusters for the low, mid and high Ngal subsamples respectively.

We compare the caustic and X-ray hydrostatic masses in Figure 4.28 (as we did in Figure 4.6, but now in subsamples based on Ngal value). We show the mean M_X/M_C profiles of the 44 clusters in Ngal subsamples separately in Figure 4.29. These profiles do not agree with each other, with the M_X/M_C ratios (given in Table 4.5) in the low, mid and high Ngal subsamples being 1.69±0.30, 1.28±0.15 and 1.12±0.10 respectively. These results confirm our findings from §4.8, that the caustic masses for clusters with low Ngal appear to be systematically underestimated, leading to artificially high M_X/M_C values for these low Ngal clusters.

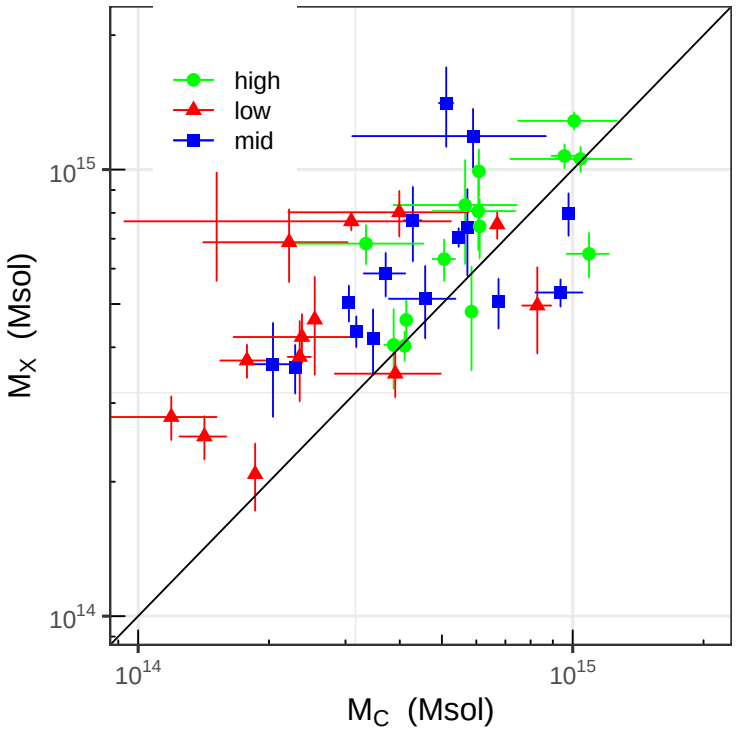

Figure 4.28: We show the hydrostatic masses versus the caustic masses for all 44 clusters in our sample, both calculated at the hydrostatic R_{500}, and divided into subsamples by their Ngal value. The high, mid and low Ngal bins are green squares, blue triangles and red circles respectively. The 1:1 line is also plotted as a solid black line. 1σ errors are shown.

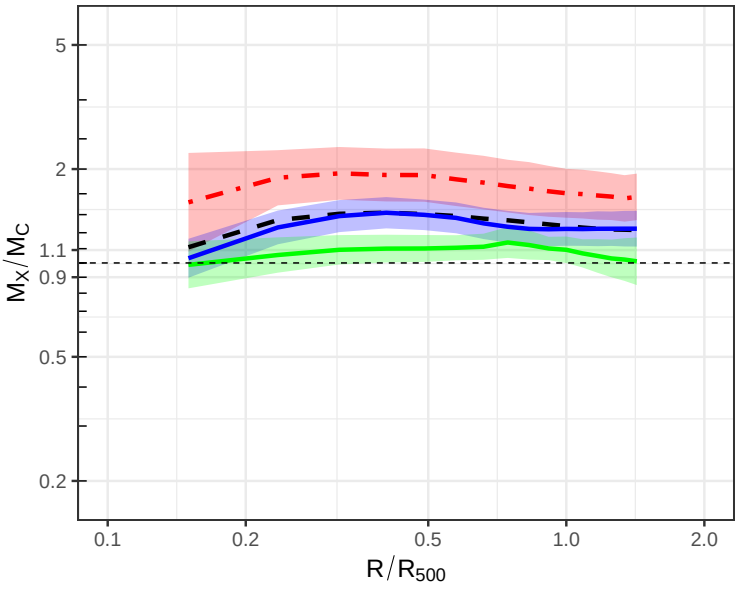

Figure 4.29: We show the M_X/M_C profiles for all 44 clusters in our sample, calculated at the hydrostatic R_{500}, for the high Ngal, mid Ngal and low Ngal clusters in green lines, blue lines, and red dot-dashed lines respectively. The 1:1 line is also plotted as a solid black line. The dashed black line shows the average M_X/M_C ratio for all clusters, and the shaded region shows the 1σ uncertainty.

Aperture	Subset	N_C	κ	M_X/M_C	δ (%)
R_{500}	All	14	0.050±0.043	$1.12^{+0.11}_{-0.10}$	25^{+12}_{-08}
R_{500}	RCC	5	0.138±0.150	$1.38^{+0.32}_{-0.27}$	33^{+54}_{-17}
R_{500}	NRCC	9	0.024±0.062	$1.06^{+0.14}_{-0.13}$	32^{+16}_{-10}
1 Mpc	All	14	0.052±0.044	$1.13^{+0.11}_{-0.10}$	28^{+12}_{-09}

Table 4.6: Summary of the results from our mass comparison analysis when we include just the 14 clusters with the highest Ngal values. The columns are the same as in Table 4.4.

4.9.2 14 clusters with highest Ngal

We also repeated our M_X/M_C profile analysis with just the 14 clusters in the high Ngal subsample (Ngal $\geq$ 210, see Table 4.3), and split them according to their dynamical state into RCC and NRCC clusters, summarizing our results in Table 4.6. In this case, the NRCC clusters (9 in total) give a M_X/M_C ratio of 1.06±0.14 and the RCC clusters (5 in total) give a ratio of 1.38±0.30 - the constraints are very weak, but are consistent with each other and this is true for all radii (see Figure 4.31). These results agree with those from §4.7.

Inspecting Figure 4.31 closely, at ~ 0.8 R_{500}, there is an increase in the M_X/M_C ratio for the RCC clusters. This is due to the inclusion of the cluster A1835 in the RCC subsample, and the effect is due to the shape of A1835's hydrostatic mass profile (see Figure 4.5; the reason for the unphysical decrease is discussed in §4.7), being especially strong due to the small size of the RCC subsample (5 clusters). Thus, we repeat our M_X/M_C profile analysis yet again on the 14 clusters in the high Ngal bin, this time removing A1835 from the RCC subsample such that we only have 13 clusters in the analysis, and this completely removes this 'bump' from the average M_X/M_C profile (see Figure 4.32). In this case we obtain 1.32±0.65 for the M_X/M_C ratio at the hydrostatic R_{500} for the RCC clusters, still consistent with the results for the NRCC clusters.

4.10 Discussion of results

We have compared the X-ray hydrostatic and caustic mass profiles of a sample of 44 clusters. The hydrostatic mass method uses measurements of the ICM to estimate the cluster mass, whereas the caustic method uses information on the dynamics of the member galaxies. Thus, the assumptions, and therefore systematics, in both methods are completely independent, which

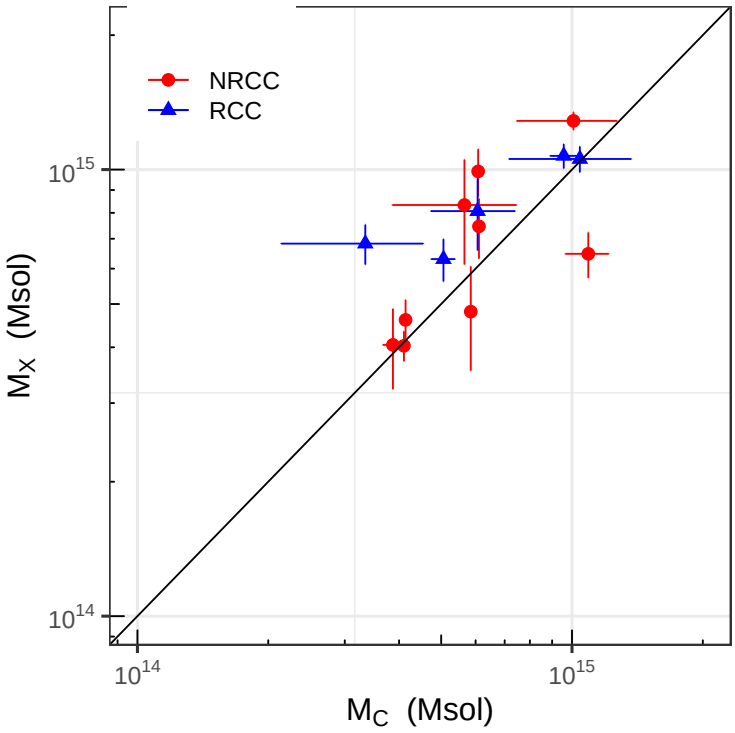

Figure 4.30: We show the hydrostatic masses versus the caustic masses for the 14 clusters in our sample with the highest Ngal values, both calculated at the hydrostatic R_{500}. We show RCC clusters as blue triangles and NRCC clusters as red circles. The 1:1 line is also plotted as a solid black line. 1σ errors are shown.

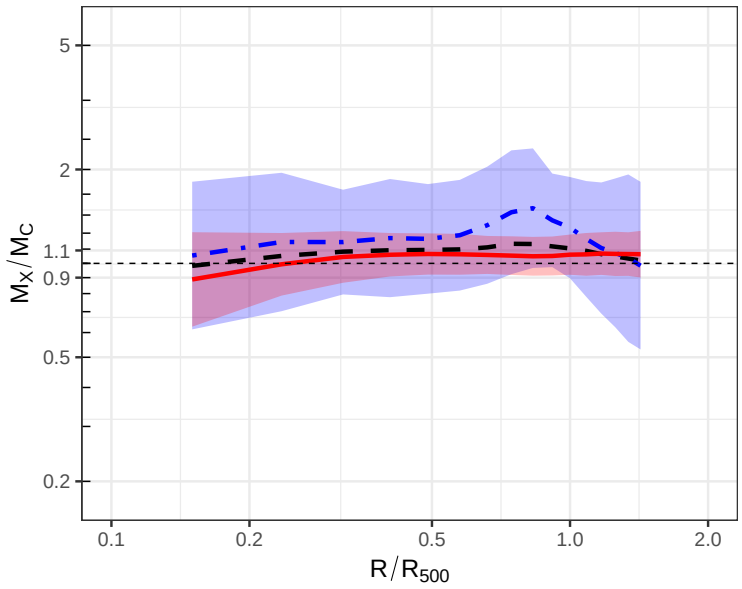

Figure 4.31: We show the average M_X/M_C profiles for the 14 clusters with the highest Ngal values in our sample, for RCC clusters as the blue dot-dashed lines and NRCC clusters as the red solid line. The blue shaded region shows the 1σ uncertainty for the RCC clusters, and the red shaded region shows the same uncertainty for the NRCC clusters.

makes for an ideal comparison. It was expected that the hydrostatic to caustic mass ratio, M_X/M_C, would be significantly less than 1 at the main comparison radius that we used of R_{500}, as the hydrostatic mass method is expected to underestimate the cluster mass by $\sim$10 - 30%, and the caustic mass is expected to overestimate the cluster mass by $\sim$10 - 20% at R_{500} (Serra et al., 2011). However, for the 44 clusters in our sample we found the M_X/M_C ratio at R_{500} to be 1.33$\pm$0.10, which was suggestive of another, unexpected systematic in either the X-ray hydrostatic masses, or the caustic masses.

We then used a variety of methods to test for this potential unexpected systematic, in both the hydrostatic and caustic mass measurements. First we compared our X-ray hydrostatic masses with those in common with two cluster samples that already had hydrostatic masses measured for them (see §4.8.1), also using *Chandra* data. We found an excellent agreement between our mass measurements, and those from the other two works. This provided strong evidence that our hydrostatic mass estimates did not contain a major systematic. However, we note that the clusters that were used in the comparison did not contain the lower mass clusters in our sample, so this test did not probe the validity of the hydrostatic mass measurements for the lower mass

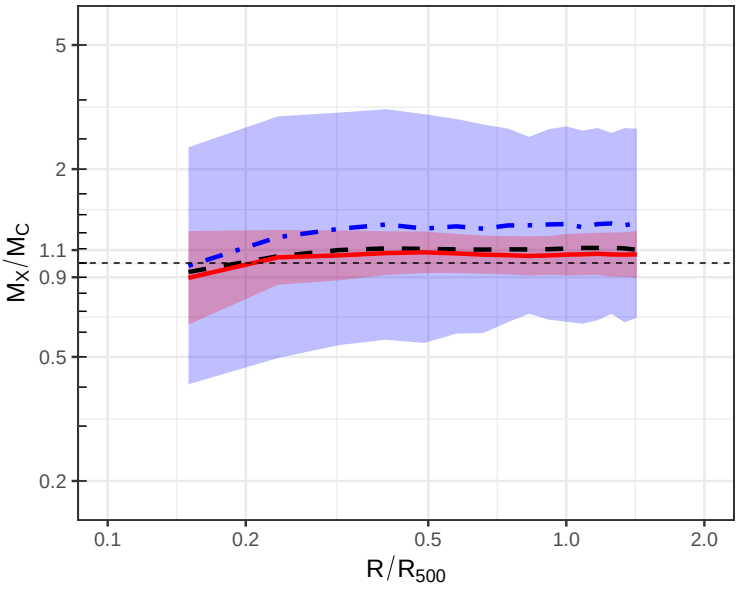

Figure 4.32: We show the average M_X/M_C profiles for the clusters with the highest Ngal values, excluding A1835 from this subsample (so leaving us with 13 clusters) due to reasons discussed in §4.9. Results for RCC clusters are shown as the blue dot-dashed lines and NRCC clusters as the red solid line. The blue shaded region shows the 1σ uncertainty for the RCC clusters, and the red shaded region shows the same uncertainty for the NRCC clusters.

clusters.

We also looked at scaling relations (see §4.8.2) for all 44 clusters in our sample. The offset that we found in the $L-T$ relation highlighted the fact that our analysis had higher temperature measurements than previous analyses, however, this was found to be due to changes in our analysis (calibration updates and changes in the statistical fitting of the spectra). Though these higher temperatures could lead to a higher mass estimate for the X-ray hydrostatic masses, we do fully understand why, and it is due to valid updates to the analysis. The $M_X - T$ relation agreed well with literature results, however, the $M_C - T$ relation had a much steeper slope than $M_X - T$ relations from literature, which suggested a systematic that varied with caustic mass. This was the first hint of an issue with the caustic masses that varied with some other quantity. In this case the quantity was temperature, an X-ray quantity, though this is due to the fact that the temperature of a cluster correlates with the Ngal value of a cluster (as less massive clusters generally have both lower temperature and lower Ngal values). The $M - Y_X$ relation agreed well with literature, providing no evidence for a systematic in our X-ray analysis. Finally, we looked at the $\sigma - T$ relation, for which our sample had a large discrepancy in the intercept of the scaling relations when compared with results from literature. Either we were overestimating the temperature for clusters in our sample which would lead to an overestimation in the hydrostatic masses, or underestimating σ values for a given temperature, which would likely lead to the caustic masses being underestimated. We see from the results from the $L - T$ relation that we have higher temperatures than previous analyses, however, it it is unlikely to be affecting our hydrostatic mass estimates significantly, as when we compare our hydrostatic masses to those from literature, we obtain a very good agreement (see §4.8.1).

We then looked at gas fractions for our sample of clusters using the hydrostatic mass and then the caustic mass in the denominator of the gas fraction (see §4.8.3). To test if there were any systematic that was dependent on the value of Ngal, a quantity relevant in the caustic method, which is the number of galaxies within the caustics, we split the cluster sample into subsamples of roughly even size based on their Ngal value. We found that when we used the hydrostatic mass as the denominator, the $f_{\rm gas}$ values obtained for all Ngal subsamples were as expected, whereas when we used the caustic masses as the denominator, we found that the $f_{\rm gas}$ values were too high for the lower Ngal clusters in our sample. This was substantial evidence that there was a systematic in the caustic mass method, which led to underestimating the cluster masses for low Ngal clusters.

As a final check on the validity of our hydrostatic masses, in §4.8.4, we compared the X-ray hydrostatic mass estimates that we had obtained and used for the main analysis using the forward-fitting mass method (see §3.2), with X-ray hydrostatic mass estimates obtained using a slightly different technique (but still using exactly the same X-ray data) called the backward-fitting mass method (see §3.3). We found that even though our forward-fitting method obtained slightly higher masses for the low mass clusters, and slightly lower masses for the high mass clusters, this was not a strong effect, and not nearly enough to account for the underestimation of the caustic masses for low Ngal clusters. We also saw no difference in the cluster mass estimates between the forward- and backward-fitting methods with Ngal. Taking into consideration the results from all the tests that we undertook, we conclude that there is likely to be a systematic in the caustic masses that is leading to an underestimation of the cluster masses for clusters with low Ngal.

Due to the strong evidence of an apparent systematic in the caustic masses of clusters with a low Ngal value, that was not evident in the higher Ngal clusters, we split the 44 clusters in our sample into three subsamples based on their Ngal value to give low, mid, and high Ngal subsamples with 15, 15, and 14 clusters respectively. We found that the M_X/M_C ratio was indeed strongly affected by the Ngal value of the clusters, with the M_X/M_C ratio of the low, mid, and high Ngal subsamples being 1.69$\pm$0.30, 1.28$\pm$0.15, and 1.12$\pm$0.11 respectively. These results were expected given the systematic in the caustic masses that we had previously found. We also split the 14 clusters in the highest Ngal subsample into two further subsamples based on their dynamical state, and calculated the M_X/M_C ratio at R_{500} of these subsamples, finding that there was no strong evidence for a dependence on the M_X/M_C ratio with dynamical state, though the errors on the ratios were large due to the small numbers of clusters in each subsample: we found the RCC clusters to have a ratio of 1.38$\pm$0.30 and the NRCC clusters to have a ratio of 1.06$\pm$0.14.

The results from the high Ngal subsample of 14 clusters are those for which we are most confident that the caustic masses are not affected by the systematic that we found. We therefore now focus on the results from the high Ngal subsample, and assume that the caustic masses in this subsample are reliable. The M_X/M_C ratio at R_{500} for the high Ngal subsample, 1.12$\pm$0.11, is greater than 0.9 at 3σ, which, given the expected overestimation of the caustic mass by $\sim$10% at R_{500}, suggests a low value of the hydrostatic bias, of $\leq$10% at 3σ. This value of the hydrostatic bias is not sufficient to fully account for the tension between the constraints placed on the

cosmological parameters by *Planck* clusters and those from the CMB. Our results from the high Ngal subsample also do not point to a dependence of the hydrostatic bias with radius, as the M_X/M_C ratio is roughly constant with radius. This is in contrast to some results from other work, especially from work using simulations to probe the dependence of the hydrostatic bias as a function of radius (e.g. Nagai et al., 2007). However we note that as the caustic mass itself is expected to have a bias that overestimates the cluster mass which decreases with radius (Serra et al., 2011), it could be the case that the hydrostatic (negative) bias is in fact increasing at higher radius, and the caustic mass (positive) bias is decreasing, and this leads to the roughly constant M_X/M_C ratio with radius.

Results from literature give a wide range of values for the hydrostatic bias, but they are generally in the range of $\sim$10-30%. Our results certainly favour a value of the hydrostatic bias at the low end of this range, if not lower. A recent study by Eckert et al. (2019) finds a value of the hydrostatic bias of $\sim$6% at R_{500}, consistent with our results.

Comparing our results from the high Ngal subsample to other studies that have compared X-ray and caustic masses, we find quite different results. Andreon et al. (2017), Foëx et al. (2017), and Armitage et al. (2019) find the the caustic mass to be *larger* than the X-ray masses by $\sim$15 - 30%, which would imply a M_X/M_C ratio of $\sim$0.75-0.85. However, we note that in Foëx et al. (2017), eight of the ten clusters in their sample had Ngal values such that they would appear in our high Ngal bin. Despite this, these results are significantly lower than our M_X/M_C ratio of 1.12$\pm$0.11 that we find in the 14 cluster high Ngal bin.

An earlier study by Rines et al. (2003) compares X-ray and caustic masses for eight clusters, finding the caustic to X-ray mass ratio to be 1.03$\pm$0.11, equivalent to a value of M_X/M_C of 0.97$\pm$0.11 at R_{500}, significantly higher than the more recent comparisons of X-ray and caustic masses, but closer to our results, and also suggesting a small value of the hydrostatic bias. However, we note that the X-ray masses from Rines et al. (2003) are estimated using a mass-temperature relation, and not from a full hydrostatic mass analysis, meaning that the results from this work are completely dependent on the normalisation of the mass-temperature relation used to estimate X-ray masses.

4.11 Summary and Conclusions

For 44 clusters in the CHeCS sample, we have compared the hydrostatic and caustic mass profiles.

- We find the value for the M_X/M_C ratio is 1.33±0.10 at the X-ray R_{500}, and find it is 1.32±0.10 at 1 Mpc for all 44 clusters in our sample.
- We find no evidence for a dependence of the M_X/M_C ratio on the dynamical state for the 44 clusters in our sample, with the M_X/M_C ratio being 1.32±0.29 for the 10 RCC clusters, and 1.33±0.11 for the 34 NRCC clusters.
- We find that the M_X/M_C ratio is dependent on the number of cluster members with measured redshifts used for the caustic mass estimation method, Ngal, as the caustic mass appears to be significantly underestimated for clusters with low Ngal values.
- Splitting our 44 cluster sample into equally sized subsamples by Ngal value, we find the M_X/M_C ratio at the X-ray R_{500} to be 1.69±0.30, 1.28±0.15 and 1.12±0.11 for the low, mid and high Ngal subsamples.
- For the 14 clusters in the Ngal $\geq$ 210 subsample, we find no evidence for a dependence of the M_X/M_C ratio on dynamical state, with the M_X/M_C ratio being 1.38±0.30 for the 5 RCC clusters (or 1.32±0.65 when A1835 is dropped from the sample), and 1.06±0.14 for the 9 NRCC clusters.
- For the 14 clusters in the high Ngal subsample, there is no evidence that the hydrostatic bias is a function of radius, and both the caustic and hydrostatic mass profiles agree within 10% at all radii (except the inner radii where the caustic masses are expected to significantly overestimate the true cluster mass).
- We present a subsample of 14 clusters that have reliable caustic masses and hydrostatic masses to be used for further study.

5

Gas fractions in cluster outskirts

5.1 Introduction

Galaxy clusters are formed from the primordial density fluctuations via gravitational collapse and subsequent hierarchical merging of smaller structures, until they reach their gargantuan sizes seen at later cosmic times. Even at recent times, cluster growth is an ongoing process, predominantly occurring via the accretion of smaller objects and gas, either from filaments or the environment surrounding the cluster.

The effect of this accretion on the cluster is greatest in the cluster outskirts, where there is a transition between the virialised structure of the cluster and the region where material is still infalling (see Reiprich et al., 2013 or Walker et al., 2019 for recent reviews), causing inhomogeneities in the intra-cluster medium (ICM) in this region. Although accretion still occurs at lower cluster radius, the effects of accretion are seen most strongly in the outer regions of the cluster as here the ICM density is lower. The effects of accretion can have a large impact on derived quantities obtained from observations of clusters in the X-ray energy band, due to these inhomogeneities in the ICM (or clumpiness) in the cluster outskirts (e.g. Simionescu et al., 2011). In addition, the assumption that the cluster is in hydrostatic equilibrium (which is required for X-ray hydrostatic mass estimates) is poorer in this region than at smaller cluster radii (e.g. Nagai et al., 2007; Lau et al., 2009; Eckert et al., 2019). The breakdown of hydrostatic

equilibrium implies that when using X-ray data to measure the hydrostatic mass profile, the hydrostatic bias is likely to be stronger at the cluster outskirts than at smaller radii, and this is seen in work using both observations (e.g. Eckert et al., 2019) and simulations (e.g. Nagai et al., 2007). The potential clumping of the ICM would lead to overestimates of not only the hydrostatic mass at these radii, but also, more strongly, the gas mass. This overestimation of the gas mass occurs as if the ICM is clumpy, then the emissivity will be overestimated (as the X-ray emissivity of the ICM scales as the square of the gas density), which in turn leads to the gas mass being overestimated. Thus, when using X-ray hydrostatic mass profiles and gas mass profiles to obtain profiles of the gas mass fraction, $f_{\rm gas}$ (calculated as the gas mass divided by the total cluster mass), higher than expected $f_{\rm gas}$ values may be measured at large cluster radii, either due to the underestimation of the hydrostatic mass due to the assumption of hydrostatic equilibrium breaking down, or the overestimation of the gas mass due to clumping of the ICM, or both. This has been seen in studies such as Simionescu et al. (2011) where super-cosmic values of $f_{\rm gas}$ are seen at large cluster radii ($\gtrsim 0.6\ R_{200}$).

We can therefore measure the value of $f_{\rm gas}$ at large cluster radii in order to probe the physical processes occurring in the cluster outskirts. However, to do this using X-ray data alone can be a difficult task, due to the fact that in order to directly measure the mass profile of a cluster using the X-ray hydrostatic mass method, spatially resolved measurements of the ICM temperature are required, and the radius out to which these can be obtained is limited due to the faintness of the cluster emission at larger radii. One way to circumvent this issue is to use a cluster mass estimation method that is not extrapolated and is unbiased out to larger radii, such as the caustic mass method, which we use in this chapter.

Additionally, measurements of the value of $f_{\rm gas}$ at large cluster radii can also be used to constrain cosmological parameters, by treating clusters as 'standard buckets' (e.g. Ettori et al., 2009). This can be done as clusters are so large that the matter distribution within them is expected to be representative of the matter distribution in the Universe as a whole (White et al., 1993; Evrard, 1997; Kravtsov et al., 2005). Baryon depletion due to AGN feedback, leading to a lower than expected $f_{\rm gas}$ using this 'standard bucket' technique, is only significant in lower mass systems ($\gtrsim 2{\times}10^{14}$ M$_{\odot}$, Planelles et al., 2013). $f_{\rm gas}$ measurements at smaller cluster radii (R_{2500}) can also be used to constrain cosmological parameters (e.g. Allen et al., 2004, 2008, 2011), with the exact choice of radius used being a trade-off between not being too small such that non-gravitational effects lead to a large scatter, and not being too large that the measurement

error becomes too large (as the limits of the instrument are reached). For this method, due to the $f_{\rm gas}$ values being measured at smaller radii, the level of baryon depletion does impact the results, and has to be modelled.

For this study, we use the universal baryon fraction (Ω_b / Ω_m) determined by Planck Collaboration et al. (2018) with high precision to be 0.157±0.003 to calculate a universal gas fraction using:

$$f_{\rm gas,univ}(r) = Y_{\rm b}(r)\frac{\Omega_b}{\Omega_m} - f_{\rm star} \tag{5.1}$$

where $f_{\rm star}$ is the fraction of baryons found in stars, and $Y_{\rm b}$ is the baryon depletion factor. The baryon depletion factor gives the fraction of baryons that is enclosed within a given radius of a cluster relative to the cosmic mean baryon fraction. The baryon depletion factor is a function of radius, as one effect that leads to baryon depletion in clusters is blowout of gas by AGN activity (e.g. McNamara & Nulsen, 2012), which has a larger effect at smaller radii. We note that baryons are present in other forms, such as warm and molecular gas, but these account for a fraction of a percent of the total baryon mass (e.g. Edge et al., 2002), so we ignore their contribution. We use the value of 0.95 for $Y_{\rm b}$, similar to the value used by Eckert et al. (2019). Although, as noted, there is in fact some dependence of $Y_{\rm b}$ on radius, we are interested in the value of $f_{\rm gas}$ in the cluster outskirts where the value of $Y_{\rm b}$ does not vary significantly. Using the conservative value of $f_{\rm star}$ of 0.015±0.005 as in (Eckert et al., 2019), and assuming it to be constant at all radii (again not true, but fine for these purposes), we obtain a universal gas fraction of 0.134±0.006. We refer to this value as the universal gas fraction, or cosmic value of the gas fraction, throughout this study.

In this chapter we use gas mass profiles obtained from X-ray data and caustic total mass profiles to measure $f_{\rm gas}$ out to large radii. We calibrate the caustic mass profiles at smaller cluster radii using X-ray hydrostatic mass profiles, in the region where the X-ray hydrostatic mass profiles are not extrapolated. Given the uncertainties on possible systematics in the caustic masses themselves (see Chapter 4), this approach does not rely on the absolute value of the caustic mass measurements, instead we are primarily interested in using the *shape* of the caustic profiles to effectively 'extrapolate' the hydrostatic masses to larger radii.

5.2 Sample and X-ray Data Analysis

Our cluster sample is derived from the Hectospec Cluster Survey (HeCS, Rines et al., 2013) as mentioned in §4.2, and has *Chandra* observations for all clusters in HeCS, and so is termed CHeCS (*Chandra* observations of the HeCS clusters). In Chapter 4 we calculated the X-ray hydrostatic mass profiles and $m_{\rm gas}$ profiles for these clusters, and we use these profiles for the work in this chapter. We note that the X-ray hydrostatic mass profiles used throughout this chapter are calculated using the forward-fitting hydrostatic mass method, detailed in §3.2. The clusters also have mass profiles obtained using the caustic method (from Rines et al., 2013), described in §4.5. For this chapter we focus on the sample of 14 clusters described in §4.8.5 (i.e. those labelled 'high' in the Ngal subsample column in Table 4.3), as we expect these clusters to be unaffected by the systematic that we found in Chapter 4 that leads to the caustic mass method underestimating the cluster mass for clusters with low Ngal values. However, we note that in §5.4 we do also use the full sample of clusters as presented in Chapter 4.

5.3 Modelling

Using the gas masses and hydrostatic masses as calculated from the `emcee` chains in Chapter 4, in addition to the caustic mass profiles from Rines et al. (2013), we use a Bayesian framework to constrain the gas fraction. As in Chapter 4 we scale our mass profiles by the hydrostatic estimate of R_{500}.

As mentioned in §4.6, the errors on our cluster masses (and now also gas mass), are better described by a lognormal distribution, so we model these errors in base 10 log space. However, we note that we still present our results (in Table 5.1 and in all figures) in linear space unless explicitly stated otherwise, as this facilitates comparison with literature values, which are often reported as such. We relate the mean and standard deviation of the X-ray and caustic masses in linear space (M and S respectively) to log space (μ and σ respectively) as in equations 4.4 and 4.5.

As in §4.6, we model the 'true' hydrostatic and caustic masses as having a bias (κ_X and κ_C respectively) compared to the true mass of the cluster, μ, as well as a scatter (δ_X and δ_C respectively). The 'true' hydrostatic and caustic masses are modelled as

$$\mu_X \sim \mathcal{N}(\mu + \kappa_{\rm X}, \delta_{\rm X}) \tag{5.2}$$

$$\mu_C \sim \mathcal{N}(\mu + \kappa_{\rm C}, \delta_{\rm C}) \tag{5.3}$$

where the $\mathcal{N}$ refers to a normal distribution, and '$\sim$' is equivalent to 'is distributed as'. We also model the X-ray gas mass, $m_{\rm gas}$, in a similar way, where instead of a bias, we have $\log_{10}(f_{\rm gas})$, which effectively converts the total mass in log space into a gas mass, and a scatter, $\delta_{\rm mgas}$, around the 'true' gas mass value, $m_{\rm gas}$, as

$$m_{\rm gas} \sim \mathcal{N}(\mu + \log_{10}(f_{\rm gas}), \delta_{\rm mgas}) \tag{5.4}$$

These 'true' hydrostatic and caustic masses, and X-ray gas masses are related to the corresponding observed quantities (marked with a $\hat{}$) as follows:

$$\hat{\mu_X} \sim \mathcal{N}(\mu_{\rm X}, \sigma_{\rm X}) \tag{5.5}$$

$$\hat{\mu_C} \sim \mathcal{N}(\mu_{\rm C}, \sigma_{\rm C}) \tag{5.6}$$

$$\hat{m_{\rm gas}} \sim \mathcal{N}(m_{\rm gas}, \sigma_{\rm mgas}) \tag{5.7}$$

where σ_X and σ_C represent the standard deviation of the lognormal error for the X-ray hydrostatic and caustic masses respectively, and $\sigma_{\rm mgas}$ represents the same but for the gas masses.

In the following sections, we will explore different methods for the specific modelling setup used, following the framework as outlined in this section.

5.3.1 Two separate regimes model

As the X-ray hydrostatic mass profiles are extrapolations outside the radius to which the measured temperature profile extends (shown for each cluster in Figure 5.1), we decide to use two slightly different models depending on the radius at which we are looking.

a) 0.2 - 1 R_{500} Regime

We use the following model in the 0.2 - 1 R_{500} radius range as it is the regime where the X-ray hydrostatic masses are largely not extrapolated. We summarize the model in Figure 5.2 (top panel), and it can be defined as follows:

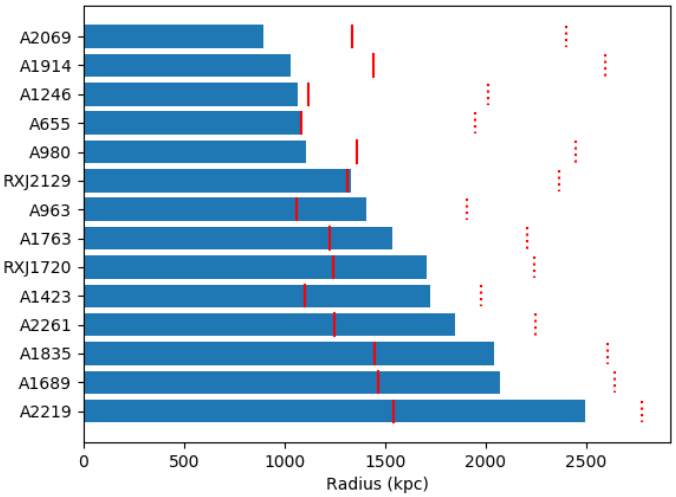

Figure 5.1: We show the radius out to which there is observational data for each cluster's temperature profile as a blue bar, which is used to obtain X-ray hydrostatic mass profile (see §3.2 and §4.3.8). Outside the outer radius, a cluster's hydrostatic mass profile is extrapolated. We plot each cluster's R_{500} as a solid red line, and 1.8 $\times$ R_{500} as a red dotted line.

$$
\begin{aligned}
m_{\rm gas} &\sim \mathcal{N}(\mu + \log_{10}(f_{\rm gas}), \delta_{\rm mgas}) \\
\hat{m_{\rm gas}} &\sim \mathcal{N}(m_{\rm gas}, \sigma_{\rm mgas}) \\
\mu_C &\sim \mathcal{N}(\mu + \kappa_{\rm C}, \delta_{\rm C}) \\
\hat{\mu_C} &\sim \mathcal{N}(\mu_{\rm C}, \sigma_{\rm C}) \\
\mu_X &\sim \mathcal{N}(\mu + \kappa_{\rm X}, \delta_{\rm X}) \\
\hat{\mu_X} &\sim \mathcal{N}(\mu_{\rm X}, \sigma_{\rm X})
\end{aligned}
\tag{5.8}
$$

For the 0.2 - 1 R_{500} model, the priors that we use on the quantities related to the X-ray hydrostatic and caustic masses are the same as in §4.6, and we place mostly weak priors on the model parameters, with the exception of κ_X (which represents the bias between the true mass, μ, and the 'true' hydrostatic mass, μ_X), where we place a tighter prior. The loose priors placed on the majority of the model parameters are as follows:

- μ, μ_X and μ_C are assigned a uniform prior with a range of 12 - 17.
- $m_{\rm gas}$ is assigned a uniform prior with a range of 11 - 16.
- κ_C is assigned a normal prior with a mean of 0 and standard deviation of 1.

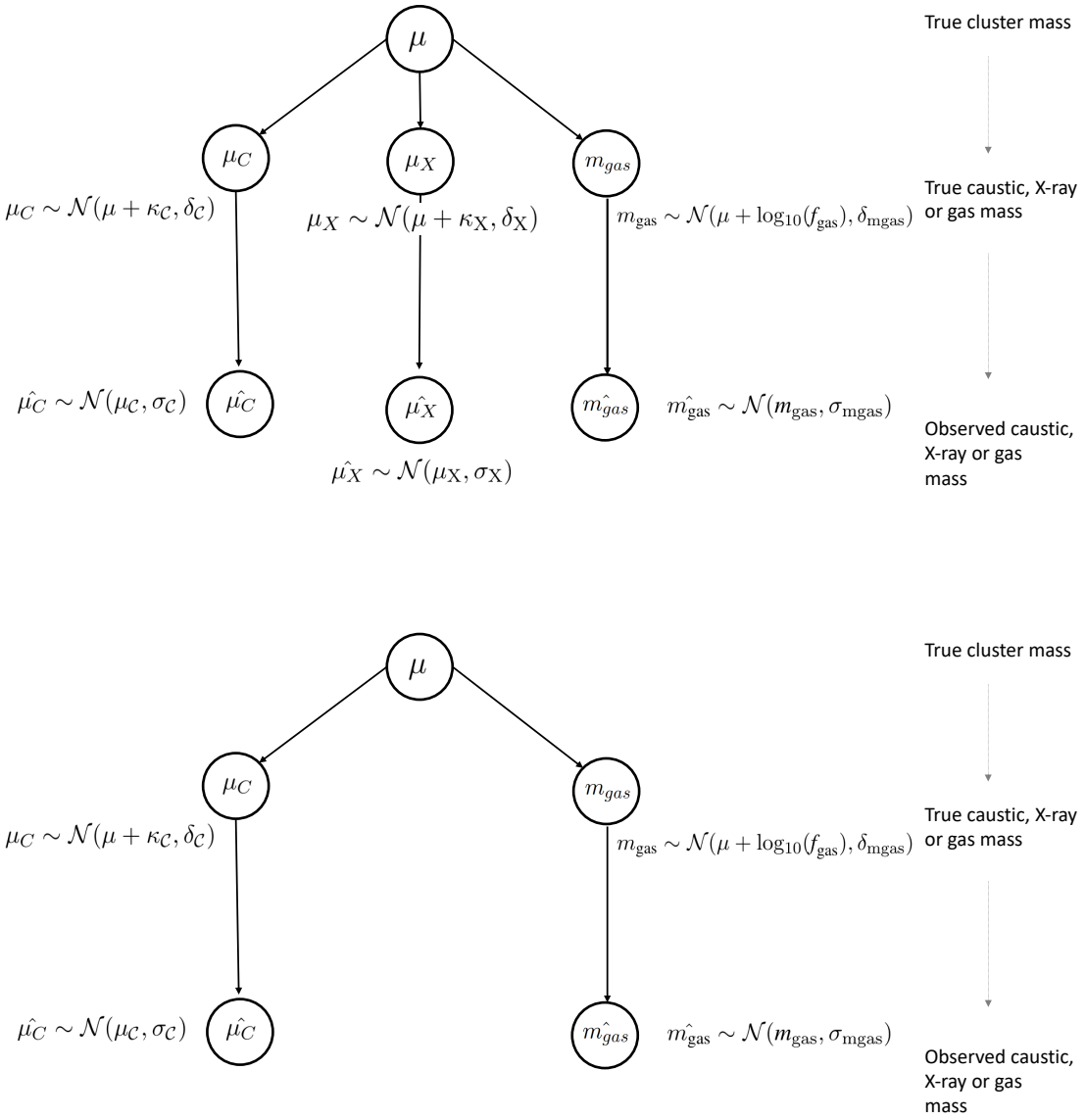

Figure 5.2: Graphical models summarising the mass modelling framework that we present in §5.3.1. The upper panel shows the setup for when the X-ray hydrostatic masses are used in the modelling in addition to the caustic masses and gas masses (in the 0.2 - 1 R_{500} regime), the lower panel shows the setup for when only the caustic and gas masses are used in the modelling (in the 1 - 1.8 R_{500} regime). The symbols are also described in that section; in summary, $\log_{10}(f_{\mathrm{gas}})$ is the logarithm (in base 10) of the gas fraction, the symbol μ represent mass, κ is the bias term, δ is the scatter term and σ is the measurement error term, and the c and x subscripts denote whether the quantity refers to that from the X-ray hydrostatic or caustic method. The $\hat{}$ refers to an observed quantity. $\mathcal{N}$ refers to a normal distribution, and "$\sim$" is equivalent to "is distributed as".

- The intrinsic scatter terms, δ_X, δ_C and $\delta_{\rm mgas}$ are assigned normal priors (truncated at zero) with a mean of 0.09 and standard deviation of 2.2, corresponding to a weak prior with a mean of $\sim$20% in normal space.
- $\log_{10}(f_{\rm gas})$ is given a normal prior with a mean of -1 and standard deviation of 1.

For the prior on κ_X, we use a Gaussian prior centred on -0.046 (in log space) which corresponds to underestimating the true mass by 10%. There is not a strong agreement in literature about the value of the hydrostatic bias, but as our results in Chapter 4 suggest a low value of the hydrostatic bias, we use this value initially. However, we note that we explore the impact of using other values for the centre of the Gaussian prior on κ_X in §5.3.5. For the strength, or tightness, of the prior on κ_X, we explored various options. The priors we tested were $\kappa_X \sim \mathcal{N}$ (-0.046, σ), with the values of σ as 0.01, 0.05 and 1 (while using the other priors set with the values as presented above). We present the results in Figure 5.3, and it can clearly be seen that a tight prior is needed in order to constrain $f_{\rm gas}$; this is expected as κ_X, κ_C and $\log_{10}(f_{\rm gas})$ are degenerate. Thus, we could have placed a tight prior on κ_C instead, and a loose prior on κ_X, but choose to place the tighter prior on κ_X, as more work has been done exploring the value of this bias. In summary, due to the above considerations, we choose to place a tight prior on κ_X, such that $\kappa_X \sim \mathcal{N}$ (-0.046, 0.01). We always use this prior on κ_X unless explicitly stated otherwise.

b) 1 - 1.8 R_{500} Regime

For the regime where most hydrostatic masses become extrapolated (1 - 1.8 R_{500} - we probe the $f_{\rm gas}$ profile up to 1.8 R_{500} as the outer radii measurements of most surface brightness profiles are outside this radius, discussed in the following section) we simply remove any terms that require the X-ray hydrostatic mass from our model. We summarize this model in Figure 5.2 (bottom panel), and it can be defined as follows:

$$\begin{aligned} m_{\rm gas} &\sim \mathcal{N}(\mu + \log_{10}(f_{\rm gas}), \delta_{\rm mgas}) \\ \hat{m_{\rm gas}} &\sim \mathcal{N}(m_{\rm gas}, \sigma_{\rm mgas}) \\ \mu_C &\sim \mathcal{N}(\mu + \kappa_{\rm C}, \delta_{\rm C}) \\ \hat{\mu_C} &\sim \mathcal{N}(\mu_{\mathcal{C}}, \sigma_{\rm C}) \end{aligned} \tag{5.9}$$

For the 1 - 1.8 R_{500} model, we use the same priors as in the 0.2 - 1 R_{500} model on the shared quantities, with the exception of κ_C, which we set to the value that is learned (i.e. the posterior of κ_C) from the modelling step run at 1 R_{500} (with the X-ray data included, i.e. so it depends

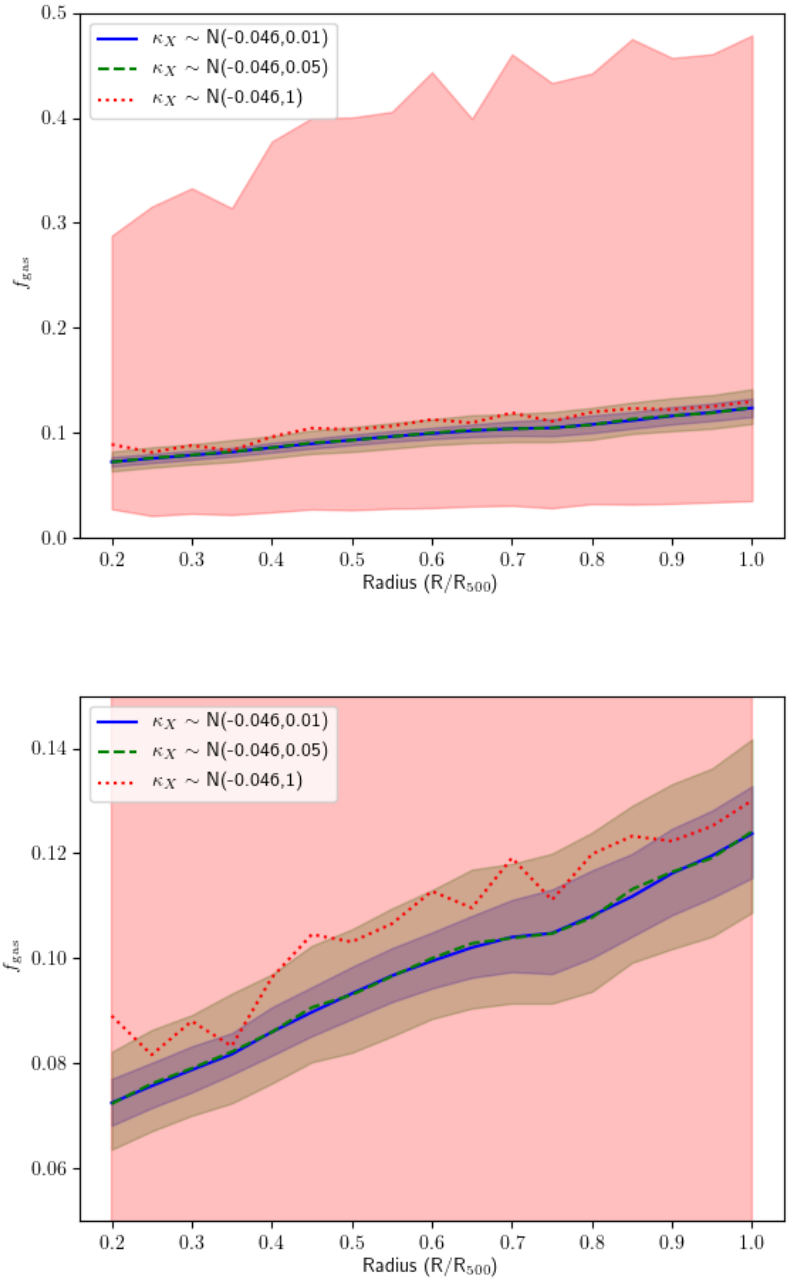

Figure 5.3: We show the $f_{\rm gas}$ profile obtained when we vary the tightness of the prior on the hydrostatic bias κ_X. The bottom panel is a zoomed in version of the top panel. 1 σ errors are shown by the shaded region.

on the assumed κ_X). We find that at 1 R_{500} that value of κ_C is -0.09 with standard deviation of 0.04, corresponding to a negative caustic bias of 19$\pm$8% , which is the same as the value found at 0.95 and 0.9 R_{500}, so we see it is not changing much around this radius. This negative caustic bias is expected, despite Serra et al. (2011) predicting that the caustic method will give a $\sim$ 15% overestimation of true mass at R_{500}. This is because we have set a tight prior on κ_X which corresponds to a 10% underestimation of true mass, and given the results from Chapter 4 where we find M_X/M_C to be $\sim$ 1.1$\pm$0.1 at R_{500}, we would therefore expect caustic bias to be an underestimation of the true mass of around 20%, which is what we find.

c) Model Implementation

Using the models described above, we used the observational data that we have for each cluster, $\hat{\mu_C}$, σ_C, $\hat{m_{\rm gas}}$, $\sigma_{\rm mgas}$ (and $\hat{\mu_X}$ and σ_X for the 0.2 - 1 R_{500} regime), to constrain the scatter terms, δ_C, $\delta_{\rm mgas}$ (and δ_X for the 0.2 - 1 R_{500} regime), and bias terms, κ_C (and κ_X for the 0.2 - 1 R_{500} regime), as well as $\log_{10}(f_{\rm gas})$ for all of the clusters in our sample (thus we sometimes refer to the scatter, bias and $\log_{10}(f_{\rm gas})$ terms as global parameters of our model).

We use the probabilistic programming language STAN [1] to implement the models described above. We use the PyStan [2] interface for this. We sampled the parameters in our model with four chains of 5,000 steps each. By repeating this analysis at increasing radii, we can produce a profile of the gas fraction.

We note that the gas mass profiles will also be extrapolated at large radii, and show the outer radius for which there is observational data for the surface brightness profile (from which the gas mass is derived) in Figure 5.4. However, for almost all clusters, this is either above the 1.8 R_{500} outer radius to which we model the gas fraction profile, or very close to it. Given that when extracting the surface brightness profile (see §4.3.9) we required each radial bin to have net counts of at least 50 but used no signal-to-noise ratio constraint, some of the surface brightness profiles extend to very large radii, as seen in Figure 5.4. Our binning also means that for all clusters we have potentially overestimated the extent to which we are, in reality, constraining the surface brightness profiles. Thus, to explore the extent to which we have *robust* observational data, we place the additional constraint on the binning such that the signal-to-noise ratio has to be at least two in each radial bin, and see that the surface brightness profiles

[1] `http://mc-stan.org`
[2] `https://pystan.readthedocs.io/en/latest/`

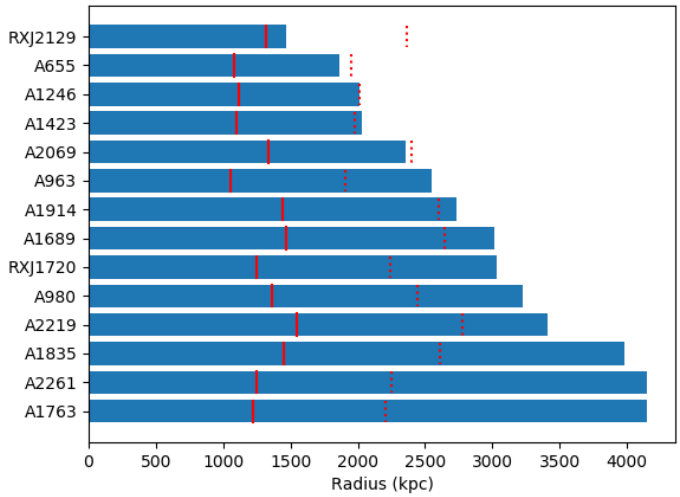

Figure 5.4: We show the radius out to which there is observational data for each cluster's surface brightness profile as a blue bar, which is used to obtain the gas mass profile (see §3.2 and §4.3.9). Outside the outer radius, a cluster's gas mass profile is extrapolated. We plot each cluster's R_{500} as a solid red line, and 1.8 $\times$ R_{500} as a red dotted line.

that previously extended to very high radii do not extend to such extreme outer radii. This added constraint does not affect our choice of outer radius to which we model the $f_{\rm gas}$ profiles, as it is still the case that for our sample of 14 clusters that we use for the primary analysis, most of the surface brightness profiles are adequately constrained out to 1.8 R_{500}. Additionally, even for those surface brightness profiles that are in reality effectively extrapolated from radii significantly below 1.8 R_{500}, the uncertainties on the $m_{\rm gas}$ profiles will increase at these larger radii, meaning that the global parameter $f_{\rm gas}$ will not be so affected by these clusters as compared to those that have well constrained $m_{\rm gas}$ profiles out to 1.8 R_{500}.

We note that the conversion of the surface brightness profile to an emission measure profile (which is then converted into a gas density profile, which in turn is used to obtain the gas mass profile - see §3.2.4 and 4.3.9) requires the 3D temperature profile, which is obtained from the measured projected temperature profile (see §4.3.8). Most temperature profiles are extrapolated at a smaller radius than the surface brightness profiles, however the conversion from surface brightness to emission measure is weakly dependent on the temperature, so the fact that most temperature profiles are extrapolated to some extent in the 1 - 1.8 R_{500} regime will not strongly affect the measured gas mass profiles.

	Two separate regimes model	Smooth transition model (not updating prior on κ_C)	Smooth transition model (updating prior on κ_C)	Smooth transition with covariance model
$f_{\rm gas}$ at R_{500}	$0.124^{+0.009}_{-0.009}$	$0.132^{+0.010}_{-0.009}$	$0.129^{+0.008}_{-0.007}$	$0.130^{+0.008}_{-0.008}$
$f_{\rm gas}$ at 1.4 R_{500}	$0.141^{+0.018}_{-0.016}$	$0.187^{+0.079}_{-0.056}$	$0.155^{+0.014}_{-0.013}$	$0.156^{+0.016}_{-0.015}$
$f_{\rm gas}$ at 1.8 R_{500}	$0.153^{+0.021}_{-0.018}$	$0.192^{+0.545}_{-0.142}$	$0.168^{+0.016}_{-0.015}$	$0.167^{+0.017}_{-0.016}$

Table 5.1: We show the value of $f_{\rm gas}$ at different radii for the different methods that we present in §5.3. All errors are 1 σ errors. The two separate regimes model is descrbied in §5.3.1, the smooth transition model in §5.3.2, and the smooth transition with covariance model in §5.3.3. We present the 1.4 and 1.8 R_{500} values as $R_{200} \sim 1.4R_{500}$ and R_{200} is around where the virial region of a cluster ends, and 1.8 R_{500} is significantly outside this radius and is the outer radius to which we don't extrapolate the majority of our $m_{\rm gas}$ profiles.

We show the gas fraction profile for the modelling method described in this section (we reiterate that the prior placed on κ_X is $\kappa_X \sim \mathcal{N}$ (-0.046, 0.01)) in Figure 5.5 and present the $f_{\rm gas}$ values in Table 5.1. We find that the value of $f_{\rm gas}$ increases with radius but is consistent with the cosmic $f_{\rm gas}$ value of 0.134$\pm$0.006, even at 1.8 R_{500}, where it has a value of 0.153$\pm$0.020. A clear feature is that the 1 σ error on the $f_{\rm gas}$ profile suddenly increases at 1 Mpc. This is expected due to the fact that it is only the caustic masses that are now constraining the value of $f_{\rm gas}$, as the X-ray hydrostatic mass profiles are dropped from the analysis at 1 Mpc; additionally, the prior on κ_C used when the X-ray masses are dropped is less tight than the prior on κ_X that is used before the X-ray masses are dropped from the analysis, and this also leads to increased errors on the $f_{\rm gas}$ profile after 1 Mpc.

Another noticeable feature in Figure 5.5 is that there is a flattening in the $f_{\rm gas}$ value at around 0.8 R_{500} that appears just before the $f_{\rm gas}$ value rises again. This is due to the unphysical hydrostatic mass profile of A1835 (discussed further in §4.7) which is included in our 14 cluster sample. In §5.3.4 we investigate the impact of dropping A1835 from our cluster sample, and find that the impact of including it in our sample is small.

We find that the value at which we choose the cut-off radius affects our results significantly (see Figure 5.6), and as such present an improved model in the next section.

5.3.2 Smooth transition model

A disadvantage of the previous approach is that using a single cut-off radius for the inclusion of the hydrostatic mass profiles does not make full use of the information available for those profiles which extend to larger radii. Thus, we extend the modelling technique that is described

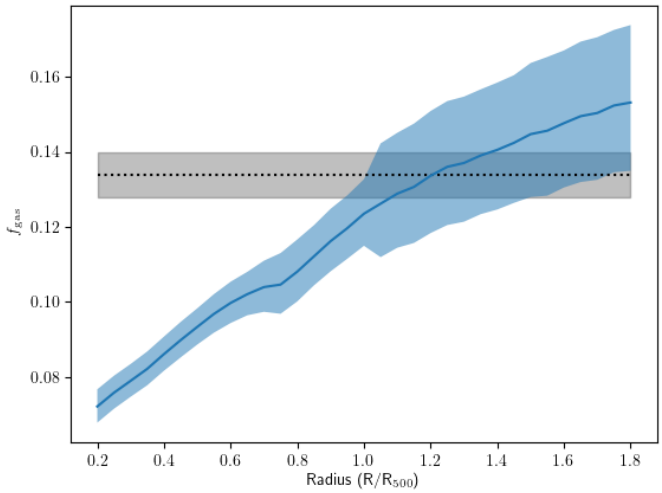

Figure 5.5: We show the $f_{\rm gas}$ as a function of radius for when we drop the X-ray hydrostatic mass profiles after 1 Mpc (i.e. the two separate regimes model, described in §5.3.1) as a blue line. 1 σ errors are shown by the shaded region. We show the universal gas fraction as a black dotted line, with the 1σ errors shown as shaded.

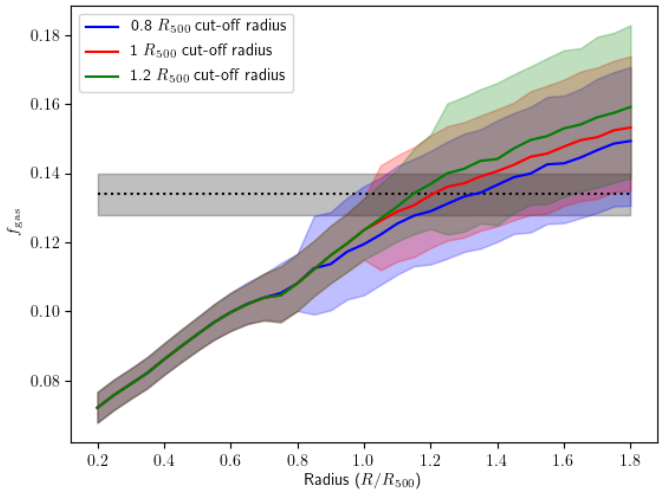

Figure 5.6: We show the $f_{\rm gas}$ as a function of radius for our modelling setup as in §5.3.1 (i.e. the two separate regimes model) for when we choose different cut-off radii. 1 σ errors are shown by the shaded region. We show the universal gas fraction as a black dotted line, with the 1σ errors shown as shaded.

in §5.3.1 (i.e. the two separate regimes model), and instead of using a single 'cut-off' radius of 1 R_{500} for where we stop using the X-ray hydrostatic mass profile data, we drop the X-ray hydrostatic mass data on a cluster-by-cluster basis, depending on whether the hydrostatic mass profile is being extrapolated or not at that radius (shown in Figure 5.1). We implement this by increasing the errors to a very large number on the hydrostatic mass after the extrapolation radius. For each radius after which at least one cluster's hydrostatic mass has been dropped from the analysis, the prior on κ_C for the next step is updated to be the value that is learned (i.e. the posterior of that parameter after model fitting) from the previous step. This means that the bias on the caustic masses, κ_C, is being constrained by the hydrostatic masses, and hence depends on the prior placed on the hydrostatic bias, κ_X.

We show the $f_{\rm gas}$ profile as a function of radius in Figure 5.7 and give values at certain radii in Table 5.1 (column 'Smooth transition model (updating prior on κ_C)'). We obtain consistent results (within 1 σ error) as the model presented in §5.3.1 (i.e. the two separate regimes model), but the $f_{\rm gas}$ values for the model presented in this section are slightly higher at large radii, and are in tension with the cosmic value of $f_{\rm gas}$.

In Figure 5.8 we show the $f_{\rm gas}$ profile obtained using the approach above (i.e. dropping the X-ray hydrostatic mass data on a cluster-by-cluster basis, and using the prior $\kappa_X \sim \mathcal{N}$ (-0.046, 0.01)), but without updating the prior on κ_C as the X-ray mass profiles are dropped from the model (i.e. we keep the loose κ_C prior at all radii). We also give values at certain radii in Table 5.1 (column 'Smooth transition model (not updating prior on κ_C)'). These results show how the constraints that we can put on the $f_{\rm gas}$ profile become weaker as the X-ray hydrostatic mass profiles are dropped from the analysis as we get to larger radii, as expected. We note that we only present the results obtained when using a loose prior on κ_C at all radii for the reader's interest, and that it is not the model adopted in this section.

5.3.3 Smooth transition model with covariance

We build further on the modelling approach used in §5.3.2 (i.e. the smooth transition model, where we updated the prior on κ_C at each step after which the first X-ray mass profile had been dropped), and now introduce a covariance term between the X-ray hydrostatic mass and the gas mass. We do this, as it is likely that if the gas mass is overestimated (or underestimated), the hydrostatic mass is also likely to be overestimated (or underestimated), as they both require the use of the gas density profile in their measurement (obtained from the surface brightness profile,

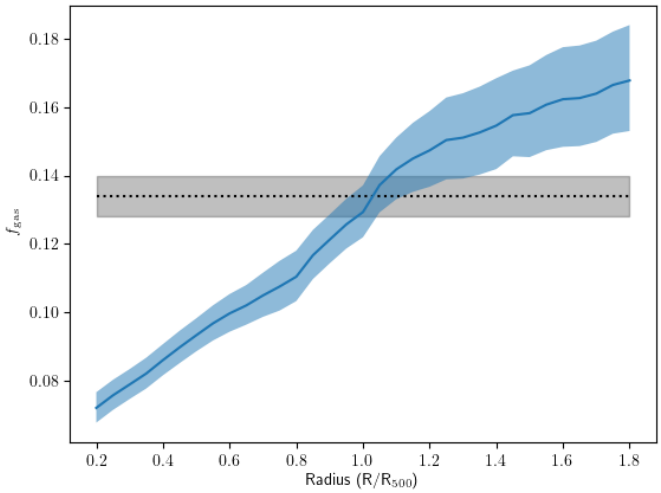

Figure 5.7: We show the $f_{\rm gas}$ as a function of radius for when we drop the X-ray hydrostatic mass profiles one by one and update the prior on κ_C each step after the first X-ray mass is dropped (i.e. the smooth transition model, described in §5.3.2) as a blue line. 1 σ errors are shown by the shaded region. We show the universal gas fraction as a black dotted line, with the 1σ errors shown as shaded.

see 4.3.9).

Thus, instead of modelling the hydrostatic mass and $m_{\rm gas}$ independently, they can be modelled as follows[3]:

$$\begin{pmatrix} \mu_X \\ \mu_C \\ m_{\rm gas} \end{pmatrix} \sim \mathcal{N}\left[\begin{pmatrix} \mu + \kappa_X \\ \mu + \kappa_C \\ \mu + \log_{10}(f_{\rm gas}) \end{pmatrix}, \begin{pmatrix} \delta_X^2 & 0 & \sigma_{\rm mxmg} \\ 0 & \delta_C^2 & 0 \\ \sigma_{\rm mxmg} & 0 & \delta_{\rm mgas}^2 \end{pmatrix} \right] \tag{5.10}$$

where $\mathcal{N}$ refers to a multivariate Gaussian distribution, and $\sigma_{\rm mxmg}$ is the covariance between the hydrostatic mass and the gas mass. The above model replaces equations 5.2, 5.3 and 5.4 in the previous analysis. We note that the modelling of the caustic mass is effectively unchanged, as we continue to treat the caustic mass as independent from the other quantities.

[3]We note that for the implementation of modelling the covariance in STAN, we apply Cholesky decomposition to the covariance matrix. Cholesky decomposition decomposes a positive definite matrix (a covariance matrix has to be positive definite), **M**, into the product of two matrices, a lower triangular matrix, **L**, and its transpose, such that $\mathbf{M} = \mathbf{L}\mathbf{L}^T$. We apply Cholesky decomposition to the covariance matrix for use in the MCMC modelling performed in STAN, as it is more numerically stable or efficient to use the Cholesky decomposition rather than the covariance matrix.

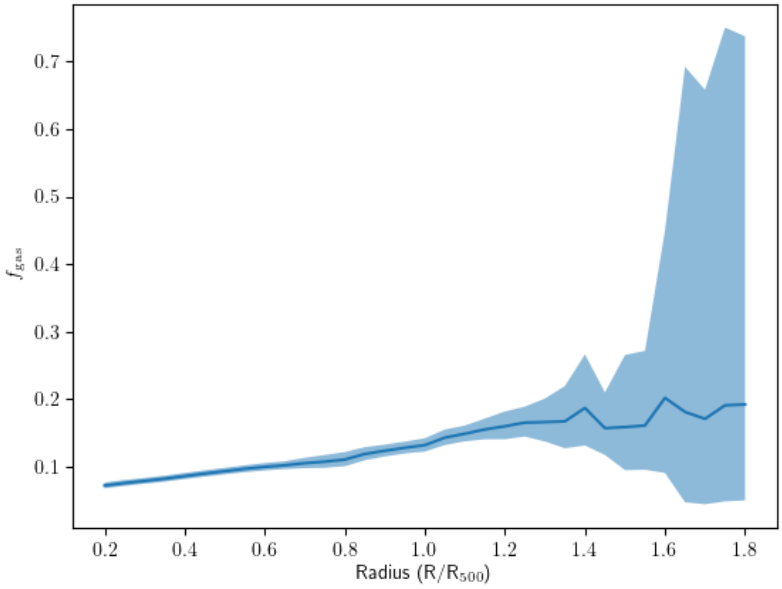

Figure 5.8: We show the $f_{\rm gas}$ as a function of radius for when we drop the X-ray hydrostatic mass profiles one by one and *don't* update the prior on κ_C (mentioned in last paragraph of §5.3.2). 1 σ errors are shown by the shaded region.

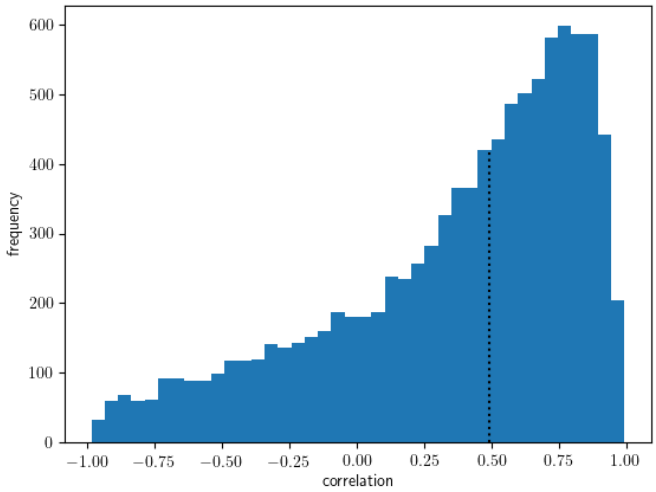

Figure 5.9: We show the range of values in our `emcee` chains for the correlation between the hydrostatic mass and the gas mass as computed from the covariance matrix (see §5.3.3) at 1 R_{500}. The dashed black line shows the median value of 0.49. The 16th and 84th percentiles are -0.17 and 0.81 respectively.

After performing the modelling process, we convert the obtained covariance matrix (from equation 5.10) into a correlation matrix and show the distribution of the correlation between the hydrostatic mass and the gas mass in Figure 5.9. The median is 0.49 and the 16th and 84th percentiles are -0.17 and 0.81 respectively, which is evidence for a weak positive covariance between the hydrostatic and gas mass, and suggests that we should indeed be modelling this covariance explicitly; however, we also see that it does not significantly impact the model parameters that we are most interested in (primarily $\log_{10}(f_{\rm gas})$ and the bias terms).

As is shown in Figure 5.10 and Table 5.1, we obtain very similar results as the model presented in §5.3.2 (i.e. the smooth transition model). We also compare the posteriors of the global parameters after the modelling process is completed at 1 R_{500} in Table 5.2. We see that by adding covariance to the model, only the parameters describing the intrinsic scatter of the gas mass and X-ray hydrostatic change significantly. We find that κ_C decreases slightly, and both δ_X and $\delta_{\rm mgas}$ increase. This indicates that we should indeed model the covariance between the hydrostatic mass and the gas mass, with this covariance being a necessary component of the model. As the hydrostatic and gas mass are correlated, such that an overestimated hydrostatic mass would correspond to an overestimated gas mass, the scatter between each and the true mass

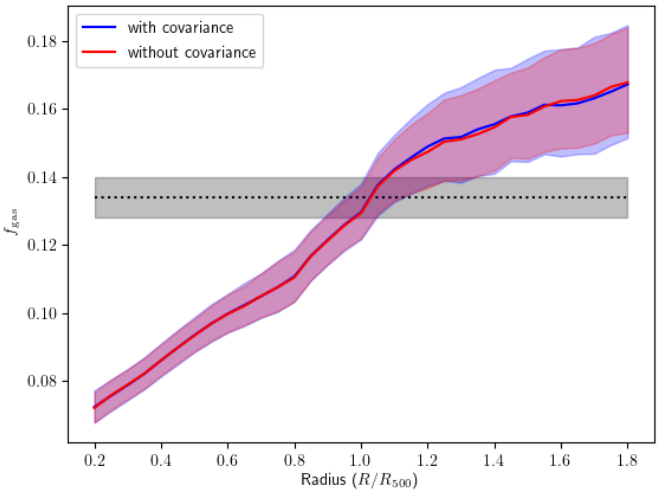

Figure 5.10: We show the $f_{\rm gas}$ as a function of radius for when we drop the X-ray hydrostatic mass profiles on by one and update the prior on κ_C each step after the first X-ray mass is dropped (described in §5.3.2) for when covariance is used (§5.3.3) and when it is not used (§5.3.2). 1 σ errors are shown by the shaded region. We show the universal gas fraction as a black dotted line, with the 1σ errors shown as shaded.

is suppressed when no covariance is modelled, and when it is modelled we see that this scatter is in fact higher.

We show corner plots for all global model parameters after the modelling process is completed at 1 R_{500} in Figure 5.11. In these corner plots we see that there exists a degeneracy between $\log_{10}(f_{\rm gas})$ and κ_X as expected. However this degeneracy is subtle, because we have a tight prior on the κ_X parameters (which still has a strong effect at the 1 R_{500} radius, as most X-ray masses are still not extrapolated, and there is already quite a tight prior on κ_C at that radius). The expected degeneracy between $\log_{10}(f_{\rm gas})$ and κ_C is also apparent. From the corner plots we also see that the relation between δ_X and $\delta_{\rm mgas}$ changes; when the $m_{\rm gas}$ and hydrostatic masses are modelled independently, we see no degeneracy between the corresponding scatter term δ_X and $\delta_{\rm mgas}$, but when the covariance between the $m_{\rm gas}$ and hydrostatic masses *are* modelled, we see that there is indeed a degeneracy between these two scatter terms.

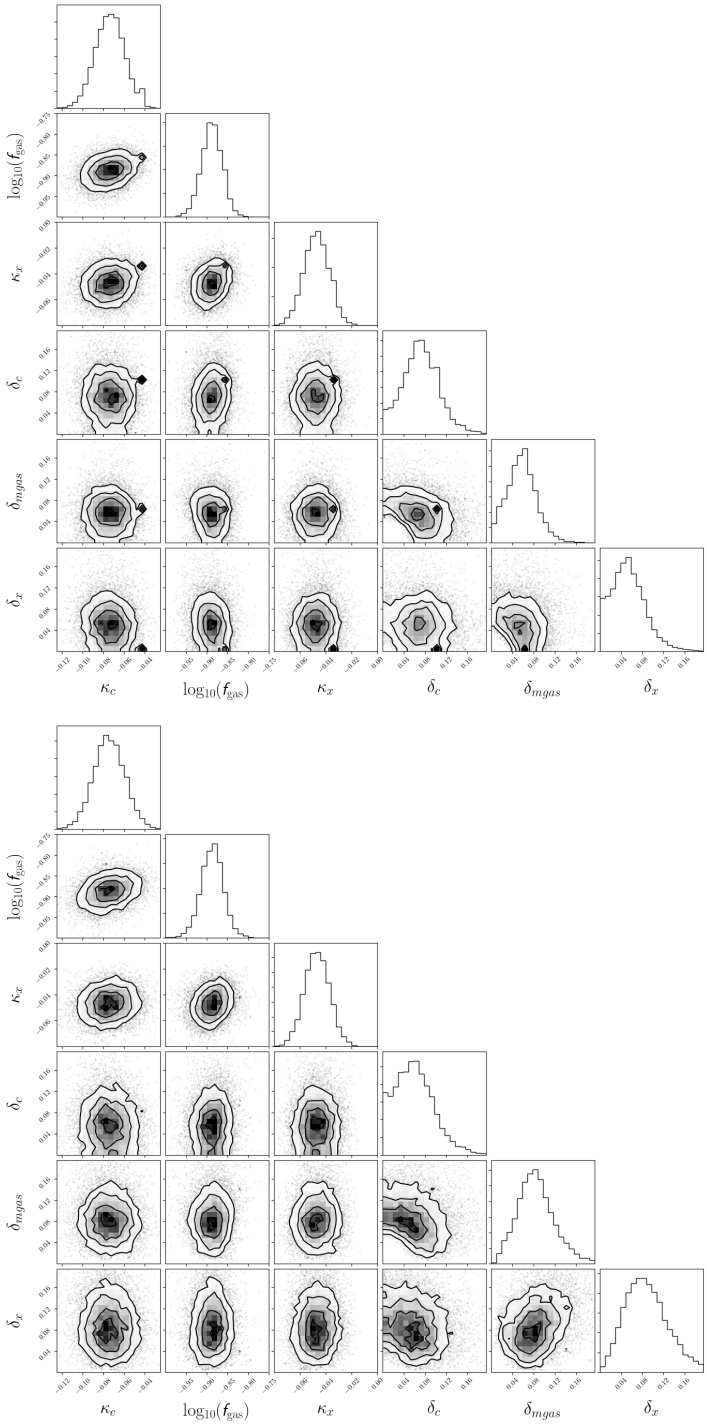

Figure 5.11: We show the corner plot for when we don't include (top panel) and do include (bottom panel) the covariance between the hydrostatic mass and the gas mass in the model setup. All values are given in log space. The model setup we use for not including and including covariance are detailed in §5.3.2 and 5.3.3 respectively. We note that the 'dot' in the corner plots in the top panel is an artefact of the MCMC fitting method, where some walkers over congregate in a part of the parameter space, and does not affect the results (this is clear in the marginal distribution plotted in the top panel).

Model	κ_X	κ_C	$\log_{10}(f_{\rm gas})$	δ_X	δ_C	$\delta_{\rm mgas}$
§5.3.2 model (log space)	-0.047±0.098	-0.073±0.014	-0.888±0.025	0.052±0.034	0.071±0.035	0.057±0.028
§5.3.2 model (linear space)	$0.897^{+0.021}_{-0.020}$	$0.844^{+0.028}_{-0.028}$	$0.129^{+0.008}_{-0.007}$	$13^{+9}_{-8}\%$	$18^{+10}_{-9}\%$	$14^{+8}_{-7}\%$
§5.3.3 model (log space)	-0.047±0.010	-0.075±0.015	-0.887±0.027	0.086±0.044	0.058±0.37	0.082±0.037
§5.3.3 model (linear space)	$0.897^{+0.020}_{-0.020}$	$0.842^{+0.030}_{-0.029}$	$0.130^{+0.008}_{-0.008}$	$22^{+13}_{-12}\%$	$14^{+10}_{-9}\%$	$21^{+11}_{-10}\%$

Table 5.2: We show the median values of the (global) parameters in our model that are confined by the data from all of our clusters simultaneously at 1 R_{500} (if those clusters have temperature data measured at this radius) for the modelling setup in §5.3.2 (i.e. the smooth transition model) and §5.3.3 (i.e. the smooth transition with covariance model) where the difference in modelling setups is that in §5.3.3 we include covariance between the hydrostatic mass and the gas mass. We show the model parameters in both log space and linear space: the log space rows show the output from the model, and the linear space rows show the model output parameter values converted into linear space for ease of interpretation. All errors are 1 σ errors.

5.3.4 Dropping A1835

As can be seen in Figure 5.10 there is a slight flattening of the $f_{\rm gas}$ profile arond 0.6 - 0.8 R_{500}, which is due to the inclusion of A1835 in our cluster sample, due to its unphysical X-ray mass profile (discussed in §4.7). We therefore use the same modelling setup as in §5.3.3 (i.e. the smooth transition with covariance model), now dropping A1835 from out cluster sample. We show our results in Figure 5.12 and note that the results are entirely consistent with when A1835 is included in the cluster sample.

5.3.5 Changing the value of the hydrostatic bias

Thus far, in all of our different approaches for the modelling, we have used the same prior on the hydrostatic bias term, κ_X, i.e. $\kappa_X \sim \mathcal{N}$ (-0.046, 0.01). We now explore the impact of placing different priors on κ_X, using the modelling approach otherwise specified in §5.3.3 (i.e. the smooth transition with covariance model). We vary the centre of the prior on κ_X to be 0, -0.046, -0.097 and -0.155 which corresponds to a negative hydrostatic bias of 0%, 10%, 20% and 30% respectively. We choose these values as they are in the range of values found from literature (e.g. Nagai et al., 2007; Rasia et al., 2012; Eckert et al., 2019).

As is shown in Figure 5.13, the models scale simply with κ_X. These results suggest that a κ_X value representing $\sim$ 20 - 30% hydrostatic bias is necessary in order to ensure that the $f_{\rm gas}$ value at 1.8 R_{500} does not exceed the cosmic $f_{\rm gas}$ value of 0.134.

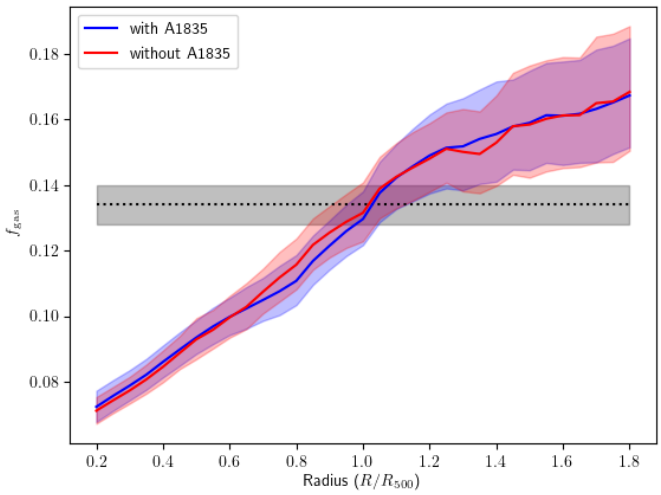

Figure 5.12: We show the $f_{\rm gas}$ as a function of radius for when we drop A1835 from our cluster sample (discussed in §5.3.4) and for when we do include it (modelled using the setup as in §5.3.3, i.e. the smooth transition with covariance model). 1 σ errors are shown by the shaded region. We show the universal gas fraction as a black dotted line, with the 1σ errors shown as shaded.

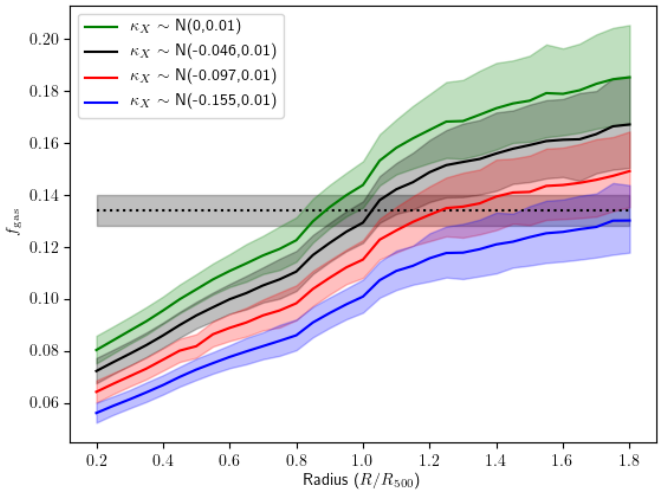

Figure 5.13: We show the $f_{\rm gas}$ as a function of radius for when we drop the X-ray hydrostatic mass profiles on by one and update κ_C each step and explicitly model covariance between the X-ray hydrostatic mass and gas mass (i.e. the smooth transition with covariance model, described in §5.3.3) for when we vary κ_X as in §5.3.5. 1 σ errors are shown by the shaded region. We show the universal gas fraction as a black dotted line, with the 1σ errors shown as shaded.

5.3.6 Modelling the prior on κ_C as in Serra 2011

In the previous analyses, the bias on the X-ray hydrostatic masses has been constrained with an informative prior, with the bias on the caustic masses being constrained by the data. However, bias on the caustic masses has been modelled in simulations (Serra et al., 2011), so here we investigate adopting the calibration of the caustic bias from those simulations and correspondingly loosening the prior on the hydrostatic bias. We do this following the same approach as in §5.3.3 (i.e. the smooth transition with covariance model), now relaxing the prior on κ_X from $\kappa_X \sim \mathcal{N}$ (-0.046, 0.01) to $\kappa_X \sim \mathcal{N}(0, 1)$, so it is now a very loose prior. The prior on κ_C is changed such that it is now the tight prior required in our model to be able to constrain $f_{\rm gas}$. The values we use for the $\kappa_C \sim \mathcal{N}$ (mean, sigma) in the prior are from Fig 12 in Serra et al. (2011), with the sigma being the 68% line in that figure divided by $\sqrt{n}$ where n = 100 as this is the sample size used in Serra et al. (2011); we note that the caustic bias changes with radius, and as such that prior on κ_C changes with radius. Figure 5.14 shows the $f_{\rm gas}$ profile derived using the prior on the caustic bias from Serra et al. (2011), compared with when the hydrostatic bias is constrained with an informative prior (as in §5.3.3); we see that using the Serra et al. (2011) results as the prior on κ_C gives a significantly higher $f_{\rm gas}$ at all radii. One explanation for this is that the caustic mass is significantly underestimated (see Chapter 4) compared to expectations from (Serra et al., 2011). Alternatively, or in addition to this, the super-cosmic values of $f_{\rm gas}$ found at large radii could be seen as evidence for significant levels of clumping of the ICM at the cluster outskirts, which would lead to an overestimation of the gas mass.

5.4 Constraints from the full CHeCS sample

We have thus far focussed on the sub-sample of 14 clusters with high Ngal values as defined in Chapter 4, where we found that these are the clusters for which the systematics on the mass determinations appeared to be minimised. Specifically, our results suggested that the caustic masses calculated for the other clusters with lower Ngal values were underestimating the true cluster mass. However, in Chapter 4 it is only the total mass that is being underestimated by the caustic mass method, and there is no suggestion that the *shape* of the caustic mass profiles themselves are wrong. In fact, for the radial range where both the X-ray and caustic masses can be measured, the X-ray divided by caustic mass profiles are flat (see Figure 4.8), suggesting that both methods agree on the shape of the cluster mass profiles (again, at least for the radial range

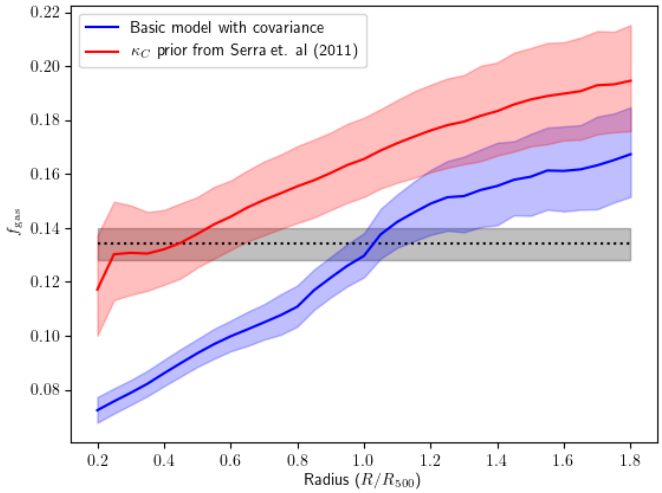

Figure 5.14: We show the $f_{\rm gas}$ as a function of radius for our modelling setup as in §5.3.3 (i.e. the smooth transition with covariance model) compared to the modelling setup in §5.3.6 when we loosen the prior on κ_X and tighten it on the κ_C prior and use the values for the κ_C prior mean as in Serra et al. (2011). 1 σ errors are shown by the shaded region. We show the universal gas fraction as a black dotted line, with the 1σ errors shown as shaded.

where both can be measured). This implies that if we can scale the caustic mass profiles for all 44 clusters appropriately, by using the X-ray mass profiles as previously, it should not matter that the caustic masses for the lower Ngal clusters underestimate the true total cluster mass, and so we explore using all 44 clusters in our analysis. To do this, we use the modelling setup as presented in §5.3.3 (i.e. the smooth transition with covariance model), where a tight prior of 10% hydrostatic bias is assumed, and the covariance between the X-ray hydrostatic mass and the gas mass is modelled.

We show the extent of observational data for all 44 clusters in Figures 5.15 and 5.16, with Figure 5.16 showing that we have data for the surface brightness profiles up to 1.8 R_{500} for the majority of the 44 clusters, so are still not extrapolating the gas mass for most clusters in our sample at this point (as discussed for the 14 cluster sample in §5.3.1).

5.4.1 Complete CHeCS sample

In Figure 5.17 we show the gas fraction as a function of radius for all 44 clusters in the CHeCS sample. We also present $f_{\rm gas}$ values at specific radii in Table 5.3. The $f_{\rm gas}$ values agree within

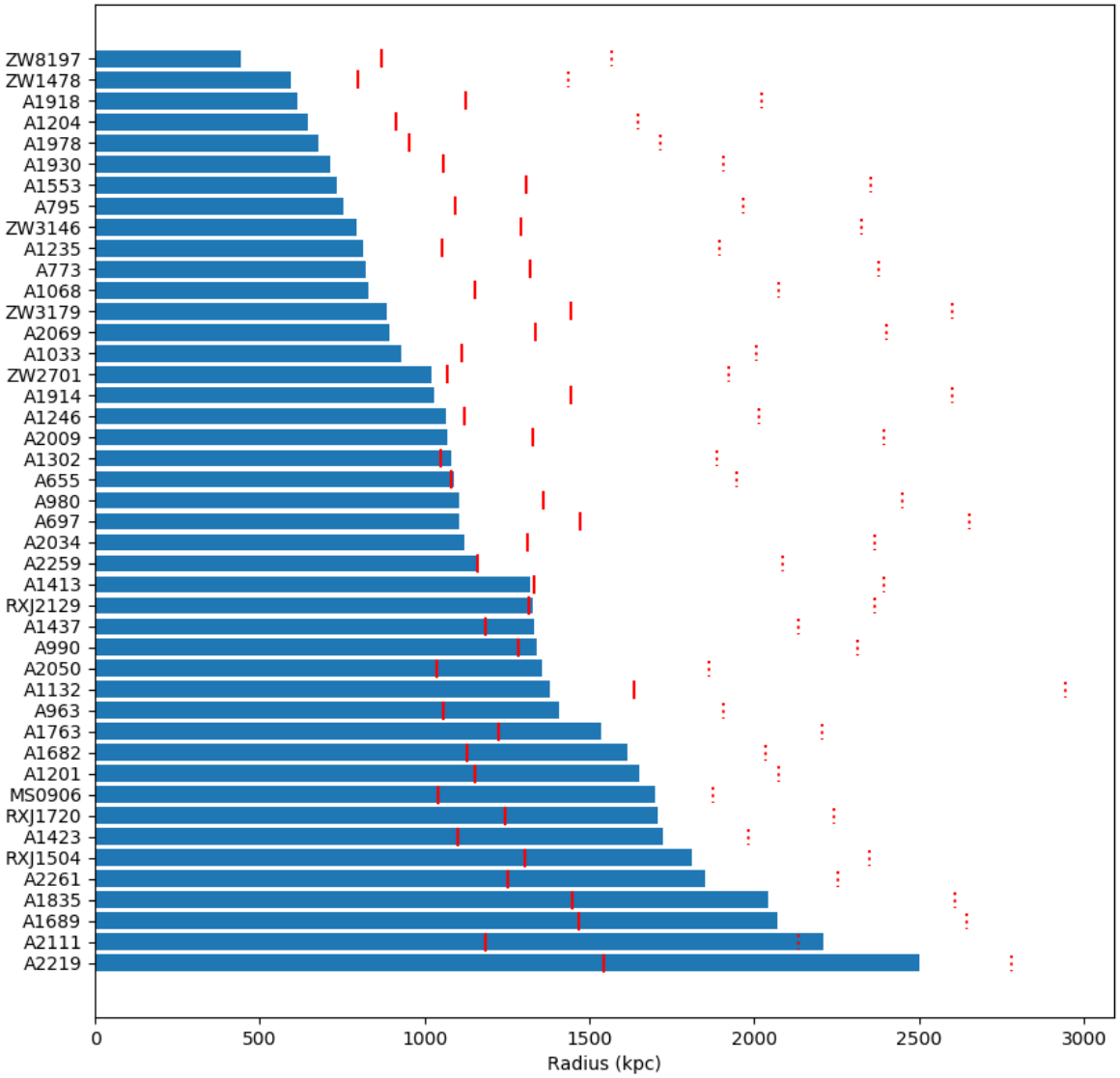

Figure 5.15: We show the radius out to which there is observational data for each cluster's temperature profile as a blue bar for all 44 clusters, which is used to obtain X-ray hydrostatic mass profile (see §3.2 and §4.3.8). Outside the outer radius, a cluster's hydrostatic mass profile is extrapolated. We plot each cluster's R_{500} as a solid red line, and $1.8 \times R_{500}$ as a red dotted line.

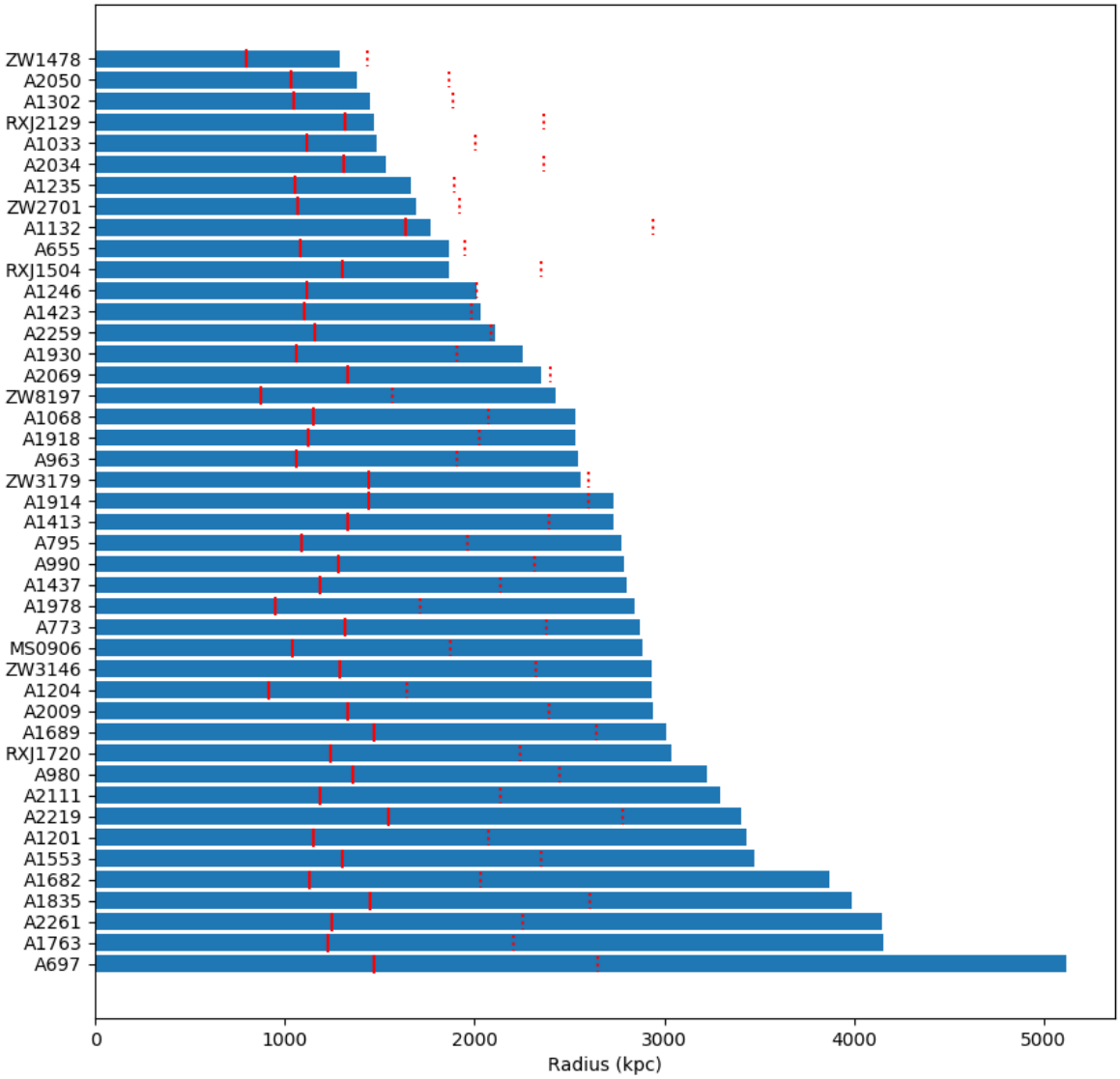

Figure 5.16: We show the radius out to which there is observational data for each cluster's surface brightness profile as a blue bar for all 44 clusters, which is used to obtain the gas mass profile (see §3.2 and §4.3.9). Outside the outer radius, a cluster's gas mass profile is extrapolated. We plot each cluster's R_{500} as a solid red line, and $1.8 \times R_{500}$ as a red dotted line.

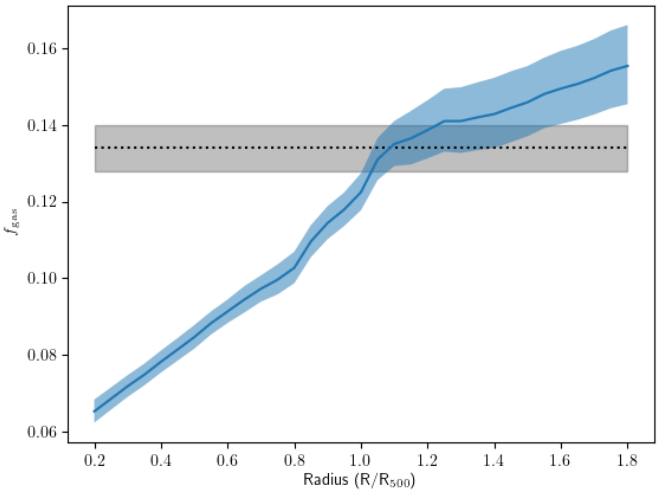

Figure 5.17: We show the $f_{\rm gas}$ as a function of radius using the modelling setup as in §5.3.3 (i.e. the smooth transition with covariance model), but for the 44 cluster sample as described in §5.4 as a blue line. 1 σ errors are shown by the shaded region. We show the universal gas fraction as a black dotted line, with the 1σ errors shown as shaded.

error at 1, 1.4 and 1.8 R_{500} with the results for when just 14 clusters were used in the same model setup (see §5.3.3). We do however see that that the average $f_{\rm gas}$ values are slightly lower when the sample of 44 clusters is used. One explanation for this is that because the 14 cluster sample consists of those clusters with the highest Ngal value, which in general correspond to the more massive clusters (see Figure 4.13); the median X-ray hydrostatic mass at 1 Mpc for the 14 cluster sample is 5.5 $\times 10^{14}$M$_{\odot}$ for the 14 cluster sample (with a 16 - 84 percentile range of 4.1 - 8.1 $\times 10^{14}$M$_{\odot}$), and 4.9 $\times 10^{14}$M$_{\odot}$ for the 44 cluster sample (with a 16 - 84 percentile range of 3.5 - 7.1 $\times 10^{14}$M$_{\odot}$). Because the effect of baryon depletion via gas removal from clusters due to AGN activity (Planelles et al., 2013) is stronger in lower mass clusters, the average $f_{\rm gas}$ in the 44 cluster sample is expected to be lower due to the inclusion of these lower mass clusters.

5.4.2 Split by cool core state

We now split the 44 cluster sample by cool-core (CC) state, as defined in Table 4.2, such that we have 10 RCC clusters and 34 NRCC clusters. We show the results for these two subsamples in Figure 5.18 and Table 5.3, and find that the $f_{\rm gas}$ values are consistent above 0.8 R_{500}. Below this radius, we see that the RCC clusters have a higher gas fraction than the NRCC clusters, which

	All 44 clusters	RCC clusters	NRCC clusters	low mass clusters	mid mass clusters	high mass clusters	14 high Ngal clusters
$f_{\rm gas}$ at R_{500}	$0.122^{+0.005}_{-0.005}$	$0.121^{+0.017}_{-0.015}$	$0.121^{+0.006}_{-0.005}$	$0.111^{+0.014}_{-0.012}$	$0.127^{+0.006}_{-0.006}$	$0.117^{+0.017}_{-0.015}$	$0.130^{+0.008}_{-0.008}$
$f_{\rm gas}$ at 1.4 R_{500}	$0.143^{+0.009}_{-0.009}$	$0.150^{+0.051}_{-0.038}$	$0.142^{+0.009}_{-0.008}$	$0.125^{+0.020}_{-0.017}$	$0.150^{+0.012}_{-0.011}$	$0.130^{+0.019}_{-0.017}$	$0.156^{+0.016}_{-0.015}$
$f_{\rm gas}$ at 1.8 R_{500}	$0.155^{+0.011}_{-0.010}$	$0.157^{+0.055}_{-0.041}$	$0.154^{+0.010}_{-0.009}$	$0.138^{+0.024}_{-0.020}$	$0.168^{+0.020}_{-0.017}$	$0.140^{+0.019}_{-0.017}$	$0.167^{+0.017}_{-0.016}$

Table 5.3: We show the value of $f_{\rm gas}$ at different radii for the modelling setup as shown in §5.3.3 (i.e. the smooth transition with covariance model) for different subsamples of clusters from the 44 clusters used in Chapter 4 (i.e. the full CHeCS sample). We also show the result for the 14 high Ngal cluster sample (as presented in Table 5.1) for comparison in the final column. The subsamples are described in §5.4. All errors are 1 σ errors.

is somewhat expected, as RCC clusters have very high gas densities in their central regions due to the cooling in the core of the cluster. Alternatively, the lower value for the gas fraction for the NRCC clusters could be due to increased non-thermal pressure support in these disturbed clusters. Landry et al. (2013) measured the gas fraction out to R_{500} for a sample of 35 clusters, finding no difference in the $f_{\rm gas}$ values between relaxed and disturbed clusters at small radii, however our results are entirely consistent, within errors, with those from Landry et al. (2013).

5.4.3 Split by hydrostatic mass

We also split the 44 cluster sample into three subsamples according to their X-ray hydrostatic mass at 1 Mpc. We show the X-ray hydrostatic masses at 1 Mpc for all 44 clusters in Figure 5.19, and the subsamples are as follows:

- low mass subsample: ZW1478, A655, MS0906, A795, ZW2701, A963, A1033, A1068, A1204, A1235, A1302, A1930, A1978, A2050, ZW8197
- mid mass subsample: A990, A1201, A1246, A1423, A1437, A1682, A1763, A1918, A2009, A2069, A2111, A2259, RXJ1720, A2261, RXJ2129
- high mass subsample: A697, A773, A980, ZW3146, ZW3179, A1132, A1413, A1553, A1689, A1835, A1914, RXJ1504, A2034, A2219

In Figure 5.20 we show the gas fraction as a function of radius for each mass bin, and show the $f_{\rm gas}$ values at specific radii in Table 5.3. We see that the low mass clusters have the lowest $f_{\rm gas}$ values (though consistent with the other mass subsamples at mid to large cluster radii), as expected, as any AGN activity that blows out the gas in the cluster will have the strongest effect in low mass clusters (McCarthy et al., 2010), and indeed we see the that the $f_{\rm gas}$ profile for the low mass clusters is also lower at smaller radii. However, we see that the mid mass clusters have a consistently higher $f_{\rm gas}$ profile than the high mass clusters, which is unexpected, but we

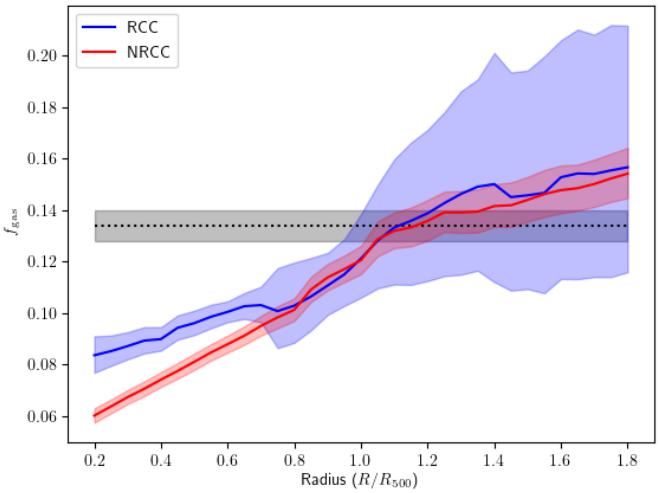

Figure 5.18: We show $f_{\rm gas}$ as a function of radius for the 44 clusters when split by their cool core state. The blue line shows the results for RCC clusters, and the red line shows the results for NRCC clusters. 1 σ errors are shown by the shaded region. We show the universal gas fraction as a black dotted line, with the 1σ errors shown as shaded.

note that their $f_{\rm gas}$ profiles are actually consistent, so there is not a strong difference between the two subsamples. In summary, the $f_{\rm gas}$ profiles for the different mass subsamples are largely consistent, with a weak suggestion of mass segregation.

In Figure 5.21 we show the individual $f_{\rm gas}$ values at 1 Mpc of all 44 clusters in our sample versus their measured X-ray hydrostatic mass at 1 Mpc. We see that there is no clear trend in $f_{\rm gas}$ with mass, which agrees with the results from the $f_{\rm gas}$ profiles in Figure 5.20.

5.5 Discussion

In this work, we utilise the low scatter and reliability of X-ray hydrostatic mass measurements of a sample of 14 clusters to 'anchor' their caustic mass profiles. The caustic mass profiles are then used, in addition to the gas mass profiles, to probe the gas fraction at large cluster radii where the caustic mass profiles are still expected to be a reliable probe of the shape of the cluster mass profiles. We try a number of modelling approaches throughout the paper, with the optimal method, the smooth transition with covariance model, presented in §5.3.3. In this setup, termed the 'smooth-transition with covariance' model, we 'anchor' the caustic masses

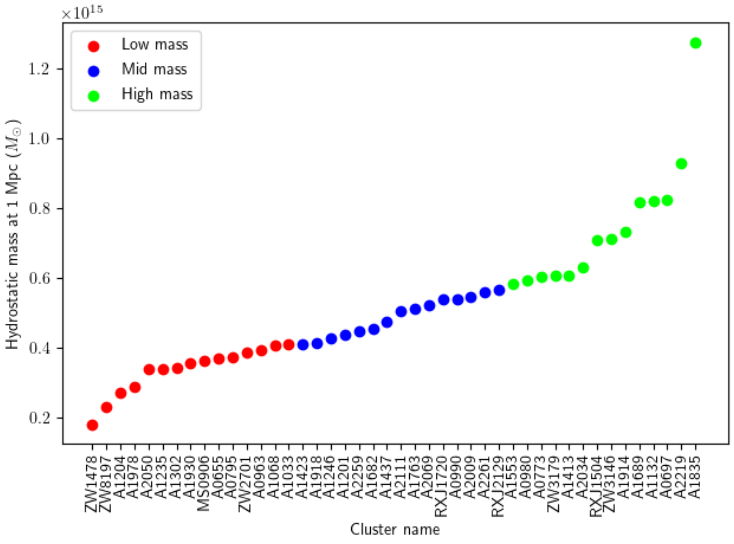

Figure 5.19: We show the X-ray hydrostatic masses for all 44 clusters at 1 Mpc.

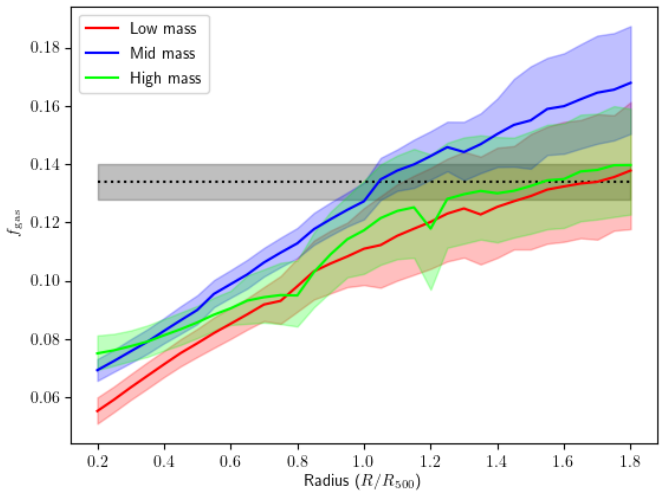

Figure 5.20: We show $f_{\rm gas}$ as a function of radius for the 44 clusters when split by their X-ray hydrostatic mass at 1 Mpc. The red, blue and green line shows the results for the low, mid, and high mass clusters respectively. 1 σ errors are shown by the shaded region. We show the universal gas fraction as a black dotted line, with the 1σ errors shown as shaded.

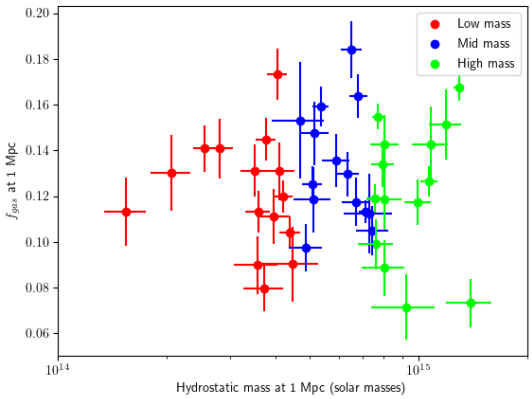

Figure 5.21: We show the $f_{\rm gas}$ values (calculated as observed $m_{\rm gas}$ / observed X-ray hydrostatic mass) versus observed X-ray hydrostatic mass, for all 44 clusters for the three cluster subsamples split by their observed X-ray hydrostatic mass at 1 Mpc. All quantities are calculated at 1 Mpc.

using the X-ray mass information for each cluster until the X-ray mass profile for that cluster becomes extrapolated, at which point we drop that cluster's X-ray mass profile information from the modelling setup, and do this on a cluster by cluster basis. For the radius regime where any X-ray mass information is used, we have a tight prior on the X-ray bias, κ_X, and as soon as the first cluster's X-ray mass profile starts being extrapolated, we tighten the previously loose prior on the caustic mass, κ_C, to the posterior calculated at the previous radius in the modelling process. One feature in the 'smooth-transition with covariance' model setup is that we model the covariance between the X-ray mass and gas mass, as these are expected to be correlated to some extent. We explore the effect of varying the tight prior on the X-ray mass bias in §5.3.5, but in the smooth transition with covariance model, we use a negative bias of 10%.

Inspecting the results obtained using the smooth transition with covariance model, we find that the $f_{\rm gas}$ value at 1.4 R_{500} is 0.156±0.016 and at 1.8 R_{500} is 0.167±0.017 (see Table 5.1), which is inconsistent with the universal gas fraction of 0.134±0.006 (see Figure 5.10). This suggests that there is some gas clumping that starts around 1.4 R_{500}, and the clumping increases out to larger radii, or that the hydrostatic bias is in fact larger at this radius than what we account for in our prior of 10% on the hydrostatic bias for these results.

Though we don't attempt to measure directly the exact level of clumping in the outskirts of the clusters in our sample (such as in, for example, Eckert et al., 2015), we can infer an average level of clumping. The clumping factor, C, can be defined as $C = \langle\rho^2_{\rm gas}\rangle/\langle\rho_{\rm gas}\rangle^2$, where $\rho_{\rm gas}$

is the gas density and $\langle\rangle$ denotes the mean inside spherical shells. For a cluster with clumping factor C, the gas mass, as estimated from the surface brightness profile, as in this work, will be overestimated by a factor of $\sqrt{C}$. Thus, if the high value of $f_{\rm gas}$ at 1.4 and 1.8 R_{500} measured using the smooth transition with covariance model is solely due to clumping, we would infer a clumping factor of 1.36$\pm$0.15 and 1.55$\pm$0.17 respectively. Focussing on the inferred value of C at 1.4 R_{500} ($\sim R_{200}$) of 1.36, this compares favourably with the clumping value of 1.3 at R_{200} found from work using simulations (Nagai & Lau, 2011). Our inferred clumping factor also agrees with the clumping factor of 1.2 at R_{200} from (Eckert et al., 2013), where they used observational data (X-ray from *ROSAT* and SZ from *Planck*) for a sample of 18 clusters to measure $f_{\rm gas}$ at this radius. Our value for C strongly disagrees with the inferred value of C from Simionescu et al., 2011 of 3-4, where X-ray data (from *Suzaku*) was used to study gas fractions in the Perseus cluster (though we note that their results may not be representative of the cluster as a whole, as they only looked at two narrow azimuthal ranges of the cluster). We note that the clumping factors that we find are, of course, dependent on the choice of prior on the X-ray hydrostatic mass that we use. For the results described above, we place a tight prior of 10% on the hydrostatic bias in our modelling, which is on the lower range of expected values for this bias. Of course, assuming a higher level of X-ray bias would lead to a lower value of clumping. We do explore the effect of changing the value of the tight prior placed on the hydrostatic bias, κ_X, in our modelling setup in §5.3.5, and find that in order to obtain a value of $f_{\rm gas}$ at 1.8 R_{500} that matches the universal gas fraction value, we would need to place a prior of $\sim$ 20-30% on the hydrostatic bias (see Figure 5.13), which is on the high end of what is generally found from literature, both from simulations (e.g. Nagai et al., 2007; Lau et al., 2009; Vazza et al., 2018) and observational work (e.g. Siegel et al., 2018; Eckert et al., 2019). Thus, despite the exact level of clumping inferred being highly dependent on the choice of prior on the hydrostatic bias in our model, we conclude that there is likely some level of clumping present in the cluster outskirts.

Our results also give us some insight into the value of the baryon depletion factor, $Y_{\rm b}$, and how it varies with radius. We see that for all modelling setups explored in this work, the gas fraction increases with radius, in line with the vast majority of work on gas fractions in clusters. Specifically, for the optimal modelling setup used in §5.3.3, we see that the gas fraction increases from roughly half the universal gas fraction at 0.2 R_{500} until it reaches the universal gas fraction at around R_{500}. This would imply that the value of $Y_{\rm b}$ rises from 0.5 to 1 over this radial range, which agrees with results from simulations that the baryon depletion factor increases with radius

(e.g. Planelles et al., 2013).

We also explored placing a tight prior that varies with radius on the caustic bias, and relaxing the tight prior on the X-ray bias. We compare this method to the method where the tight prior is placed instead on the X-ray bias, and our results are presented in Figure 5.14. We see that the $f_{\rm gas}$ profile is higher at all radii when the tight prior is placed on the caustic bias. The difference between the two approaches (i.e. placing a tight prior on the caustic bias based on values for this bias from Serra et al., 2011 versus placing a tight bias on the hydrostatic bias of 10%) is strongest at lower radii. The caustic mass is expected to be overestimated at smaller radii, up to 70% (Serra et al., 2011), and we have placed this knowledge of the caustic bias from (Serra et al., 2011) into our tight prior on the caustic bias. The high values for $f_{\rm gas}$ found at small radii for the approach where we place a tight prior on the caustic bias suggests that the caustic mass is actually not overestimated in the lower radius regime as strongly as suggested in (Serra et al., 2011). Furthermore, as the $f_{\rm gas}$ value is higher than the universal gas fraction from $\sim 0.5\ R_{500}$ and continues increasing out to larger radii, we could interpret this result as indicating that the caustic mass is in fact underestimated compared to expectations not just at small cluster radii, but even at larger radii, where it is expected, at least from the simulations of (Serra et al., 2011), to be an unbiased estimator of mass.

Our main results do not depend on the absolute value of the caustic mass profiles, but instead rely on their shape, as they are 'anchored' by the X-ray mass profiles at lower radii. However, given the values we find for $f_{\rm gas}$ when we use a tight prior on the caustic bias, which indicate that the assumption of little to no caustic bias at large radii from (Serra et al., 2011) may be wrong, it could also be the case that the shape of the caustic mass profiles are not as reliable as previously believed, which is a limiting factor on our results, thus, this work further highlights the need to better understand the caustic mass method and the level of bias present in this method.

We note that our sample is an X-ray selected sample, and thus will preferentially select systems that have a higher gas fraction than average, as these systems will be brighter than average for a given mass. However, as the scatter between the total cluster mass and the gas mass at a fixed cluster mass is low (Mantz et al., 2016, 2018a), this effect should not be strong, and hence we do not account for this effect explicitly in our analysis.

5.6 Conclusion

In this chapter we use caustic mass profiles and gas mass profiles to measure $f_{\rm gas}$ out to large radii. We incorporate the accuracy and precision of our X-ray hydrostatic mass profiles into the analysis by calibrating the caustic mass profiles at smaller cluster radii with our X-ray hydrostatic mass profiles where the X-ray hydrostatic mass profiles are not extrapolated.

Our main results, which focus on the 14 clusters in the high Ngal subsample, give a value of $f_{\rm gas}$ at 1.8 R_{500} of 0.167$\pm$0.017 using our optimal model (defined in §5.3.3), suggesting that there is some clumping of the ICM at this radius.

6

Conclusions

In this thesis, I use *Chandra* data to study galaxy clusters throughout, however each of the three projects that I carried out during my PhD had quite different foci, and I summarise the results from each project in §6.1. I then describe any further work that could be done, with the results of my thesis in mind, in §6.2.

6.1 Summary of work

The first project in my PhD, presented in Chapter 2, was as part of the XXL survey (Pierre et al., 2016), an *XMM-Newton* cluster survey with 10ks exposures covering two 25 deg^2 contiguous areas. I followed up the 21 of the 24 most likely XXL cluster candidates at redshift $z > 1$ using *Chandra* (Logan et al., 2018), utilizing its superior angular resolution, to vet this sample for cluster candidates that were actually misclassified AGN, and identify clusters that were contaminated by previously unresolved AGN. This work enhanced the value of the high redshift sample of XXL, by confirming real clusters - seven were rejected and one was identified as being significantly contaminated by previously unresolved AGN.

I then present my work on the calibration of the X-ray hydrostatic mass estimation method in Chapter 4. Using *Chandra* data of a complete sample of 44 clusters that had accompanying caustic mass measurements, I developed a full hydrostatic mass code that used Affine-Invariant

Monte Carlo Markov Chain (MCMC) for the modelling, to estimate the X-ray mass. I compared the X-ray masses to the caustic masses, using a hierarchical Bayesian framework, in order to probe the hydrostatic bias. After making this comparison, it was clear that there was an issue with one (or both) of the mass measurement methods as the X-ray to caustic mass ratio was much higher than expected. After some investigation, I found that the caustic mass method significantly underestimated the mass of clusters with a low number of galaxy redshifts (Ngal). This was expected, albeit at a lower level, from simulations Serra et al. (2011). I then focussed on the 14 clusters with the highest Ngal values in our sample, for which I found the X-ray to caustic mass ratio to be 1.12 ± 0.11, suggesting a low to zero value for the hydrostatic bias. I note that it may be the case that this unexpectedly strong dependence of the caustic bias with Ngal has some effect even in the high Ngal cluster subsample, leading to this still rather high value for the X-ray to caustic mass ratio in this subsample. Additionally, by comparing the mass profiles, I found that this bias did not have a strong dependence on the radius, ignoring the low radius regime, where the caustic mass measurement is biased high (Serra et al., 2011).

Unlike the X-ray hydrostatic mass method, the caustic mass method enables accurate estimation of the mass profile out to large cluster radii. By combining the caustic mass profile with a gas mass profile which *can* by measured to large radii with the X-ray data, I probed the gas mass fraction (f_{gas}) in the cluster outskirts in Chapter 5. I explore a variety of different modelling setups, and the results from the optimal setup, where I set a tight prior of 10% on the X-ray hydrostatic bias show the gas fraction to reach the universal gas fraction at $\sim R_{500}$, and continue increasing until 1.8 R_{500}, the largest radius that was studied, suggesting that there is some level of clumping in the cluster outskirts.

6.2 Future work

The work presented in Chapter 2 that used *Chandra* follow-up observations of high-redshift XXL cluster candidates to probe the level of AGN contamination in these clusters led to a successful *XMM-Newton* proposal to further follow up three of the brightest clusters in the vetted sample that we presented in that chapter. The clusters are XLSSC 634, 3XLSS J021325.0-042000 and XLSSC 048, and in addition there existed archival observations of XLSSC 072, XLSSC 029, and XLSSC 122 (see Table 2.1 for information on these clusters). The length of the observations were calculated so as to be able to measure robust ICM properties of the high

redshift clusters (mass, temperature and luminosity). The results of this work will be presented in Duffy et al (in prep.). This work also showed the utility of *Chandra* to vet high-redshift clusters detected with lower resolution X-ray telescopes, and this technique will likely be used to follow up interesting systems detected with *eROSITA* in the coming years. The work highlights the need for a high resolution X-ray telescope to complement *XMM-Newton*, and it will be a shame that *Chandra* will almost certainly not be functioning when *ATHENA*, with an expected resolution similar to *XMM-Newton*, is launched.

Our results in Chapter 4 suggest that the hydrostatic bias is small, however, the validity of our results depends on how much we trust the caustic mass method. We found issues with the caustic mass method (specifically that it was underestimating the true cluster mass), with a strong dependence of the caustic bias with the numbers of member galaxies for which spectroscopic redshifts have been measured per cluster. This level of dependence was significantly higher than what was previously believed from work using observations (Serra et al., 2011), and a more detailed investigation into this dependence using observational data would be an important next step. One way in which this dependence could be investigated is to select clusters that have large numbers of member galaxies with measured spectroscopic redshifts, and randomly subsample cluster members from each cluster, and re-calculate their mass using this lower number of spectroscopic redshifts per cluster, to see how strong this dependence is.

In Chapter 5, where we measure the gas fraction in the cluster outskirts using both X-ray hydrostatic and caustic mass measurements of the clusters in our sample, the strength of our conclusions regarding the level of clumping of the ICM in the cluster outskirts is largely dependent on the level of hydrostatic bias that we choose to model. Looking to the future, with the advent of *ATHENA*, a high sensitivity X-ray telescope with low instrumental background, we should be able to directly measure the level of clumping in the cluster outskirts, as well as measure the hydrostatic mass accurately out to larger radii. Until then, we will have to use other indirect methods to infer the level of gas clumping in the cluster outskirts.

This work presented in this thesis demonstrates the utility of *Chandra* as an X-ray telescope in its own right, as well as highlighting the powerful studies of clusters and extra information that can be gained about these systems when using *Chandra* alongside an X-ray observatory that is more sensitive, but with lower spatial resolution (i.e. *XMM-Newton*).

Bibliography

Abell G. O., 1958, ApJS, 3, 211

Adami C., et al., 2011, A&A, 526, A18

Adami C., et al., 2018, A&A, 620, A5

Adelman-McCarthy J. K., et al., 2008, ApJS, 175, 297

Akritas M. G., Bershady M. A., 1996, ApJ, 470, 706

Allen S. W., Schmidt R. W., Ebeling H., Fabian A. C., van Speybroeck L., 2004, MNRAS, 353, 457

Allen S. W., Rapetti D. A., Schmidt R. W., Ebeling H., Morris R. G., Fabian A. C., 2008, MNRAS, 383, 879

Allen S. W., Evrard A. E., Mantz A. B., 2011, ARA&A, 49, 409

Ameglio S., Borgani S., Pierpaoli E., Dolag K., 2007, MNRAS, 382, 397

Anders E., Grevesse N., 1989, Geochim. Cosmochim. Acta, 53, 197

Andreon S., Trinchieri G., Moretti A., Wang J., 2017, A&A, 606, A25

Applegate D. E., et al., 2016, MNRAS, 457, 1522

Armitage T. J., Kay S. T., Barnes D. J., Bahé Y. M., Dalla Vecchia C., 2019, MNRAS, 482, 3308

Arnaud M., Pointecouteau E., Pratt G. W., 2007, A&A, 474, L37

Asplund M., Grevesse N., Sauval A. J., 2005, in Barnes III T. G., Bash F. N., eds, Astronomical

Society of the Pacific Conference Series Vol. 336, Cosmic Abundances as Records of Stellar Evolution and Nucleosynthesis. p. 25

Asplund M., Grevesse N., Sauval A. J., Scott P., 2009, ARA&A, 47, 481

Bahcall J. N., Sarazin C. L., 1977, ApJ, 213, L99

Bartelmann M., Maturi M., 2017, Scholarpedia, 12, 32440

Battaglia N., Bond J. R., Pfrommer C., Sievers J. L., 2012, ApJ, 758, 74

Bauer F. E., Fabian A. C., Sanders J. S., Allen S. W., Johnstone R. M., 2005, MNRAS, 359, 1481

Becker M. R., Kravtsov A. V., 2011, ApJ, 740, 25

Biffi V., et al., 2016, ApJ, 827, 112

Biffi V., Dolag K., Merloni A., 2018, preprint, (arXiv:1804.01096)

Birkinshaw M., 1999, Phys. Rep., 310, 97

Blandford R. D., Saust A. B., Brainerd T. G., Villumsen J. V., 1991, MNRAS, 251, 600

Bleem L. E., et al., 2015, ApJS, 216, 27

Boehringer H., Werner N., 2009, arXiv e-prints, p. arXiv:0907.4277

Böhringer H., et al., 2000, ApJS, 129, 435

Böhringer H., et al., 2001, A&A, 369, 826

Böhringer H., et al., 2004, A&A, 425, 367

Bonamente M., Landry D., Maughan B., Giles P., Joy M., Nevalainen J., 2013, MNRAS, 428, 2812

Borgani S., Guzzo L., 2001, Nature, 409, 39

Borgani S., Kravtsov A., 2011, Advanced Science Letters, 4, 204

Bradt H., Mayer W., Naranan S., Rappaport S., Spada G., 1967, ApJ, 150, L199

Branchesi M., Gioia I. M., Fanti C., Fanti R., 2007, A&A, 472, 727

Bremer M. N., et al., 2006, MNRAS, 371, 1427

Buote D. A., Tsai J. C., 1996, ApJ, 458, 27

Burenin R. A., Vikhlinin A., Hornstrup A., Ebeling H., Quintana H., Mescheryakov A., 2007, ApJS, 172, 561

Butler A., et al., 2017, preprint, p. XXL Paper XVIII (arXiv:1703.10296)

Byram E. T., Chubb T. A., Friedman H., 1966, Science, 152, 66

Cash W., 1979, ApJ, 228, 939

Cavaliere A., Fusco-Femiano R., 1976, A&A, 49, 137

Cavaliere A. G., Gursky H., Tucker W. H., 1971, Nature, 231, 437

Chen Y., Reiprich T. H., Böhringer H., Ikebe Y., Zhang Y. Y., 2007, A&A, 466, 805

Chiappetti L., Fotopolou S., Lidman C., Faccioli L., 2018, A&A, Accepted

Clerc N., Sadibekova T., Pierre M., Pacaud F., Le Fèvre J.-P., Adami C., Altieri B., Valtchanov I., 2012, MNRAS, 423, 3561

Condon J. J., Cotton W. D., Greisen E. W., Yin Q. F., Perley R. A., Taylor G. B., Broderick J. J., 1998, AJ, 115, 1693

Cox A. N., 2000, Allen's astrophysical quantities

Desai S., et al., 2012, ApJ, 757, 83

Diaferio A., 1999, MNRAS, 309, 610

Diaferio A., 2009, arXiv e-prints, p. arXiv:0901.0868

Diaferio A., Geller M. J., 1997, ApJ, 481, 633

Donahue M., et al., 2014, ApJ, 794, 136

Douspis M., Salvati L., Aghanim N., 2019, arXiv e-prints, p. arXiv:1901.05289

Ebeling H., Edge A. C., Bohringer H., Allen S. W., Crawford C. S., Fabian A. C., Voges W., Huchra J. P., 1998, MNRAS, 301, 881

Ebeling H., Edge A. C., Henry J. P., 2001, ApJ, 553, 668

Ebeling H., Edge A. C., Mantz A., Barrett E., Henry J. P., Ma C. J., van Speybroeck L., 2010, MNRAS, 407, 83

Eckert D., et al., 2012, A&A, 541, A57

Eckert D., Ettori S., Molendi S., Vazza F., Paltani S., 2013, A&A, 551, A23

Eckert D., Roncarelli M., Ettori S., Molendi S., Vazza F., Gastaldello F., Rossetti M., 2015, MNRAS, 447, 2198

Eckert D., et al., 2016, A&A, 592, A12

Eckert D., et al., 2019, A&A, 621, A40

Edge A. C., 2001, MNRAS, 328, 762

Edge A. C., Wilman R. J., Johnstone R. M., Crawford C. S., Fabian A. C., Allen S. W., 2002, MNRAS, 337, 49

Eisenhardt P. R. M., et al., 2008, ApJ, 684, 905

Erben T., et al., 2013, MNRAS, 433, 2545

Ettori S., De Grandi S., Molendi S., 2002, A&A, 391, 841

Ettori S., Morandi A., Tozzi P., Balestra I., Borgani S., Rosati P., Lovisari L., Terenziani F., 2009, A&A, 501, 61

Ettori S., Donnarumma A., Pointecouteau E., Reiprich T. H., Giodini S., Lovisari L., Schmidt R. W., 2013, Space Sci. Rev., 177, 119

Evans I. N., et al., 2010, ApJS, 189, 37

Evrard A. E., 1997, MNRAS, 292, 289

Fabian A. C., 1994, ARA&A, 32, 277

Fabian A. C., Daines S. J., 1991, MNRAS, 252, 17P

Fabian A. C., Nulsen P. E. J., 1977, MNRAS, 180, 479

Fabricant D., et al., 2005, PASP, 117, 1411

Faccioli L., et al., 2018, A&A, 620, A9

Fassbender R., et al., 2011, A&A, 527, L10

Foëx G., Böhringer H., Chon G., 2017, A&A, 606, A122

Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, PASP, 125, 306

Fotopoulou S., et al., 2016, A&A, 592, A5, XXL Paper IV

Freeman P., Doe S., Siemiginowska A., 2001, Sherpa: a mission-independent data analysis application. pp 76–87, doi:10.1117/12.447161

Fritz G., Davidsen A., Meekins J. F., Friedman H., 1971, ApJ, 164, L81

Fruscione A., et al., 2006, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series. p. 62701V, doi:10.1117/12.671760

Fusco-Femiano R., Lapi A., 2018, MNRAS, 475, 1340

Galametz A., et al., 2009, ApJ, 694, 1309

Gavazzi R., Soucail G., 2007, A&A, 462, 459

Giacconi R., Murray S., Gursky H., Kellogg E., Schreier E., Tananbaum H., 1972, ApJ, 178, 281

Giacconi R., et al., 1979, ApJ, 230, 540

Giles P. A., Maughan B. J., Birkinshaw M., Worrall D. M., Lancaster K., 2012, MNRAS, 419, 503

Giles P. A., et al., 2016, A&A, 592, A3, XXL Paper III

Giles P. A., et al., 2017, MNRAS, 465, 858

Giodini S., Lovisari L., Pointecouteau E., Ettori S., Reiprich T. H., Hoekstra H., 2013, Space Sci. Rev., 177, 247

Gioia I. M., Maccacaro T., Schild R. E., Wolter A., Stocke J. T., Morris S. L., Henry J. P., 1990a,

ApJS, 72, 567

Gioia I. M., Henry J. P., Maccacaro T., Morris S. L., Stocke J. T., Wolter A., 1990b, ApJ, 356, L35

Gladders M. D., Yee H. K. C., 2005, ApJS, 157, 1

Goodman J., Weare J., 2010, Communications in Applied Mathematics and Computational Science, 5, 65

Gruen D., et al., 2014, MNRAS, 442, 1507

Gursky H., Kellogg E. M., Leong C., Tananbaum H., Giacconi R., 1971, ApJ, 165, L43

Hao J., et al., 2010, ApJS, 191, 254

Hasselfield M., et al., 2013, J. Cosmology Astropart. Phys., 2013, 008

Henson M. A., Barnes D. J., Kay S. T., McCarthy I. G., Schaye J., 2017, MNRAS, 465, 3361

Hilton M., et al., 2010, ApJ, 718, 133

Hilton M., et al., 2018, ApJS, 235, 20

Hinshaw G., et al., 2013, ApJS, 208, 19

Hitomi Collaboration et al., 2016, Nature, 535, 117

Hoekstra H., 2001, A&A, 370, 743

Hoekstra H., 2003, MNRAS, 339, 1155

Hoekstra H., 2007, MNRAS, 379, 317

Hoekstra H., Hartlap J., Hilbert S., van Uitert E., 2011, MNRAS, 412, 2095

Hoekstra H., Bartelmann M., Dahle H., Israel H., Limousin M., Meneghetti M., 2013, Space Sci. Rev., 177, 75

Hoekstra H., Herbonnet R., Muzzin A., Babul A., Mahdavi A., Viola M., Cacciato M., 2015, MNRAS, 449, 685

Hoffman M. D., Gelman A., 2011, arXiv e-prints, p. arXiv:1111.4246

Hubble E., 1929, Proceedings of the National Academy of Science, 15, 168

Hudson D. S., Mittal R., Reiprich T. H., Nulsen P. E. J., Andernach H., Sarazin C. L., 2010, A&A, 513, A37

Israel H., Reiprich T. H., Erben T., Massey R. J., Sarazin C. L., Schneider P., Vikhlinin A., 2014, A&A, 564, A129

Kaastra J. S., Paerels F. B. S., Durret F., Schindler S., Richter P., 2008, Space Sci. Rev., 134, 155

Kaiser N., 1992, ApJ, 388, 272

Kalberla P. M. W., Burton W. B., Hartmann D., Arnal E. M., Bajaja E., Morras R., Pöppel W. G. L., 2005, A&A, 440, 775

Kettula K., et al., 2013, ApJ, 778, 74

Kim M., et al., 2007, ApJS, 169, 401

Koester B. P., et al., 2007, ApJ, 660, 239

Koulouridis E., et al., 2014, A&A, 567, A83

Koulouridis E., et al., 2017, preprint, (arXiv:1709.01926)

Kravtsov A. V., Borgani S., 2012, ARA&A, 50, 353

Kravtsov A. V., Nagai D., Vikhlinin A. A., 2005, ApJ, 625, 588

Kravtsov A. V., Vikhlinin A., Nagai D., 2006, ApJ, 650, 128

Landry D., Bonamente M., Giles P., Maughan B., Joy M., Murray S., 2013, MNRAS, 433, 2790

Lau E. T., Kravtsov A. V., Nagai D., 2009, ApJ, 705, 1129

Leccardi A., Molendi S., 2008, A&A, 487, 461

Leccardi A., Rossetti M., Molendi S., 2010, A&A, 510, A82

Lieu M., et al., 2016, A&A, 592, A4

Limousin M., Morandi A., Sereno M., Meneghetti M., Ettori S., Bartelmann M., Verdugo T., 2013, Space Sci. Rev., 177, 155

Logan C. H. A., et al., 2018, A&A, 620, A18

Mahdavi A., Hoekstra H., Babul A., Bildfell C., Jeltema T., Henry J. P., 2013, ApJ, 767, 116

Mantz A., Allen S. W., 2011, arXiv e-prints, p. arXiv:1106.4052

Mantz A. B., et al., 2014, ApJ, 794, 157, XXL Paper V

Mantz A. B., Allen S. W., Morris R. G., Schmidt R. W., 2016, MNRAS, 456, 4020

Mantz A. B., Allen S. W., Morris R. G., von der Linden A., 2018a, MNRAS, 473, 3072

Mantz A. B., et al., 2018b, A&A, 620, A2

Martínez M., Del Olmo A., Coziol R., Perea J., Focardi P., 2010, in Verdes-Montenegro L., Del Olmo A., Sulentic J., eds, Astronomical Society of the Pacific Conference Series Vol. 421, Galaxies in Isolation: Exploring Nature Versus Nurture. p. 125

Martini P., et al., 2013, ApJ, 768, 1

Martino R., Mazzotta P., Bourdin H., Smith G. P., Bartalucci I., Marrone D. P., Finoguenov A., Okabe N., 2014, MNRAS, 443, 2342

Marziani P., et al., 2017, A&A, 599, A83

Maughan B. J., Reiprich T. H., 2019, The Open Journal of Astrophysics, 2, 9

Maughan B. J., Giles P. A., Randall S. W., Jones C., Forman W. R., 2012, MNRAS, 421, 1583

Maughan B. J., Giles P. A., Rines K. J., Diaferio A., Geller M. J., Van Der Pyl N., Bonamente M., 2016, MNRAS, 461, 4182

Mazzotta P., Rasia E., Moscardini L., Tormen G., 2004, MNRAS, 354, 10

McCarthy I. G., et al., 2010, MNRAS, 406, 822

McCarthy I. G., Schaye J., Bird S., Le Brun A. M. C., 2017, Monthly Notices of the Royal Astronomical Society, 465, 2936

McDonald M., et al., 2012, Nature, 488, 349

McNamara B. R., Nulsen P. E. J., 2012, New Journal of Physics, 14, 055023

Meekins J. F., Fritz G., Chubb T. A., Friedman H., 1971, Nature, 231, 107

Mehrtens N., et al., 2012, MNRAS, 423, 1024

Merloni A., et al., 2012, preprint, (arXiv:1209.3114)

Miralda-Escude J., 1991, ApJ, 380, 1

Mitchell R. J., Culhane J. L., Davison P. J. N., Ives J. C., 1976, MNRAS, 175, 29P

Miyazaki S., Hamana T., Ellis R. S., Kashikawa N., Massey R. J., Taylor J., Refregier A., 2007, ApJ, 669, 714

Molendi S., Pizzolato F., 2001, ApJ, 560, 194

Molnar S., 2015, Frontiers in Astronomy and Space Sciences, 2, 7

Morrison G. E., Owen F. N., Ledlow M. J., Keel W. C., Hill J. M., Voges W., Herter T., 2003, ApJS, 146, 267

Murray S. G., Power C., Robotham A. S. G., 2013, Astronomy and Computing, 3, 23

Nagai D., Lau E. T., 2011, ApJ, 731, L10

Nagai D., Vikhlinin A., Kravtsov A. V., 2007, ApJ, 655, 98

Nastasi A., et al., 2011, A&A, 532, L6

Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493

Nelson K., Lau E. T., Nagai D., 2014, ApJ, 792, 25

Nelson D., et al., 2019, Computational Astrophysics and Cosmology, 6, 2

O'Hara T. B., Mohr J. J., Bialek J. J., Evrard A. E., 2006, ApJ, 639, 64

Okabe N., Zhang Y. Y., Finoguenov A., Takada M., Smith G. P., Umetsu K., Futamase T., 2010, ApJ, 721, 875

Pacaud F., et al., 2006, MNRAS, 372, 578

Pacaud F., et al., 2007, MNRAS, 382, 1289

Pacaud F., et al., 2016, A&A, 592, A2

Pesce J. E., Fabian A. C., Edge A. C., Johnstone R. M., 1990, MNRAS, 244, 58

Peterson J. R., Fabian A. C., 2006, Phys. Rep., 427, 1

Piccinotti G., Mushotzky R. F., Boldt E. A., Holt S. S., Marshall F. E., Serlemitsos P. J., Shafer R. A., 1982, ApJ, 253, 485

Pierre M., et al., 2004, J. Cosmology Astropart. Phys., 9, 011

Pierre M., et al., 2006, MNRAS, 372, 591

Pierre M., et al., 2016, A&A, 592, A1

Pillepich A., Porciani C., Reiprich T. H., 2012, MNRAS, 422, 44

Pizzolato F., Soker N., 2005, ApJ, 632, 821

Planck Collaboration et al., 2014, A&A, 571, A20

Planck Collaboration et al., 2016a, A&A, 594, A13

Planck Collaboration et al., 2016b, A&A, 594, A24

Planck Collaboration et al., 2016c, A&A, 594, A27

Planck Collaboration et al., 2018, arXiv e-prints, p. arXiv:1807.06209

Planelles S., Borgani S., Dolag K., Ettori S., Fabjan D., Murante G., Tornatore L., 2013, MNRAS, 431, 1487

Poole G. B., Fardal M. A., Babul A., McCarthy I. G., Quinn T., Wadsley J., 2006, MNRAS, 373, 881

Pratt G. W., Croston J. H., Arnaud M., Böhringer H., 2009, A&A, 498, 361

Rasia E., et al., 2006, MNRAS, 369, 2013

Rasia E., et al., 2012, New Journal of Physics, 14, 055018

Rasia E., et al., 2015, ApJ, 813, L17

Rasmussen J., Ponman T. J., 2007, MNRAS, 380, 1554

Rawle T. D., et al., 2012, ApJ, 747, 29

Reiprich T. H., Böhringer H., 2002, ApJ, 567, 716

Reiprich T. H., Basu K., Ettori S., Israel H., Lovisari L., Molendi S., Pointecouteau E., Roncarelli M., 2013, Space Sci. Rev., 177, 195

Rines K., Geller M. J., Kurtz M. J., Diaferio A., 2003, AJ, 126, 2152

Rines K., Geller M. J., Diaferio A., Kurtz M. J., 2013, ApJ, 767, 15

Rosati P., Della Ceca R., Norman C., Giacconi R., 1998, ApJ, 492, L21

Rossetti M., Eckert D., De Grandi S., Gastaldello F., Ghizzardi S., Roediger E., Molendi S., 2013, A&A, 556, A44

Rothschild R., et al., 1979, Space Science Instrumentation, 4, 269

Rozo E., et al., 2010, ApJ, 708, 645

Rozo E., Rykoff E. S., Bartlett J. G., Evrard A., 2014, MNRAS, 438, 49

Rozo E., Rykoff E. S., Becker M., Reddick R. M., Wechsler R. H., 2015, MNRAS, 453, 38

Ruderman J. T., Ebeling H., 2005, ApJ, 623, L81

Russell H. R., et al., 2012, MNRAS, 423, 236

Sanderson A. J. R., Ponman T. J., O'Sullivan E., 2006, MNRAS, 372, 1496

Santos J. S., Rosati P., Tozzi P., Böhringer H., Ettori S., Bignamini A., 2008, A&A, 483, 35

Santos J. S., Tozzi P., Rosati P., Böhringer H., 2010, A&A, 521, A64

Santos J. S., et al., 2011, A&A, 531, L15

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

Scharf C. A., Ebeling H., Perlman E., Malkan M., Wegner G., 1997, ApJ, 477, 79

Schellenberger G., Reiprich T. H., Lovisari L., Nevalainen J., David L., 2015, A&A, 575, A30

Sehgal N., et al., 2013, ApJ, 767, 38

Sereno M., 2015, MNRAS, 450, 3665

Sereno M., Ettori S., Moscardini L., 2015, MNRAS, 450, 3649

Serlemitsos P. J., Smith B. W., Boldt E. A., Holt S. S., Swank J. H., 1977, ApJ, 211, L63

Serra A. L., Diaferio A., Murante G., Borgani S., 2011, MNRAS, 412, 800

Siegel S. R., et al., 2018, ApJ, 861, 71

Simionescu A., et al., 2011, Science, 331, 1576

Smith R. K., Brickhouse N. S., Liedahl D. A., Raymond J. C., 2001, ApJ, 556, L91

Smith G. P., et al., 2016, MNRAS, 456, L74

Smolčić V., et al., 2016, A&A, 592, A10, XXL Paper XI

Somboonpanyakul T., McDonald M., Lin H. W., Stalder B., Stark A., 2018, ApJ, 863, 122

Springel V., et al., 2005, Nature, 435, 629

Stanford S. A., Gonzalez A. H., Brodwin M., Gettings D. P., Eisenhardt P. R. M., Stern D., Wylezalek D., 2014, ApJS, 213, 25

Suto D., Kawahara H., Kitayama T., Sasaki S., Suto Y., Cen R., 2013, ApJ, 767, 79

Tasse C., Le Borgne D., Röttgering H., Best P. N., Pierre M., Rocca-Volmerange B., 2008, A&A, 490, 879

Taylor M. B., 2005, in Shopbell P., Britton M., Ebert R., eds, Astronomical Society of the Pacific Conference Series Vol. 347, Astronomical Data Analysis Software and Systems XIV. p. 29

Tundo E., Moretti A., Tozzi P., Teng L., Rosati P., Tagliaferri G., Campana S., 2012, A&A, 547, A57

Tyson J. A., Valdes F., Wenk R. A., 1990, ApJ, 349, L1

Urban O., Werner N., Allen S. W., Simionescu A., Mantz A., 2017, MNRAS, 470, 4583

Vazza F., Brunetti G., Kritsuk A., Wagner R., Gheller C., Norman M., 2009, A&A, 504, 33

Vazza F., Angelinelli M., Jones T. W., Eckert D., Brüggen M., Brunetti G., Gheller C., 2018, MNRAS, 481, L120

Vikhlinin A., 2006, ApJ, 640, 710

Vikhlinin A., McNamara B. R., Forman W., Jones C., Quintana H., Hornstrup A., 1998, ApJ, 502, 558

Vikhlinin A., Kravtsov A., Forman W., Jones C., Markevitch M., Murray S. S., Van Speybroeck L., 2006, ApJ, 640, 691

Vikhlinin A., Burenin R., Forman W. R., Jones C., Hornstrup A., Murray S. S., Quintana H., 2007, in Böhringer H., Pratt G. W., Finoguenov A., Schuecker P., eds, Heating versus Cooling in Galaxies and Clusters of Galaxies. p. 48 (arXiv:astro-ph/0611438), doi:10.1007/978-3-540-73484-0˙9

Vikhlinin A., et al., 2009, ApJ, 692, 1060

Voges W., et al., 1999, A&A, 349, 389

Voit G. M., 2005, Reviews of Modern Physics, 77, 207

Walker S., et al., 2019, Space Sci. Rev., 215, 7

Wang Q. H. S., Markevitch M., Giacintucci S., 2016, ApJ, 833, 99

Weißmann A., Böhringer H., Šuhada R., Ameglio S., 2013, A&A, 549, A19

Wen Z. L., Han J. L., Liu F. S., 2009, ApJS, 183, 197

White S. D. M., Navarro J. F., Evrard A. E., Frenk C. S., 1993, Nature, 366, 429

White D. A., Jones C., Forman W., 1997, MNRAS, 292, 419

Willingale R., Starling R. L. C., Beardmore A. P., Tanvir N. R., O'Brien P. T., 2013, MNRAS, 431, 394

Willis J. P., et al., 2013, MNRAS, 430, 134

Wilson S., et al., 2016, MNRAS, 463, 413

Wittman D., Tyson J. A., Margoniner V. E., Cohen J. G., Dell'Antonio I. P., 2001, ApJ, 557, L89

Wojtak R., et al., 2018, MNRAS, 481, 324

Zhang Y.-Y., et al., 2010, ApJ, 711, 1033

ZuHone J. A., Markevitch M., Johnson R. E., 2010, ApJ, 717, 908

Zwicky F., 1933, Helvetica Physica Acta, 6, 110

Zwicky F., 1937, ApJ, 86, 217

Zwicky F., Herzog E., Wild P., Karpowicz M., Kowal C. T., 1961, Catalogue of galaxies and of clusters of galaxies, Vol. I

van Weeren R. J., de Gasperin F., Akamatsu H., Brüggen M., Feretti L., Kang H., Stroe A., Zandanel F., 2019, Space Sci. Rev., 215, 16

von der Linden A., et al., 2014, MNRAS, 439, 2