

A Universe-Old Puzzle: A Statistical Approach to Confirming the Maturity of Galaxy
Cluster XLSSC122

by

Ben Fredebo Rasmussen

A Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of

BACHELOR OF SCIENCE (HONOURS)

in the Department of Physics and Astronomy

© Ben Fredebo Rasmussen, 2024
University of Victoria

All rights reserved. This thesis may not be reproduced in whole or in part, by
photocopying or other means, without the permission of the author.

A Universe-Old Puzzle: A Statistical Approach to Confirming the Maturity of Galaxy
Cluster XLSSC122

by

Ben Fredebo Rasmussen

Supervisory Committee

Dr. J. Willis, Supervisor
(Department of Physics and Astronomy)

Dr. G. Steeves, Departmental Member
(Department of Physics and Astronomy)

ABSTRACT

Galaxy clusters are the largest and most massive gravitationally bound structures in the universe. Made up of distinct constituent parts each with their own astrophysical processes associated with them, clusters offer a window into the large-scale structure formation, thermodynamics, and physics of the universe at its extremes.

They are dominated, by mass ($\approx 80\%$), by over densities in dark matter that become virialized, or stable under the forces of gravity. This dark matter causes accretion of the surrounding baryonic matter in the form of both hot gas and galaxies. The gas found in the intracluster medium (ICM) emits light according to its temperature, with energies corresponding to x-ray photons. Conversely, galaxy emission is predominantly in the form of starlight which lies roughly in the visible part of the electromagnetic spectrum.

Galaxies in a cluster are largely collisionless objects that take many many years to properly establish equilibrium with the underlying gravitational potential provided by the dark matter. The gas, on the other hand, mechanically interact with itself, which combined with its high radiation, causes loss of both energy and angular momentum. This causes a much more rapid infall of gas onto the regions of dark matter over densities. The gas can then be said to thermalize much quicker than the galaxies. In an evolved cluster, the gas provides a suitable mapping of where the radiation-less dark matter is found.

The evolution of these structures is studied in extent in modern astronomy throughout different epochs of formation and stability. It is then insightful to research the properties of a well evolved cluster. The galaxy cluster XLSSC 122 is an exceptional candidate to do so as it sits at a redshift of $z = 2$ corresponding to a look-back time of 10.4 billion years while still exhibiting the maturity of clusters in the much later universe. Observations of the cluster have been made in a number of regimes using a variety of instruments. The *Hubble Space Telescope* (HST) obtained data in both the *F105W* and *F140W* wavebands using the Wide Field Camera 3 (WFC3) which sit in the optical range for the rest frame of the cluster. In addition, X-ray observations have been made in the *XMM-Newton* Large-Scale Structure survey (XMM-LSS).

Using previously determined galactic mass data and the X-ray emission data as a proxy for the distribution of dark matter, the interplay between cluster galaxies and cluster gas is studied here, leveraging a variety of statistical techniques. In doing so, we find that there exists a bimodality within the cluster member population with respect to correlations with the underlying dark matter distribution or plasma flux. There exists a population of galaxies that are both young and blue in colour, that do not correlated with the X-Ray flux. We

conclude that these galaxies are in the process of falling into the cluster and as such have not had time to reach equilibrium with their surroundings. Furthermore, there is a significant population of older and redder galaxies that is highly correlated with the background flux due to the intra-cluster gas. These galaxies formed before the formation of the cluster, and have had sufficient time to fall into the dark matter potential which is demonstrated by this increase in correlation. These two insights are consistent with the hypothesis that the galaxy cluster XLSSC 122 is indeed a mature cluster of galaxies. This result is important as this is the highest redshift galaxy cluster found that exhibits this level of maturity and gravitational relaxation.

Contents

Supervisory Committee	ii
Abstract	iii
Table of Contents	v
List of Figures	vii
Acknowledgements	xii
Dedication	xiii
1 Introduction	1
1.1 What are Galaxy Clusters?	1
1.2 Constituent Parts of a Cluster	1
1.3 Our Cluster	3
1.3.1 The Data: Hubble	5
1.3.2 The Data: XMM-Newton	7
1.3.3 Meet the Members	8
1.4 Contribution	11
1.5 Agenda	11
2 Flux Measurements	13
2.1 Manual Flux	13
2.2 Flux Using Coordinate Reprojection	14
2.3 Comparison Testing	16
2.4 Sub-Samples of the Cluster	18
3 Linear Correlation Techniques	20
3.1 A Final Sub-Sample	21

3.2 Linear Correlation Results	25
4 Radial Correlation Techniques	29
5 Spatial Correlation Techniques	35
6 Analysis and Findings	44
Bibliography	49
A Entire Field of View Images	52

List of Figures

Figure 1.1 Greyscale figure is a composite Hubble image using the $F140W$ passband. The image is displayed using a FK5 projection. The two dashed circles are the measured r_{200} (outer) and r_{500} (inner) values corresponding to the radius at which $\rho_{cluster} = 200/500 \cdot \rho_{critical}$ where $\rho_{critical}$ is the critical density of the universe. Spectroscopically confirmed 'Gold' and 'Silver' cluster members, see (Willis et al., 2020), are outlined in red and green while a foreground structure at $z = 1.93$ is highlighted in blue. Adapted from (Willis et al., 2020).	6
Figure 1.2 <i>Left:</i> Raw XMM-Newton image of the cluster. Photon energy range is 0.4 – 3.0 keV using combined data from the PN and MOS ccds on the space telescope. Image uses same projection as Fig. 1.1 with circles at r_{200} (inner) and r_{500} (outer). Some artifacts from chip gaps in the ccd can be seen. <i>Right:</i> Adaptively smoothed image of XLSSC 122 with bright foreground sources removed. Both figures adapted from (Mantz et al., 2018).	8
Figure 1.3 Histogram of found redshift values for galaxies in the region of the cluster where $F140W_{Kron} < 24$. Colours follow same identification as Fig. 1.1 with unclassified galaxies shown in grey. Vertical dashed lines are the mean values for both the $z = 1.93$ and $z = 1.98$ structures. Adapted from (Willis et al., 2020).	9

Figure 1.4 Colour Magnitude Diagram for all galaxies in the HST image. Colours follow same identification as Fig. 1.1 with all other galaxies in the field of view denoted as grey. Error bars are shown as the 1σ uncertainty associated with the magnitude determination. Horizontal axis magnitudes found with an adaptive Kron aperture (denoted with subscript Kron) while the vertical magnitude difference was found with an aperture of fixed angular size of $0.8''$. Vertical dotted lines are spectroscopic limits of $F140W_{Kron} = 24, 24.5$ while horizontal lines are the limit for red-sequence consideration (dashed-dot) and a least-squares fit to the red members of the cluster (solid). Adapted from (Willis et al., 2020).	10
Figure 2.1 Visual demonstrating the method for which the manual flux behind each galaxy in the cluster was obtained. The specific example here is for object 451, a 'Gold' cluster member. Overlaid are the coarsely resolved X-ray image pixels with counted pixels highlighted in red.	14
Figure 2.2 Segmentation map of the image in Fig. 1.1. Each object corresponds to the boundaries of an object in the Hubble data with pixel intensity values equal to their catalog position.	15
Figure 2.3 Comparison of three methods of reprojection with the segmented galaxy from the Hubble image. <i>Top Left:</i> Object 526, a 'Gold' cluster member and also the brightest galaxy in the cluster (BCG). Shape is somewhat odd due to overlapping galaxies in the centre of the cluster. <i>Top Right:</i> Object 526 reprojected into the X-ray frame using the routine here (Robitaille et al., 2023a) with a Gaussian kernel. <i>Bottom Left:</i> With a Hanning kernel. <i>Bottom Right:</i> Using an exact flux-conserving spherical polygon intersection method found here (Robitaille et al., 2023a).	16
Figure 2.4 Scatter plots of the calculated flux in counts-per-second as a function of galactic mass taken from (Trudeau et al., 2022) for the 26 confirmed members. <i>Top Left:</i> Using the aforementioned Gaussian reprojection method. <i>Top Right:</i> Using the Hanning method. <i>Bottom Left:</i> Using the flux-conserving exact method. <i>Bottom Right:</i> By hand.	17

Figure 2.5 <i>Left:</i> Comparison test for the flux from the raw image vs. the adaptively smoothed image for the 26 confirmed member galaxies. Line overlaid is $F_{Raw} = F_{Smooth}$. <i>Right:</i> Comparison test between the two most trustworthy methods, manual flux determination and flux conserving exact reprojection. Line overlaid is $F_{Manual} = F_{Exact}$.	18
Figure 3.1 Pearson r coefficient for the 4 sub-samples as well as all 26 confirmed member galaxies in the cluster. Included are all five coefficient values for the four methods of flux determination: by hand, exact reprojection, Gaussian kernel reprojection, and Hanning kernel reprojection.	21
Figure 3.2 Scatter plot of the mass in the form $\log_{10} \frac{M}{M_*}$ from (Trudeau et al., 2022) versus $F140W_{Kron}$ magnitude of the 26 member galaxies. Line overlaid is the linear least-squares fit on the data to establish a relation between the two physical quantities. Mass of galaxies left in log form so as to not generate negative mass for faint galaxies.	22
Figure 3.3 Recreation of Fig. 1.4 that includes fainter and redder objects. Blue objects are the objects that lie within the spectroscopic completeness limit of $F140W_{Kron} < 24$ while red objects are objects too faint to get a well resolved spectra. Grey objects fall below a colour threshold of $[F105W - F140W]_{ap} = 1.15$.	23
Figure 3.4 Plot of the estimated redshift with uncertainties for the member extension candidates established in Fig. 3.3 along the horizontal axis. The vertical axis is the pearson-r coefficient found using a bootstrap method that selected 69.0% of the sample and correlated, repeating 10^6 iterations. Vertical line corresponds to the cluster redshift of $z = 1.98$. Horizontal line shows a selection threshold of the 69th percentile at $r = 0.329$.	25
Figure 3.5 Scatter plot of X-ray flux in counts-per-second as a function of the galactic mass from (Trudeau et al., 2022). All four sub-samples from prior are included as well as the entire sample and the red extended sample, shown in light red, that was generated previously.	26
Figure 3.6 Pearson r coefficient bar plot for all 6 of the finalized samples.	27
Figure 3.7 Pearson r coefficient <i>p-value</i> bar plot for all 6 of the finalized samples. Plotted logarithmically.	28

Figure 4.1 Normalized cumulative radial mass distributions for each of the samples summed to last mass in each sample. Included is the X-Ray flux profile in yellow, normalized to the radial coordinate corresponding to the last mass in the red extended sample. Some ambiguity present as to where the normalization should occur.	31
Figure 4.2 KS-Test D-statistic from the distributions in 4.1. All values fall below 0.5 with the blue and young samples showing considerably larger D-statistics.	32
Figure 4.3 KS-Test p-value from the distributions in 4.1 corresponding to the critical values of the D-statistics from 4.2.	33
Figure 5.1 Spatial mass densities generated using a circular aperture KDE with varying radii (in kpc). Sample shown here is the 26 member (all) sample of galaxies. The X-Ray flux distribution is included in the top left for comparison. The generated density map has been pixel matched to the coordinate frame of the X-Ray image.	36
Figure 5.2 2D Pearson r coefficient as a function of aperture size using a circular aperture with all 6 previously defined samples. Plot gives the scale with which the two distribution types are correlated. Red extended sample is most correlated followed by the red and old samples. Blue becomes moderately correlated at high apertures while the young galaxies are uncorrelated.	37
Figure 5.3 Gaussian and Circular kernel with a parameter $a = 50$ kpc. The area under each curve remains the same, but the mass distribution after the application of the kernel for a Gaussian kernel is considerably more spread out.	38
Figure 5.4 Spatial mass densities generated using a Gaussian aperture KDE with varying 'a' parameter (in kpc). Sample shown here is the 26 member (all) sample of galaxies. The X-Ray flux distribution is included in the top left for comparison. The generated density map has been pixel matched to the coordinate frame of the X-Ray image.	39

ACKNOWLEDGEMENTS

I would like to thank Dr. Jon P. Willis for taking me on to complete this project. It has been a pleasure to be supervised by you, to learn from you, and thank you for helping me navigate this period of transition I have found myself in. Thank you to both of my parents. Your unconditional support has allowed me to begin to pursue this journey into astronomy I did not think possible. Finally, thank you to my friends and peers in the department. I could not have gotten through this degree without each and every one of you.

DEDICATION

I dedicate this text to my two wonderful and loving grandmothers. To Leona, I think it is safe to say my passion for astronomy is not going to just be a hobby. To Sina, whenever I get the chance to look into the cosmos with a telescope I will look for you, smiling and loving me from the heavens.

Chapter 1

Introduction

1.1 What are Galaxy Clusters?

Galaxy clusters are the largest self-gravitating celestial structures in the observed universe. If sufficiently evolved in time, they exhibit a characteristic known as virialization, or stability under the forces of gravity (Willis et al., 2020). The formation of galaxy clusters occur over timescales that rival the age of the universe itself.

Anisotropies in the density of the early universe grow over this time, amplified by gravity, allowing for complex dynamical systems such as these to develop (Diaferio et al., 2008). These fluctuations in the primordial density field evolve through accretion of the surrounding matter and stabilize through slow-moving and dynamical gravitational processes (Kravtsov and Borgani, 2012). The matter in the universe is diverse and takes many forms and so in this process of accretion, galaxy clusters become made up of many of the elements that make up the universe as a whole.

The intricate coordination between gravity fields, the formation of galaxies and stars, and the presence of dark matter make galaxy clusters act as a cosmic laboratory. They are then ideal for studying many physical processes in the universe such as structure formation, astrophysical thermodynamics, and galaxy evolution and contribute to our understanding of cosmology (Kravtsov and Borgani, 2012).

1.2 Constituent Parts of a Cluster

The diversity of cluster constituent elements can be simplified to three major components. As with much of the rest of the universe, the matter content of a cluster is dominated by dark

matter (Diaferio et al., 2008). Dark matter has proved to be one of the largest unsolved mysteries in both cosmology and particle physics (Garrett and Duda, 2011). Roughly 80% of the gravitating matter in the universe is *dark* insofar as it does not exhibit any detectable electromagnetic radiation and does not interact with other matter in any way apart from gravitationally. We are then only able to find this exotic material by measuring the dynamics of conventional, light emitting, components of a system. The quoted percentage of the universal makeup of matter corresponds well with that of a cluster, with approximately 85% of cluster mass being made up of dark matter (Trudeau et al., 2022). Dark matter in a cluster acts as the foundation that other matters collapses onto and is the most significant component of a cluster.

The next noteworthy constituent of a cluster is an extremely hot gas in the form of plasma that fills the intracluster medium (ICM) smoothly and emits light in the X-Ray regime (Kravtsov and Borgani, 2012). This gas, surprisingly, is the dominant form of baryonic matter in a cluster (Mascolo et al., 2023). Baryonic matter can also be described as *visible* matter, and so the mass of plasma outweighs that of the stars within the galaxy cluster members by an order of magnitude. This gas emits light characteristic of particles of temperatures in the many millions of Kelvin. This high temperature is largely driven by an energetic feedback loop supplied by supernovae and active galactic nuclei (AGNs) (Kravtsov and Borgani, 2012).

Plasma particles at these temperatures in a cluster interact within themselves immensely. When combined with its violent radiation, ICM plasma can rapidly lose energy and angular momentum in the cluster. We may then describe the ICM gas as a *collisional* population of particles (Kravtsov and Borgani, 2012). This means that the gas particles within a cluster will thermalize, or achieve gravitational equilibrium, with their surroundings quickly with respect to the overall cluster evolution. The distribution of plasma within the ICM will then be intrinsically tied to the distribution of dark matter in the cluster. More formally, the iso-density surfaces of the ICM will trace the iso-potential surfaces due to dark matter in a virialized cluster (Kravtsov and Borgani, 2012).

The final major component of a galaxy cluster is of course the galaxies themselves. Much of the little remaining mass budget of a cluster is found within galaxies, which are made up of dust, cold gas, and stars (Diaferio et al., 2008). Stars are the primary emitter within the optical regime for a clusters rest frame. In contrast to the ICM gas, galaxies rarely, if ever, interact between one another. This leads us to describe them as *collisionless* objects, and although considerably larger than gas particles, act as point masses in the greater context of the cluster (Kravtsov and Borgani, 2012). This means that they are governed purely by

their random motion and momentums that take many eons to perturb into an equilibrium state.

To summarize, the important points for our analysis here are as follows. The underlying structure of a galaxy cluster is governed by non-emitting dark matter which is traced by fast thermalizing intracluster gas. This gas emits within the X-Ray regime and can be imaged. Juxtaposed with this gas are many optically emitting galaxies that take significant time to reach the same spatial distributions of the gas and the foundational dark matter. The maturity of a cluster can then be roughly described by the relationship between how well the orientations of the overall galaxy arrangements match those described by the emission of cluster X-Rays.

1.3 Our Cluster

As mentioned previously, the search for and study of clusters at different epochs in the timeline of the universe are important to our understanding of how they evolve. This supposition leads us to search for galaxy clusters at the extreme range of what is observationally possible. This leads us to the XXL Survey, a 50deg^2 survey of the sky in multiple X-Ray wavelengths that was designed to find many hundreds of galaxy clusters out to a redshift of unity and provide a well-defined sample of extragalactic objects (Pierre et al., 2016). The parameters of this cosmological experiment have in recent years been extended to higher redshift objects leading to the discovery of the galaxy cluster XLSSC 122. This cluster is unique in that it is the first massive cluster discovered in the XXL Survey that has X-Ray emission characteristic of a redshift of ≈ 2 (Mantz et al., 2018). When we talk of redshift in both a cosmological or observational capacity, we are talking about a change in the wavelength of a photon of light with respect to the two moments of emission and observation of that photon. We may write an explicit expression relating redshift to both observational and cosmological parameters following (Willis, 2022):

$$z = \frac{\lambda_0 - \lambda_e}{\lambda_e} = \frac{\lambda_0}{\lambda_e} - 1 = \frac{a(t_0)}{a(t_e)} - 1 \quad (1.1)$$

From an observational perspective, redshift is defined in terms of the change in the wavelength of light at the time of emission, λ_e , to what we actually observe, λ_0 . This is important as it allows us a 'line-of-sight' measurement for distance in terms of just the incoming light from an object. Tangibly, we may obtain experimental values of z by finding known spectral features in the electromagnetic spectrum of an object, like a galaxy, and

comparing to where we know those features should be in the rest reference frame.

This observational redshift also corresponds to a *cosmological* redshift given by the RHS of equation (1.1). This value describes the relative increase of the universe's scale factor, $a(t)$, and is due to the actual expansion of the fabric of the universe rather than relative velocities like we are usually familiar with when talking about conventional wavelength shifting (Willis, 2022).

With an understanding of what redshift means, our cluster XLSSC 122 sits at $z = 1.98$ as a mean redshift for confirmed cluster members (Willis et al., 2020). This is significant for a number of reasons. The first is that, at least at the time of Mantz et al. [2018], this cluster is the highest redshift cluster found. Secondly, redshifts of this magnitude describe parts of the universe that are at a considerably younger age than what it is now. Due to the expansion of the universe, we must describe age as a *lookback* time or the time in the past where the light we receive in the present day was emitted. This can be found using the following from (Ringermacher and Mead, 2014):

$$t_L = t_H \int_0^z \frac{dz'}{(1+z')E(z')} \quad (1.2)$$

Where t_L is the lookback time and $E(z)$ can be written as:

$$E(z) = \sqrt{\Omega_m(1+z)^3 + \Omega_k(1+z)^2 + \Omega_\Lambda} \quad (1.3)$$

With Ω_m as the dark matter plus baryonic matter density parameter, Ω_k as the curvature parameter, and Ω_Λ as the dark energy density parameter (Ringermacher and Mead, 2014). We may solve this explicit using (Wright, 2006) with the following cosmological parameters used by (Willis et al., 2020): $\Omega_m = 0.286$, $\Omega_k = 0$ for a flat universe, and $\Omega_\Lambda = 0.714$ with $H_0 = 69.6 \text{ km s}^{-1} \text{ Mpc}^{-1}$. For a redshift of $z = 1.98$ we arrive at a lookback time for XLSSC 122 of $t_L = 10.372$ Gyr or 10.372 billion years from emission until now. Using the same cosmological parameters, we also calculate the age of the universe to be $t_{age} = 13.720$ Gyr. This means that the light observed from this cluster was emitted when the universe was only 3.348 Gyr old.

Despite the extreme age difference between us and the observed state of the cluster, it is remarkably well evolved. As we will see later, XLSSC 122 has a well established red sequence of galaxies which indicate that the formation of this cluster began well before the observations (Willis et al., 2020). Further, a well defined region of intracluster medium gas has been found that represents a reasonably well developed virialised halo for the cluster further suggesting cluster stability (Mantz et al., 2018). It has also been found that the

metallicity of the cluster with respect to solar $\frac{Z}{Z_*} = 0.33^{+0.33}_{-0.17}$ is consistent with lower redshift clusters (Mantz et al., 2018).

We know from Section 1.2 that a cluster takes very long timescales to form properly. This means that our cluster took much of the early universe to form with additional time required for the galaxy formation themselves. A mean, dust-free, stellar age of 2.98 Gyr has been found to support this claim (Willis et al., 2020) further suggesting star formation occurred before the infall of galaxies into the cluster. This corresponds to a galactic star formation history that begins when the universe was only 370 million years old. The combination of cluster parameters consistent with a virialized, mature galaxy cluster with the young age of the universe at the time of emission and very young age at the time of star formation make XLSSC 122 a unique candidate of study.

1.3.1 The Data: Hubble

XLSSC 122 has been imaged in multiple wavelength bands from different telescopes. To gain insight into the properties of the galaxies in the cluster, photometry from The Hubble Space Telescope was taken in two bands: F105W and F140W (Willis et al., 2020). The numbers correspond to the central wavelengths of 10500Å and 14000Å respectively and both filters are denoted W for wide. The observations were done on the Hubble wide field camera 3 (WFC3) and were reduced with the procedure outlined in (Willis et al., 2020).

The image in the F140W band can be seen in Fig. 1.1 with each member or nearby galaxy circled in different colours. (Willis et al., 2020) outlines a method for classifying the cluster members into their respective qualities. Integrating the redshift probability distribution within the region of $z \in [1.96, 2.00]$ gives rise to a value P_{mem} which classifies membership. If $P_{mem} > 0.5$ we may call it a *gold* member with 33 galaxies fitting this criterion labelled in red. If we increase the qualification criteria to $0.1P_{mem} < 0.5$ we obtain 4 more *silver* members highlighted in green in Fig. 1.1. A further subset of galaxies is circled in blue where their respective redshift integrals give $P'_{mem} > 0.5$ where P'_{mem} is defined as the same PDF integral on redshift in the range $z \in [1.91, 1.95]$ giving a subset of galaxies that lie in the foreground independent of our cluster.

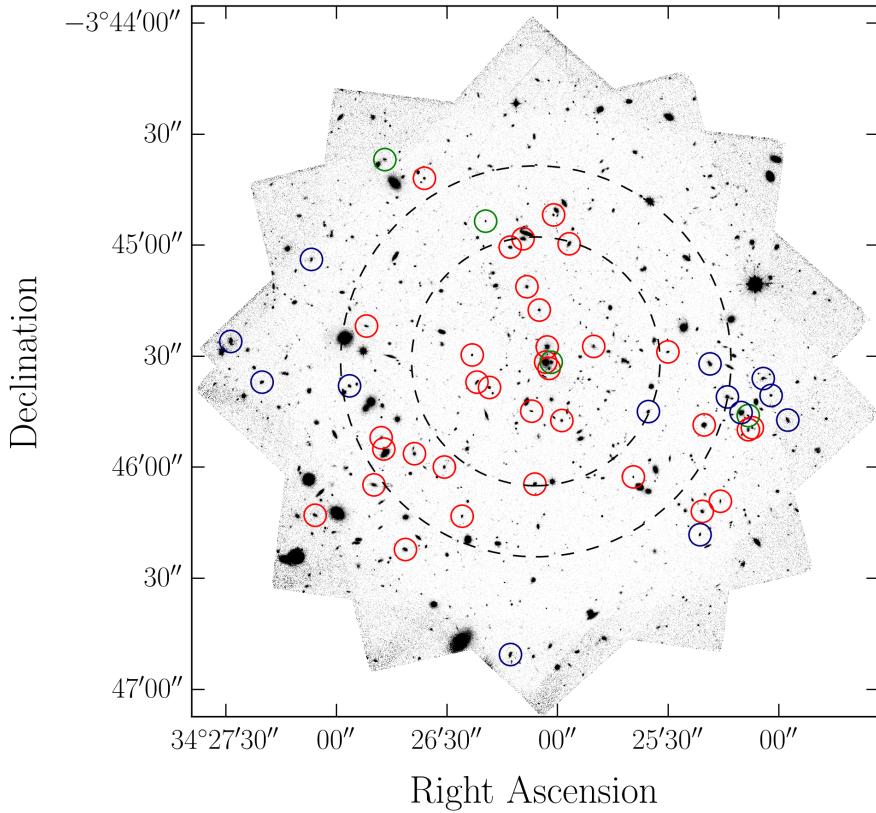


Figure 1.1: Greyscale figure is a composite Hubble image using the $F140W$ passband. The image is displayed using a FK5 projection. The two dashed circles are the measured r_{200} (outer) and r_{500} (inner) values corresponding to the radius at which $\rho_{cluster} = 200/500 \cdot \rho_{critical}$ where $\rho_{critical}$ is the critical density of the universe. Spectroscopically confirmed 'Gold' and 'Silver' cluster members, see (Willis et al., 2020), are outlined in red and green while a foreground structure at $z = 1.93$ is highlighted in blue. Adapted from (Willis et al., 2020).

For a spatial reference, the standard radii corresponding to 200 and 500 times the critical density of the universe have been included and will be used as a measuring stick for scales between different data. The values for r_{500} [r_{200}] can be found using the relation defined in (Mantz et al., 2018):

$$r_{500} = \left(\frac{3M(r_{500})}{4\pi(500 \cdot \rho_c(t))} \right)^{\frac{1}{3}} \quad (1.4)$$

But $\rho_c(z)$, the critical density will be dependent on the size of the universe at a certain time or redshift and can be written as (Willis, 2022):

$$\rho_c(t) = \frac{3}{8\pi G} \left(\frac{\dot{a}}{a} \right)^2 = \frac{3H^2(t)}{8\pi G} \quad (1.5)$$

Where G is the gravitational constant, $a(t)$ is the scale factor as before, and $H(t)$ is the time dependent Hubble parameter. We may use the following relation (Willis, 2022) (Ringermacher and Mead, 2014):

$$H(t) = \frac{\dot{a}}{a} = E(z) = \sqrt{\Omega_m(1+z)^3 + \Omega_k(1+z)^2 + \Omega_\Lambda} \quad (1.6)$$

Plugging in the same parameters as before we arrive at a critical density consistent with $\rho_c(z=1.98) = 6.4 \cdot 10^{-26} \frac{kg}{m^3}$. The value of mass for this radius was found from the X-Ray data as $M(r_{500}) = 6 \cdot 10^{13} M_*$ (Mantz et al., 2018). The radius value is then easily calculated and results in $r_{500} = 311 kpc$. Our estimate here falls well within the uncertainty of the value found by (Mantz et al., 2018) of $r_{500} = 295 \pm 23 kpc$. Similarly we find that $r_{200} = 422 kpc$ which is in agreement with (Mantz et al., 2018) of $r_{200} = 443 \pm 35 kpc$. To give an idea of scale, r_{500} largely encloses the member galaxies as seen in Fig. 1.1. In a more human readable unit, $r_{500} = 1.3 \cdot 10^{19} \text{ km}$ or about 500000 times the size of the solar system but XLSSC 122 only takes up about half of an arcminute (or $\approx 0.002^\circ$) in the sky from Earth.

1.3.2 The Data: XMM-Newton

To compliment the data already presented, and the reason detection of XLSSC 122 happened in the first place, data of X-Ray emission was taken during the XXL survey. X-Ray emission data largely maps the ICM gas as talked about before and is a combination of a number of instruments present on the XMM-Newton spacecraft (Mantz et al., 2018). Made up of data from the MOS (Metal Oxide Semi-conductor) CCDs and the PN CCD, the data seen in Fig. 1.2 corresponds to energies (wavelengths) between $0.4 - 3.0 keV$. Remnants of the gaps between CCD elements can be noticed alongside removed foreground objects. The same spatial radius of r_{500} and r_{200} can be seen superimposed to compare to Fig. 1.1.

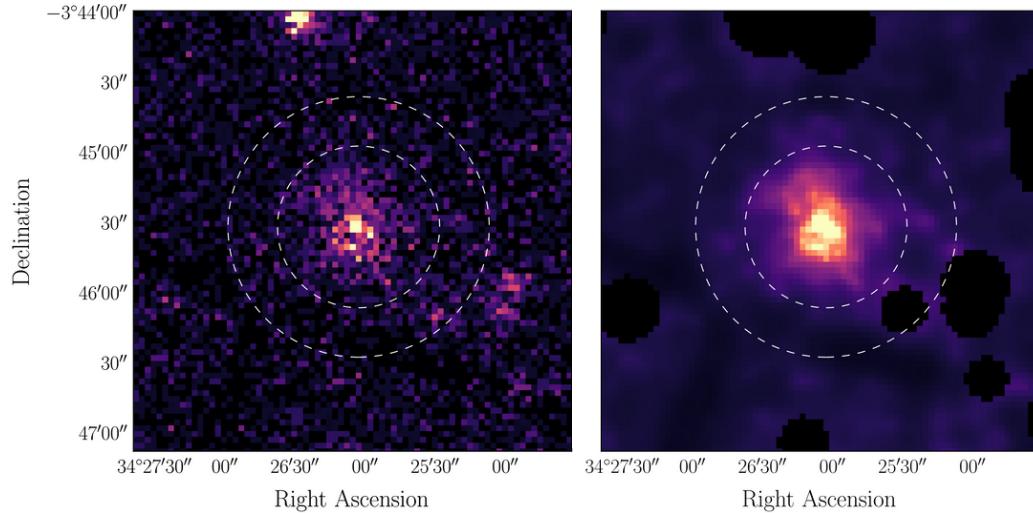


Figure 1.2: *Left:* Raw XMM-Newton image of the cluster. Photon energy range is 0.4–3.0 keV using combined data from the PN and MOS ccds on the space telescope. Image uses same projection as Fig. 1.1 with circles at r_{200} (inner) and r_{500} (outer). Some artifacts from chip gaps in the ccd can be seen. *Right:* Adaptively smoothed image of XLSSC 122 with bright foreground sources removed. Both figures adapted from (Mantz et al., 2018).

The second image in Fig. 1.2 corresponds to an adaptively smoothed version of the left image. The data found in this figure will represent the dark matter profile of the cluster system for the remainder of this analysis as described in section 1.2.

1.3.3 Meet the Members

From analysis done by (Willis et al., 2020), (Trudeau et al., 2022) we have considerable knowledge about the individual members of this cluster. The most basic properties of cluster members are their position relative to us. We may see angular displacements on the sky as in Fig. 1.1 but it is useful to probe this line-of-sight dimension using redshift. Objects in the vicinity of the cluster can be seen in Fig. 1.3 with the same colour classes as previously with the addition of gray, non-cluster galaxies. This figure clearly shows the existence of a foreground structure as highlighted by (Willis et al., 2020).

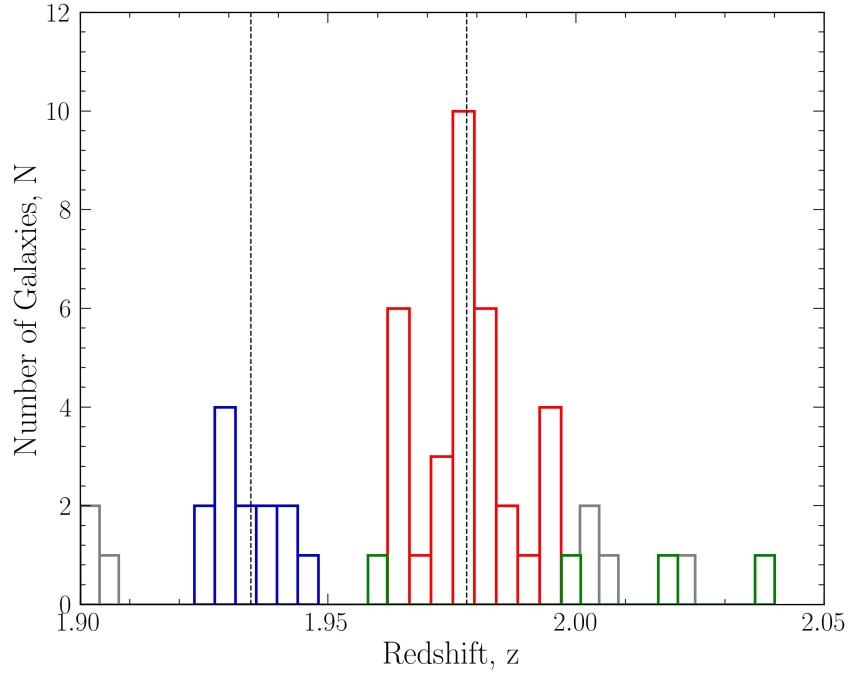


Figure 1.3: Histogram of found redshift values for galaxies in the region of the cluster where $F140W_{Kron} < 24$. Colours follow same identification as Fig. 1.1 with unclassified galaxies shown in grey. Vertical dashed lines are the mean values for both the $z = 1.93$ and $z = 1.98$ structures. Adapted from (Willis et al., 2020).

It is common to describe galaxies using what is known as a colour-magnitude diagram (CMD). Colour-Magnitude diagrams are ubiquitous in the field of astronomy, used as a reference diagnostic tool for systems as small as stars to individual galaxies all the way to groups of galaxies in clusters (Sciarratta et al., 2019). As in the name, they are plots of the *colour* of an object as a function of their magnitude (using the standard astronomy magnitude system). A colour in astronomy describes the difference between the magnitudes in two observed bands for an object. For example, the colour of the galaxies in our cluster can be described using the following (Willis et al., 2020):

$$\text{Colour} = F105W_{ap} - F140W_{ap} \quad (1.7)$$

where these are the observed bands from earlier. The subscript *ap* denotes a fixed aperture, in this case of angular size $0.8''$, that was used to integrate the flux. A fixed aperture is suitable for colour determination as it should, with correct size, get a consistent flux from the object in both bands regardless of the size of the object. This is in contrast to the

magnitude part of a CMD that is typically integrated in a variable aperture. In our case, a Kron aperture is used to compute magnitude which adapts to the size and shape of the object so as to integrate the entire galaxy flux in the band.

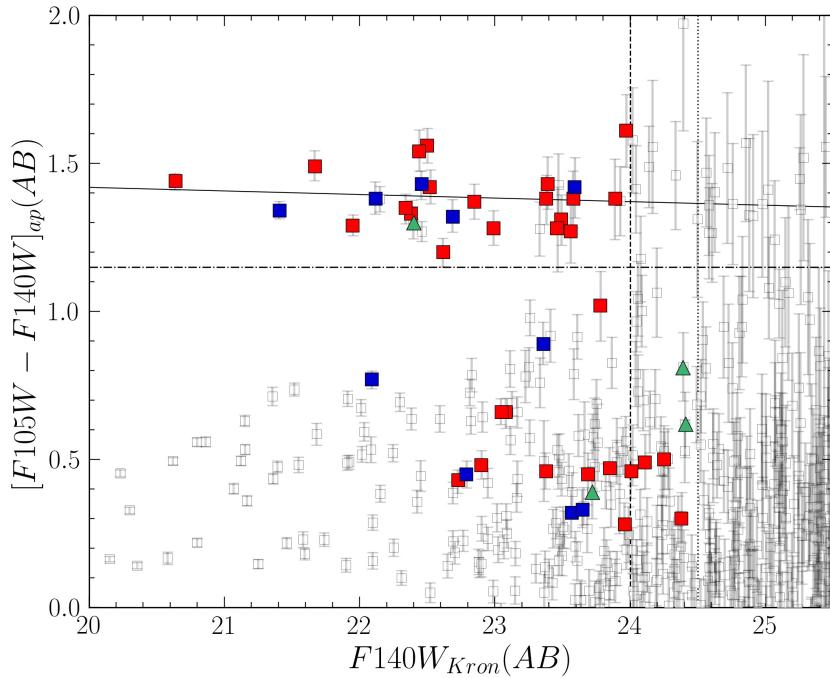


Figure 1.4: Colour Magnitude Diagram for all galaxies in the HST image. Colours follow same identification as Fig. 1.1 with all other galaxies in the field of view denoted as grey. Error bars are shown as the 1σ uncertainty associated with the magnitude determination. Horizontal axis magnitudes found with an adaptive Kron aperture (denoted with subscript Kron) while the vertical magnitude difference was found with an aperture of fixed angular size of $0.8''$. Vertical dotted lines are spectroscopic limits of $F140W_{Kron} = 24, 24.5$ while horizontal lines are the limit for red-sequence consideration (dashed-dot) and a least-squares fit to the red members of the cluster (solid). Adapted from (Willis et al., 2020).

As seen in Fig. 1.4, the colour is plotted against $F140W_{Kron}$ for all objects within the image range found in Fig. 1.1. The same convention for colour labelling of the galaxies as before is present here. The red represent the best, confirmed members of the cluster and not the *red* cluster members. Green is once again the silver members and blue are the foreground galaxies. All objects outlined in gray are within the field of view but do not have redshifts within $z \in [1.90, 2.05]$. Uncertainty corresponds to 1σ on each magnitude. The spectroscopic completeness limits from the detector are outlined with dotted lines while

the red-sequence cutoff is the dashed-dotted line that will be discussed in more detail later. Finally, the confirmed members of the cluster that lie on the red sequence were fitted using a least-squares model to produce the solid black line (Willis et al., 2020).

1.4 Contribution

As mentioned previously, (Willis et al., 2020) note this cluster as being remarkably *mature* when compared to other clusters found in similar universal epochs. Analysis done by (Willis et al., 2020) conclude that much of the star formation in the galaxies of the cluster, and especially the red-sequence, occurred at a coordinated time *before* the formation of the structure. This suggests there has been significant time for these galaxies to become virialized with the surrounding environment. Conversely, a number of confirmed cluster members are still forming stars at the time of emission and have a much younger stellar age. These galaxies have unlikely had sufficient time to equilibrate with their surroundings.

Combining the gas emission in the form of the X-Ray data with the optical stellar emission data in novel ways allows for a method to test the above claim. If it is indeed true that the X-Ray emission acts as a mapping of gravitational over-densities, we may hypothesize that these populations of old and red galaxies will find themselves correlating strongly with this map. This is due to their considerable time in the potential when compared to the newer, bluer galaxies. We may also expect that these new galaxies and their properties will be far less tied to the overall cluster environment. Confirmations of the above claims using a statistical approach will both contribute to our understanding of this particular cluster, XLSSC 122, as well as the evolution of high redshifted galaxy clusters.

1.5 Agenda

In this thesis, we investigate the interplay between cluster gas and the distribution of galactic mass in the cluster XLSSC 122 through a number of statistical techniques. Our analysis is three-fold, beginning with the simplest forms of correlation determination building up to more complex methods.

To begin, the X-Ray flux in the region of each member galaxy was determined using both a 'by-hand' approach and a number of more sophisticated algorithms. The manual method involved pixel matching the under-resolved X-Ray image with those of the optical data and computing a weighted mean flux value. For the algorithmic approach, we applied a segmentation map to the optical image, selecting each object in turn and reprojected this

selected object onto the same coordinate frame as the X-Ray image. The flux of X-Rays was then integrated for the radially weighted reprojected galaxy using a variety of kernels. These approaches were then compared.

To analyze the different populations of galaxies present within the cluster, they were split into distinct sub-samples delineated by benchmark properties in both colour and stellar age. With X-Ray flux values corresponding to each galaxy in the cluster and the masses obtained from Trudeau et al. [2022], a simple linear correlation analysis was performed for all sub-samples.

To further extend our analysis to galaxies without established stellar mass values, a bootstrap algorithm was implemented that artificially selected galaxies corresponding to the known ratios present in each sub-sample maximizing correlation with the other red cluster members. We called this the extended red sub-sample. Masses for these extra galaxies were found by extrapolating a linear fit onto the mass-magnitude relation of confirmed members.

Simple linear correlation between these quantities leaves much to be desired. To account for some spatial distribution present within both the gas and galaxies, we assumed radially symmetry and computed cumulative density functions (CDFs) in mass and flux. Our discrete one-to-one technique was transformed to a one-dimensional continuous analysis. The CDFs of each sub-sample were compared to the that of the cluster gas X-Ray CDF using the Kolmogorov-Smirnov test.

The first two techniques were limited in their ability to account for the spatial anisotropies present in the cluster as our galaxy distribution is not perfectly radial in nature. Using a kernel density estimator approach, maps of the galactic mass projected onto the region in the sky were generated for different apertures sizes and kernels for all sub-samples. These were then compared to the X-Ray image pixel-by-pixel to establish 2-D correlations. Convolving the two types of maps led to a correlation density mapping in addition.

All three methods of analysis contribute different knowledge to the morphology and statistical quantification of maturity of cluster XLSSC 122. The final component of this thesis was interpreting the results found in each section.

Chapter 2

Flux Measurements

Our first statistical test to perform on the cluster consists of relatively simple linear correlations between the galactic masses found by (Trudeau et al., 2022) and the corresponding flux in the X-Rays underneath each member object. This is motivated by an assumption that if the mass density profile of the galaxies is truly virialized, we should find these galaxies where the dark matter and hence the X-Ray emission is also denser. In order to perform this test, we need the masses found by (Trudeau et al., 2022) and the fluxes corresponding to each galaxy. Finding these fluxes will be done in a number of ways.

2.1 Manual Flux

The simplest and arguably most reliable method for obtaining the flux of X-Rays associated with the spatial position of each galaxy is to pixel match the two data sets and add up the flux by hand. This serves as not only a sanity check for further methods, but also a reliable starting point. As the resolution of the data acquired by XMM-Newton and Hubble are mismatched by a considerable amount, we find flux values by averaging the poorly resolved X-Ray data over any pixel that is also lit up by parts of the underlying galaxy. This is of course limited in its accuracy but not in its universality.

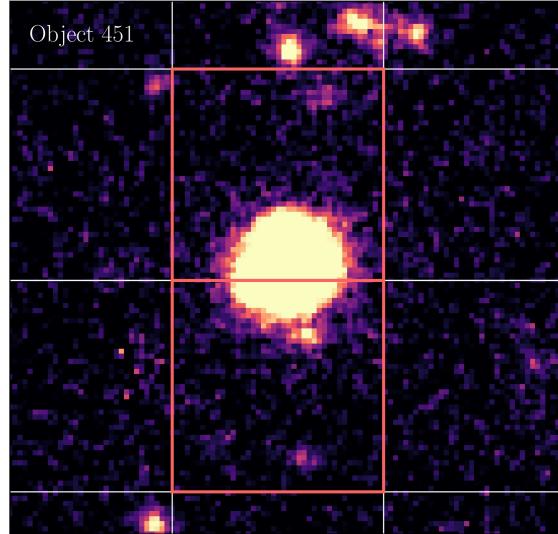


Figure 2.1: Visual demonstrating the method for which the manual flux behind each galaxy in the cluster was obtained. The specific example here is for object 451, a 'Gold' cluster member. Overlaid are the coarsely resolved X-ray image pixels with counted pixels highlighted in red.

To find the flux value, we take an unweighted mean value of all X-Ray pixels that overlap with the luminosity of the chosen galaxy. We may write this as:

$$F_{XRay} = \frac{\sum_{i=1}^{N_{pix}} F_i}{N_{pix}} \quad (2.1)$$

To demonstrate this, Fig. 2.1 shows the highlighted XMM-Newton data that would correspond to the pixels of flux for object 451, a gold cluster member. This method gives a solid estimate for the true flux values but is unweighted with respect to the radial distance from the object of interest and as such will give the same importance to each pixel involved. We then look for a method that works around this.

2.2 Flux Using Coordinate Reprojection

To circumvent the problems of the above method, we look to reprojection algorithms. Robitaille et al. [2023b] gives a number of ways to convert between astronomical coordinate

systems. In conjunction with these, a simple algorithm to find the individual flux values underneath each object is used here.

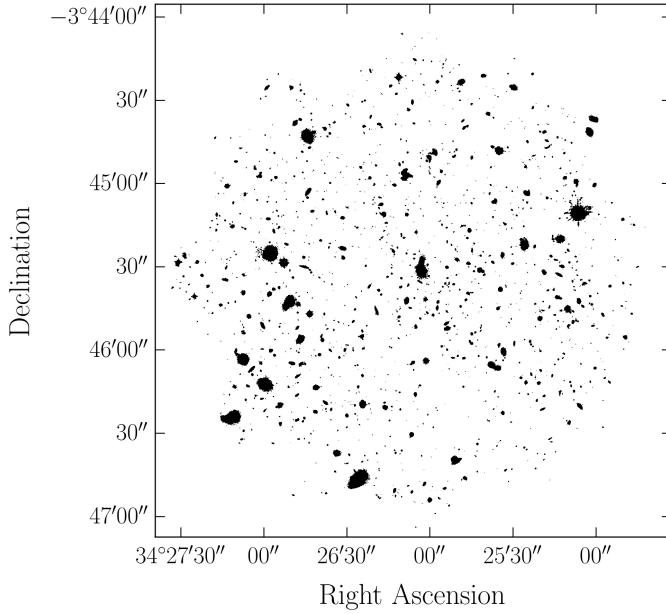


Figure 2.2: Segmentation map of the image in Fig. 1.1. Each object corresponds to the boundaries of an object in the Hubble data with pixel intensity values equal to their catalog position.

To implement our algorithm, we follow the proceeding steps:

1. Select object of interest
2. Create mask of this object using the segmentation map in Fig. 2.2
3. Reproject the mask from the Hubble coordinate frame to the X-Ray coordinate frame using one of the methods shown in Fig. 2.3 and create binary mask
4. Apply binary mask to the X-Ray data, retaining the shape of the underlying galaxy
5. Sum up all pixels in the new image to obtain the radially weighted flux value behind an object

This algorithm was implemented for a number of different reprojection algorithms defined in both (Robitaille et al., 2023a) and (Robitaille et al., 2023b). These methods included a

Gaussian kernel reprojection, a Hanning kernel reprojection, and a flux-conserving spherical polygon intersection method. We can see the effects of the methods on the brightest cluster galaxy (BCG) in figure 2.3 with the matched pixels of both data sets.

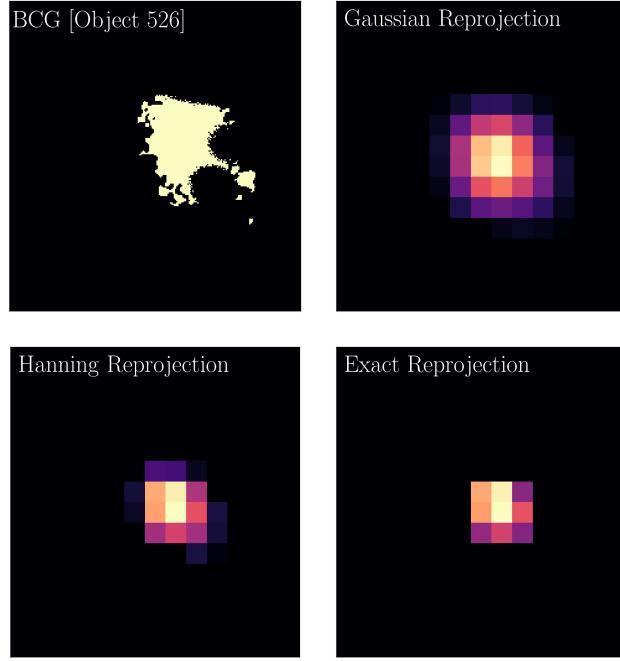


Figure 2.3: Comparison of three methods of reprojection with the segmented galaxy from the Hubble image. *Top Left:* Object 526, a 'Gold' cluster member and also the brightest galaxy in the cluster (BCG). Shape is somewhat odd due to overlapping galaxies in the centre of the cluster. *Top Right:* Object 526 reprojected into the X-ray frame using the routine here (Robitaille et al., 2023a) with a Gaussian kernel. *Bottom Left:* With a Hanning kernel. *Bottom Right:* Using an exact flux-conserving spherical polygon intersection method found here (Robitaille et al., 2023a).

Clearly, each method has a considerably different way to account for the radial size of an object. In contrast, the manual method would weight each pixel in a roughly 4x4 grid evenly.

2.3 Comparison Testing

With multiple methods used to obtain the flux of each galaxy in X-Rays, it is useful to compare the results. The first comparison test we may perform is between the four methods used. If all methods were equal, we should expect that the flux vs. mass plot should be the same for each method. As can be seen in Fig. 2.4 the same rough features are present

across all methods, like the cloud of values near the origin with two bright X-Ray sources. The dispersion in the cloud is clearly different from each method, and the intermediate data points differ in their respective flux values.

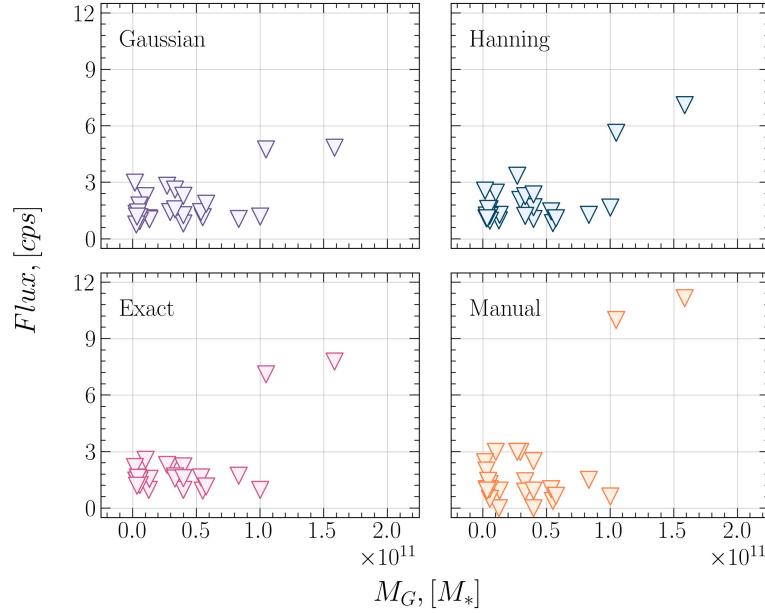


Figure 2.4: Scatter plots of the calculated flux in counts-per-second as a function of galactic mass taken from (Trudeau et al., 2022) for the 26 confirmed members. *Top Left:* Using the aforementioned Gaussian reprojection method. *Top Right:* Using the Hanning method. *Bottom Left:* Using the flux-conserving exact method. *Bottom Right:* By hand.

Regardless, it seems fairly reasonable to continue on with any of the methods as there is not a significant difference between the shapes of the plots. That being said, the benefits of a simpler Gaussian or Hanning attempt at reprojection are not apparent and for this technique it is likely that the exact reprojection is sufficient.

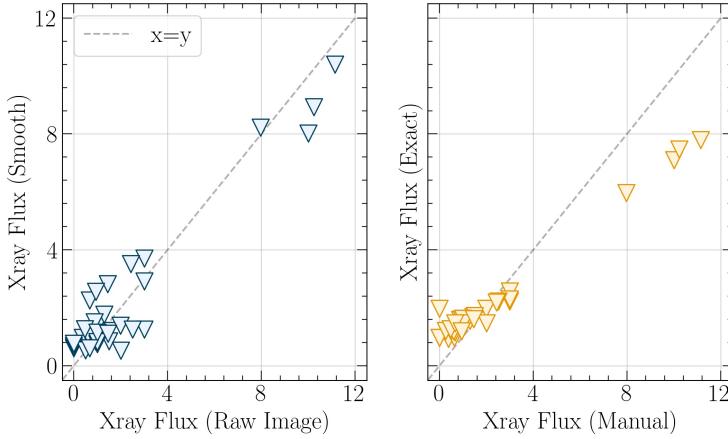


Figure 2.5: *Left:* Comparison test for the flux from the raw image vs. the adaptively smoothed image for the 26 confirmed member galaxies. Line overlaid is $F_{Raw} = F_{Smooth}$. *Right:* Comparison test between the two most trustworthy methods, manual flux determination and flux conserving exact reprojection. Line overlaid is $F_{Manual} = F_{Exact}$.

We may also compare the differences in flux obtained between the manual and exact methods. This can be seen in the right hand side of Fig. 2.5. It is clear that the manual method over estimates the flux for high mass galaxies while underestimating the low mass galaxies when compared to the exact flux method (or vice-versa). This can be explained as all pixels underneath an object were added up equally in the manual case to give an overestimate but also were diluted by pixels that only touched the object on the outskirts for the underestimate. It is likely that the exact reprojection is the best estimate.

Of note here is that there is two different sets of data for the X-Ray data. (Mantz et al., 2018) was able to adaptively smooth the original data set. This can be seen in Fig. 1.2 and flux values using the same methods outlined above were used on the smooth data set. The left hand side of Fig. 2.5 shows the manual flux between the two images. There is some dispersion but the features largely stay the same between the two images.

2.4 Sub-Samples of the Cluster

To gain insight into the features of our cluster, we must delineate between different populations of galaxy that we find within the cluster to learn about the cluster as a whole. As seen in Fig. 1.4, there is a clear bimodality in the colour of the galaxies in the sample.

This bimodality persists, as expected, into the age of the galaxies in the cluster. We then generate sub-samples of the cluster based on these two criterion. We may write these

as follows, with \tilde{t}_{age} as the median stellar age of each galaxy from (Trudeau et al., 2022):

1. Red if $[F105W - F140W]_{ap}(AB) > 1.15$
2. Blue if $[F105W - F140W]_{ap}(AB) \leq 1.15$
3. Old if $\tilde{t}_{age} > 1.0$ Gyr
4. Young if $\tilde{t}_{age} \leq 1.0$ Gyr

Creating modular sub-samples of the galaxies will allow us to draw conclusions about the clusters bimodality against itself. We will generate a fifth sub-sample later that takes into account the galaxies that do not have previously determined mass values from (Trudeau et al., 2022).

Chapter 3

Linear Correlation Techniques

Now that we have discrete sets of data that describe both the mass and integrated X-Ray flux underneath each confirmed object in the cluster, it is possible to quantitatively assess the correlation between each variable. To do this, we utilize a simple Pearson-r correlation as outlined here (Rummel, 1976) with the strength of correlation measured using the Pearson-r coefficient:

$$r_{xy} = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}} \quad (3.1)$$

Where the strength of the linear correlation between two random variables is measured between -1 and 1 in the dimensionless quantity r_{xy} and the bar denotes the mean value of the particular variable.

We then wish to use the above on our sets of discrete mass-flux data sets. Since we have some variation in the flux obtained from the four methods used, the Pearson-r is calculated and shown in Fig. 3.1.

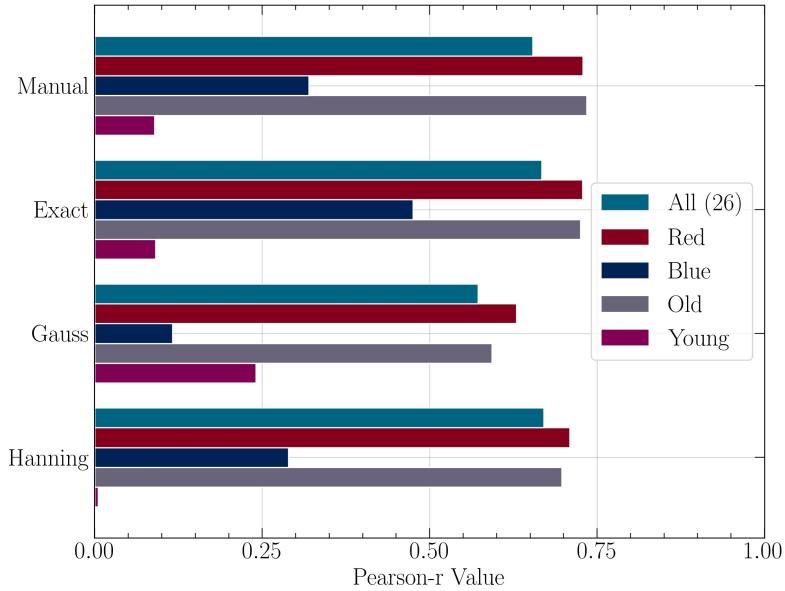


Figure 3.1: Pearson r coefficient for the 4 sub-samples as well as all 26 confirmed member galaxies in the cluster. Included are all five coefficient values for the four methods of flux determination: by hand, exact reprojection, Gaussian kernel reprojection, and Hanning kernel reprojection.

We see here that the behaviour of correlation between each of the sub-samples is largely independent of the method used. The strongest correlation is found in the red and old samples with magnitude only falling for the case of a Gaussian aperture reprojection. The young and blue sub-samples are much more weakly correlated in all cases. Once again, it seems like the choice of method does not effect the conclusions that we can draw from the data and as such we move forward with only the manual flux determination method.

Fig. 3.1 demonstrates that there exists intrinsic correlation between mass and X-Ray flux for red (old) galaxies in the cluster. This can also be shown to be a statistically significant correlation as will be seen later, and so we may leverage the assumptions of correlation of red galaxies to develop a final grouping of galaxies that we do not have mass data for.

3.1 A Final Sub-Sample

We are interested in the way that spatially distributed galaxies in our cluster match up with the X-Ray emission profile in the same region. The majority of the galaxies that have been analyzed by (Willis et al., 2020) (Trudeau et al., 2022) are the brightest and most central

members of the cluster, and as such we have mass data for the largest mass galaxies in the cluster. To get a better idea of how the fainter and spread out galaxies contribute to the overall relationship between gas and galactic mass, we look to find a sixth sub-sample that encapsulates these additional galaxies.

The first step in doing so is to estimate appropriate mass values for the fainter galaxies in the cluster. To do this, we establish a linear relation between mass and magnitude for the brightest member galaxies and extrapolate additional data points from this. As can be seen in Fig. 3.2, a fairly strong linearity can be established. Using this, mass values for all objects in Fig. 1.1 were obtained.

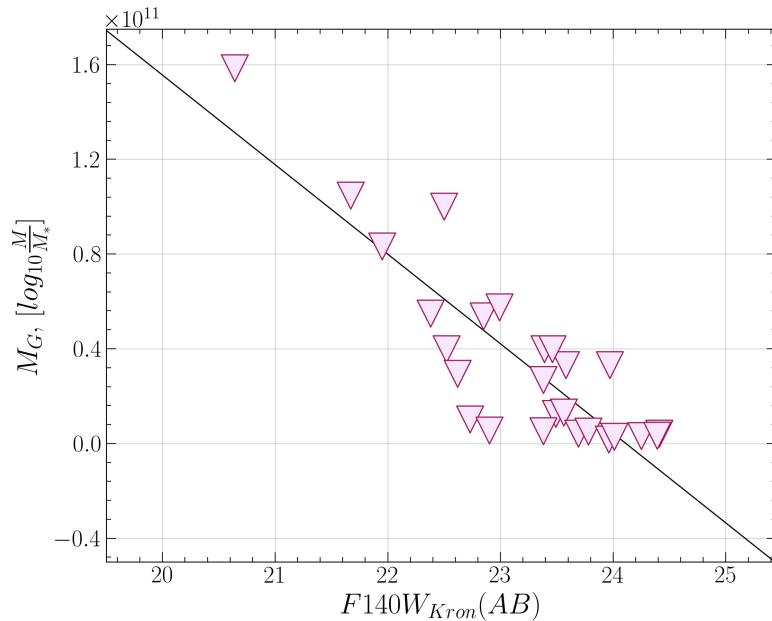


Figure 3.2: Scatter plot of the mass in the form $\log_{10} \frac{M}{M_*}$ from (Trudeau et al., 2022) versus $F140W_{Kron}$ magnitude of the 26 member galaxies. Line overlaid is the linear least-squares fit on the data to establish a relation between the two physical quantities. Mass of galaxies left in log form so as to not generate negative mass for faint galaxies.

Now to extend our sample into galaxies that fall below that spectroscopic completeness limits labelled in Fig. 1.4, we once again return to a colour-magnitude diagram. Since we are looking for galaxies that are faint, but are likely members of the cluster red-sequence, we set a colour threshold of $[F105W - F140W]_{ap} = 1.15$. We are then left with 89 additional candidates that could be cluster members with red colour.

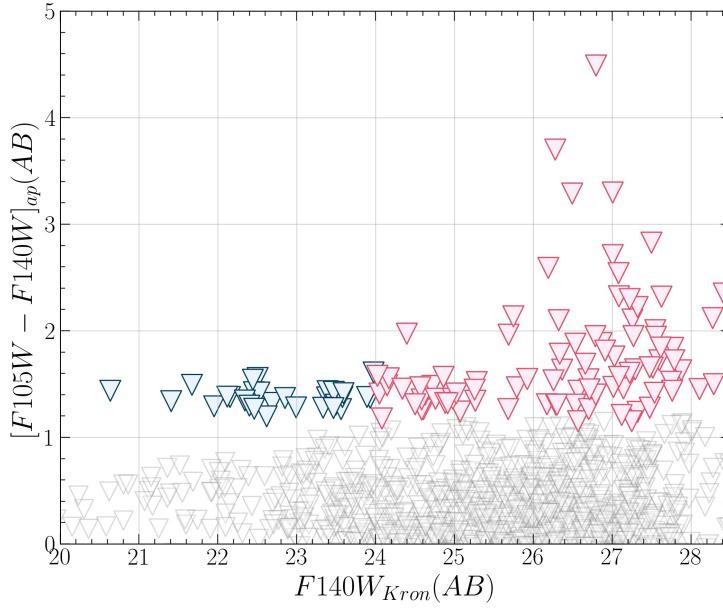


Figure 3.3: Recreation of Fig. 1.4 that includes fainter and redder objects. Blue objects are the objects that lie within the spectroscopic completeness limit of $F140W_{Kron} < 24$ while red objects are objects too faint to get a well resolved spectra. Grey objects fall below a colour threshold of $[F105W - F140W]_{ap} = 1.15$.

Fig. 3.3 shows the methodology here, as the objects that have been previously categorized are coloured blue with additional candidates highlighted in red. All objects that are too blue are coloured gray. Now that we have 89 galaxies that could be part of the cluster, we look to the original sample of red sequence galaxies found in Fig. 1.4. The total number of different types of objects on the red-sequence with $F140W_{Kron} < 24$ is 29. The ratios of gold members, silver members, foreground objects, and non-members to all objects on the red-sequence are as follows:

$$1. \frac{N_{Gold}}{N_{RS}} = 65.5\%$$

$$2. \frac{N_{Silver}}{N_{RS}} = 3.5\%$$

$$3. \frac{N_{FG}}{N_{RS}} = 17.2\%$$

$$4. \frac{N_{NM}}{N_{RS}} = 13.8\%$$

We can use these numbers to estimate the fraction of the red objects in Fig. 3.3 that will be a part of the cluster. To do this, we make a significant assumption that the ratio of

the types of objects in the cluster hold across all magnitudes. This is likely not the case, but we know that the cluster will have fainter members and so will the foreground structure, and against the background it is not unreasonable to assume that these ratios would roughly hold.

We also assume that gold and silver members contribute equally to this approach and so our target ratio of members in the candidates becomes $\frac{N_{add}}{N} = 69.0\%$. To select our final sample, we have implemented a bootstrap resampling approach that hinges on the assumption raised previously, that red members masses are strongly correlated with their respective X-Ray fluxes. The procedure is as follows:

1. Take the candidate sample of galaxies and select 69% of them at random
2. Compute the Pearson-r Correlation Coefficient for the sub-set of additional galaxies
3. Record both the r and p value for the iteration for all of the sub-set and add to a vector specific to each object
4. Repeat n times, adding values to each objects individual vector if they are included in the sub-set
5. Compute the mean Pearson-r of each object over the 69% of runs that the particular object was included for
6. Select the final 69% of objects with the largest correlation coefficients. These will become the extended red sample

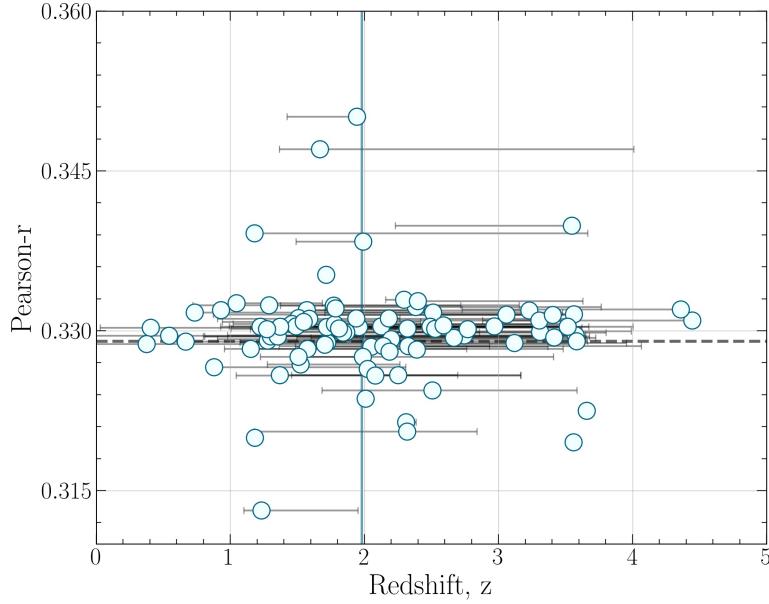


Figure 3.4: Plot of the estimated redshift with uncertainties for the member extension candidates established in Fig. 3.3 along the horizontal axis. The vertical axis is the pearson-r coefficient found using a bootstrap method that selected 69.0% of the sample and correlated, repeating 10^6 iterations. Vertical line corresponds to the cluster redshift of $z = 1.98$. Horizontal line shows a selection threshold of the 69th percentile at $r = 0.329$.

We can see the result of this method in Fig. 3.4. The effectiveness of this approach is highly dependent on the assumption that red galaxies in the cluster are the most correlated, and will be effected by the linear magnitude-mass relation established in Fig. 3.2. The procedure was ran a million times before converged correlation values were achieved. The cutoff coefficient can be seen in Fig. 3.4. This sub-sample of galaxies will be called the *red extended* sample from this point onwards.

3.2 Linear Correlation Results

Now that the final six sub-samples have been obtained, we may compute the linear correlation coefficient for each one and compare. Using equation (3.1), the Pearson-r was found for each sub-sample between the variables of galaxy mass and X-Ray flux. Of note here is that flux method for the final result is the manual one, as we can trust its effectiveness and the results largely do not change from method-to-method. Fig. 3.5 shows a scatter plot of these variables for each sample. Interesting characteristics here are the inclusion of the BCG in a

sample which may become important later on. Blue and young sub-samples are low mass and for the most part, low flux. There is considerable spread present in the red extended sample that likely comes from the mass determination extrapolation.

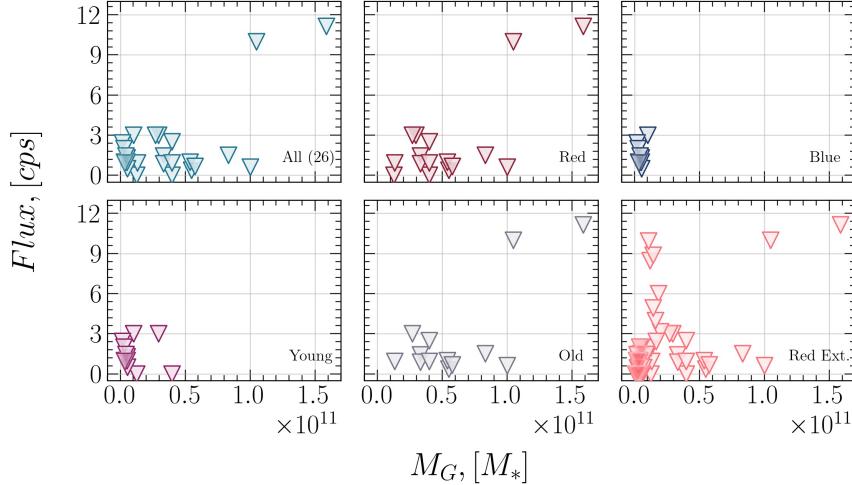


Figure 3.5: Scatter plot of X-ray flux in counts-per-second as a function of the galactic mass from (Trudeau et al., 2022). All four sub-samples from prior are included as well as the entire sample and the red extended sample, shown in light red, that was generated previously.

The calculated correlation coefficients can be found in Fig. 3.6 with the Red and Old samples showing strong linear correlations > 0.75 . The entire sample (All) includes these galaxies and as such also exhibits fairly strong linear correlation. The Blue sample shows moderate correlation while the young sample shows barely any. The extended Red sits almost at 0.5, what could be considered a moderate correlation as well.

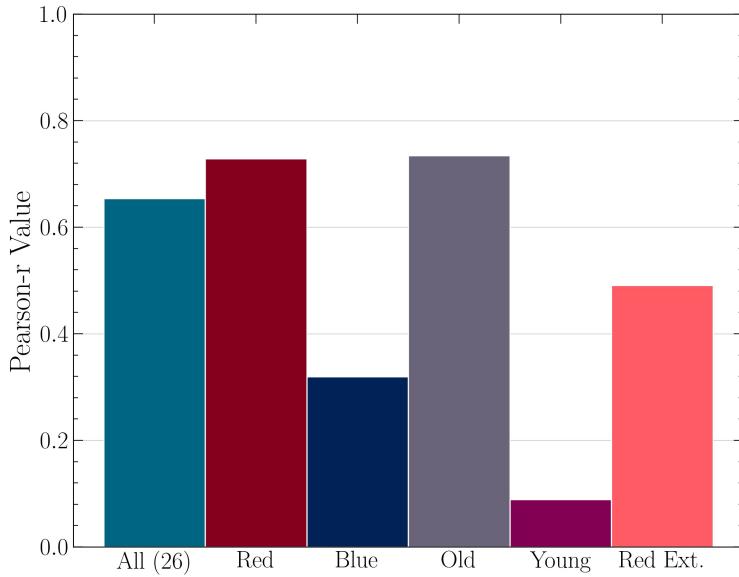


Figure 3.6: Pearson r coefficient bar plot for all 6 of the finalized samples.

To test the significance of these results, we look to find a p-value for the given statistics. The method to find the p-value here is the same as (Virtanen et al., 2020) implemented in scipy, where the distribution of the correlation coefficient is:

$$f(r) = \frac{(1 - r^2)^{n/2-2}}{B(\frac{1}{2}, \frac{n}{2} - 1)} \quad (3.2)$$

Where n is the length of our variable vector and B is the beta-function defined for real numbers as:

$$B(x, y) = \int_0^1 t^{x-1} (1-t)^{y-1} dt \quad (3.3)$$

For a given sample of random variables like we have here, the p-value is computed as the probability that $|r|$ of the sample drawn from an overall population with no correlation at all would be greater or equal to $|r|$ (Virtanen et al., 2020). Generally speaking, we may reject the null hypothesis, which is that the correlation does not differ significantly from zero, if our p-value is < 0.05 .

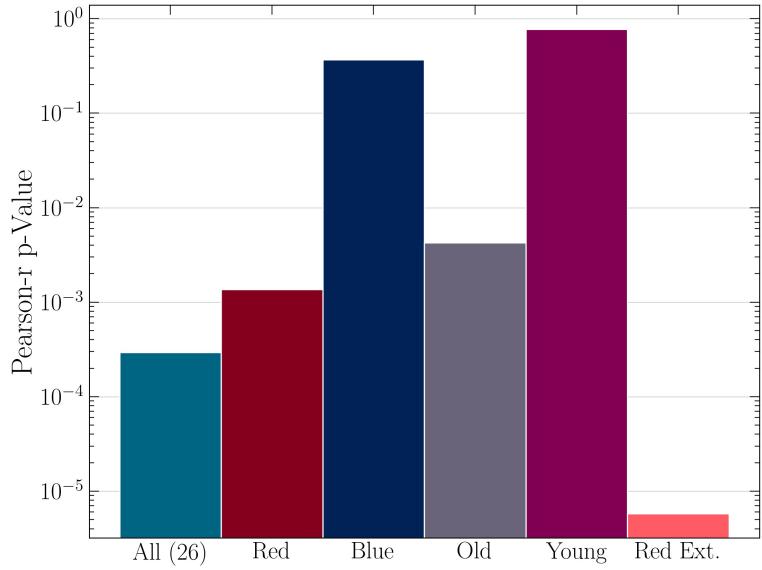


Figure 3.7: Pearson r coefficient *p-value* bar plot for all 6 of the finalized samples. Plotted logarithmically.

As we can see in Fig. 3.7, the significance of the results discussed above differ wildly. Any correlation present in the Blue and Young samples is largely irrelevant as both p-values are close to 1, meaning we cannot reject the null hypothesis. Old and red galaxies in the cluster can be said to have a statistically significant linear correlation between mass and X-Ray flux since both $p < 0.005$. Similarly, the entire sample is also significant as $p < 0.0005$. Of particular note here is the magnitude of the p-value for the extended red sample. Although the strength of the correlation of this sub-sample was not the highest, it is clearly the most significant with $p < 0.00001$. This is likely a symptom of the method of selection of the galaxies in the sample as their correlation within the cluster was part of the selection criteria. That being said, it is a good sign that this sample exhibits stronger properties than the other counterparts.

Linear correlation between galactic mass and X-Ray flux demonstrated a strong bimodality in the population of galaxies present in the cluster which is a step in the right direction in analyzing and comparing the cluster gas and galaxies. However, the limitations of this method are profound, as we do not take any spatial anisotropy into account. This is to say, all galaxies, regardless of their positions in the sky have been treated in the same way. We look to remedy this next.

Chapter 4

Radial Correlation Techniques

We now look to take into account the spatial distribution of galaxies and their X-Ray counterparts. To do this, we expand our analysis from the simple 1-Dimensional discrete linear correlation, to a 2-Dimensional radial profile analysis. In an idealized cluster, the radial asymmetry would be low as the constituents would relax, forming a radially uniform distribution of mass and flux. This of course will not be the case, but we may assume that there exists a radial symmetry to generate radial profiles of the cluster.

Using the same cosmology as before (Willis et al., 2020) as well as the calculator provided by (Wright, 2006), we can convert coordinates quoted in RA and DEC into units of kpc. The conversion factor is given as $8.530 \text{ kpc}/''$ at a redshift of $z = 1.98$. Subtracting the position of the BCG (assumed to be centre of the cluster) we arrive at a transformed set of coordinates with respect to a central origin. The angular distance of each galaxy from the centre is computed simply as:

$$\theta_{gal} = \sqrt{(\Delta RA_{gal})^2 + (\Delta DEC_{gal})^2} \quad (4.1)$$

Where $\Delta RA_{gal} = RA_{gal} - RA_{BCG}$, $\Delta DEC_{gal} = DEC_{gal} - DEC_{BCG}$. We then convert this angular distance to a radial one using the angular diameter distance found using (Wright, 2006):

$$R_{gal} = \frac{\theta_{gal}}{d_A(z = 1.98)} \quad (4.2)$$

In catalog space, we then have a radial coordinate and mass for each galaxy in each sub-sample. We can then compute the cumulative density function (CDF) as:

$$F_m(M_*) = P[M_* \leq m] \quad (4.3)$$

We can rewrite this in a computable form as:

$$F_m(R) = \frac{1}{M_{tot}} \sum_{M_* \leq m} m(R)_i \quad (4.4)$$

Where $M_{tot} = \sum_{all} M_*(R)$. This factor will normalize the functions such that they are true CDFs. These curves can be seen for the six samples in Fig. 4.1. To compare this to the X-Ray data, we also need to compute a radial density function for the flux. We once again look for something similar to the form of equation (4.2).

The flux distribution function can be found by summing up the total flux from all pixels in an annulus of specified size ΔR . This can be written as:

$$f(R) = \sum_{pix} i(R + \frac{\Delta R}{2} > R > R - \frac{\Delta R}{2}) \quad (4.5)$$

Where i is the individual pixel flux of each pixel found inside the annulus and $f(R)$ is the total flux found for each radial coordinate. Each radial coordinate is centred on the BCG. We may then find the CDF of the flux as:

$$F_i(R) = \frac{1}{f_{tot}} \sum_{f(R) \leq f_{tot}} f(R) \quad (4.6)$$

Where $f_{tot} = \sum_{all} f(r)$. With both the radial CDF of the galactic mass and of the X-Ray flux, we look to compare them.

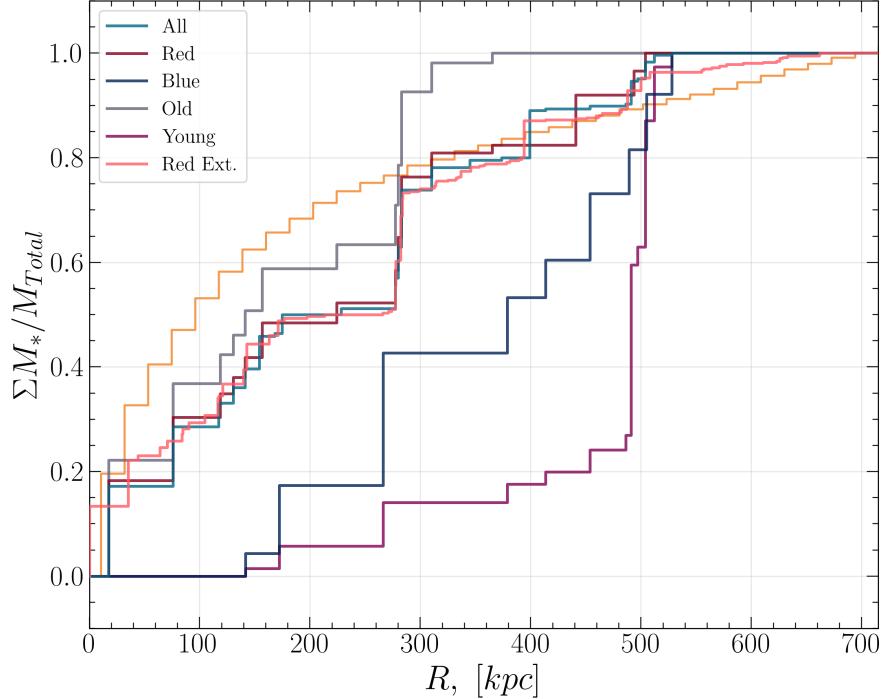


Figure 4.1: Normalized cumulative radial mass distributions for each of the samples summed to last mass in each sample. Included is the X-Ray flux profile in yellow, normalized to the radial coordinate corresponding to the last mass in the red extended sample. Some ambiguity present as to where the normalization should occur.

To do this, the two-side Kolmogorov-Smirnov is used. The goodness-of-fit test between two independent samples, compares how close to each other the samples appear. The null hypothesis in this case is that the samples are drawn from the same distribution while the alternative is that they are different entirely. The statistic that quantifies this difference is called the KS D-statistic and can be computed using the following (Dodge, 2008):

$$D = \sup_x |F(x) - H(x)| \quad (4.7)$$

Where \sup denotes the largest value and x is the independent random variable. In our case, we may rewrite this as:

$$D = \sup_R |F_m(R) - F_i(R)| \quad (4.8)$$

Since we are dealing with normalized CDFs as input to the test, the D-statistic ranges from

0 to 1 and describes the vertical distance the distributions differ by. The values for our distribution can be seen in Fig. 4.2 and largely show what can be eyeballed in Fig. 4.1. Clearly the young and blue collections of galaxies differ significantly from the X-Ray flux distribution with each of the other galaxy sub-samples following it to some degree. The feature that is most likely the value of the D-statistic is a ledge of *missing* galaxies near the $R = 200$ kpc mark, but the basic shape is followed by our more mature samples of galaxies.

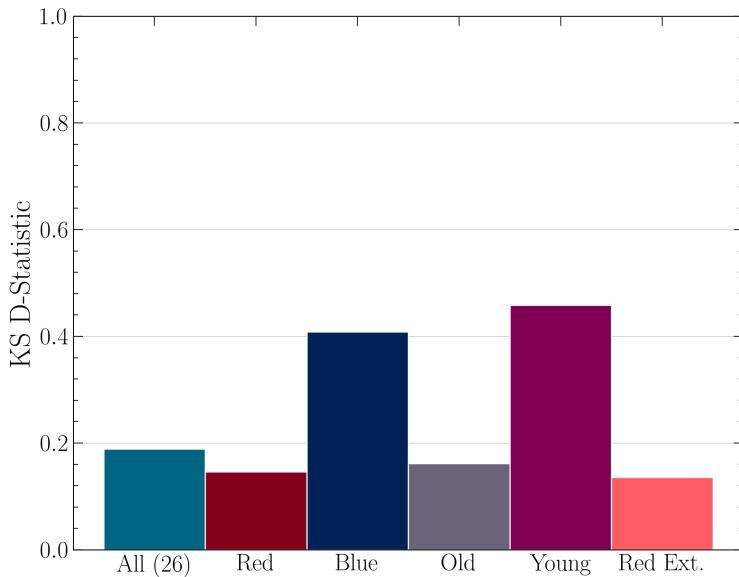


Figure 4.2: KS-Test D-statistic from the distributions in 4.1. All values fall below 0.5 with the blue and young samples showing considerably larger D-statistics.

To interpret the value of the D-statistic, we can accept or reject the null hypothesis based on a standard p-value which is derived from the underlying distribution of the D-statistic. Contrary to other statistical tests, our null hypothesis is that the distributions are in fact drawn from the same underlying one and so to say that the flux and mass distributions are in fact very different, we look for p-values < 0.05 . Large p-values close to 1 indicate that it is likely that the distributions could be drawn from the same for a particular D-statistic indicated for each sample in Fig. 4.2.

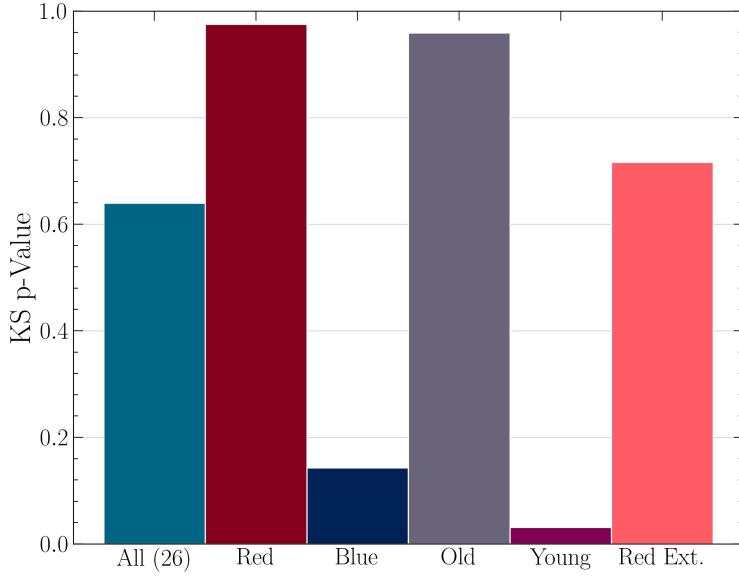


Figure 4.3: KS-Test p-value from the distributions in 4.1 corresponding to the critical values of the D-statistics from 4.2.

Using this, we can make conclusions about each of the sub-samples radial profiles with respect to the X-Ray emission from the plasma in the cluster. As can be seen in Fig. 4.3, the red and old sub-samples of galaxies indicate a considerable statistical match between the two distributions for these samples. Followed closely behind by the extended red and then the entire member sample, we are unable to reject the null hypothesis of these four sub-samples and can thus say that for the given D-statistics, the radial mass CDFs of the galaxies in these sub-samples could have been drawn from the same underlying mass distribution that the gas in the cluster follows. We once again interpret this using the flux of the X-Ray emission as a proxy for spatial mass distribution of the clusters dark matter, and can say that the old and red galaxies follow a similar *radial* profile. Of note here is that the p-value for the blue sub-sample does not exactly meet the criteria for null hypothesis rejection with a value of $p = 0.14 > 0.05$. There is not enough evidence to say that this sub-sample does not follow the underlying distribution at all, but matches considerably less than the others. Lastly, we may notice that the p-value for the young sample $p = 0.03 < 0.05$ and so we can reject the null hypothesis outright. This means that the young galaxies *do not* follow the underlying radial dark matter distribution of the cluster.

We have now seen that the galaxies that show properties of red colour and old age follow

the radially symmetric mass distribution of the cluster while the younger and bluer galaxies, with varying degrees of agreement, do not follow this profile. As the cluster is currently evolving, our assumption that the spatial placement of each constituent is symmetric along radial shells surrounding the BCG is not true but has allowed us insight into the CDFs of each sub-sample. To fix this gap in our understanding, we finally look to take all three dimensions of the cluster into account by generating spatially distributed mass densities for each sub-sample and compare them to the gas emission.

Chapter 5

Spatial Correlation Techniques

The final comparison test between the X-Ray gas emission and galaxies in the cluster comes in the form of a spatial density analysis. To do so we look to perform a kernel density estimator method onto the distribution of galaxies to generate a mapping in space that can then be compared to the spatial resolved flux of X-Rays. The purpose of a kernel density estimator (KDE) is to estimate an underlying probability density function given a sampling of points within that probability density (Weglarczyk, 2018). As it applies to our scenario, if we can compute an estimate for the mass density underneath the positions of the galaxies, we may compare this to that of the X-Rays, testing whether the galactic and dark matter mass distributions are comparable.

A kernel density estimator can be written as (Weglarczyk, 2018):

$$\hat{f}(x') = \frac{1}{n} \sum_{i=1}^n K(x_i, x') \quad (5.1)$$

Where we have n observations in the sample taken from an unknown probability density function $f(x')$. $K(x_i, x')$ assigns each i th sample point a new value according to the kernel function that is chosen (Weglarczyk, 2018). There is some choice in how sampled points get spread out to the whole distribution. For a probability density in two dimensions, we can extend our definition of the KDE as either the product of two individual kernels as below (Weglarczyk, 2018):

$$\hat{f}(x', y') = \frac{1}{nh_x h_y} \sum_{i=1}^n K\left(\frac{x_i - x'}{h_x}\right) K\left(\frac{y_i - y'}{h_y}\right) \quad (5.2)$$

Or as a radial kernel density estimator (Weglarczyk, 2018):

$$\hat{f}(x', y') = \frac{1}{nh_x h_y} \sum_{i=1}^n K \left(\sqrt{\left(\frac{x_i - x'}{h_x} \right)^2 + \left(\frac{y_i - y'}{h_y} \right)^2} \right) \quad (5.3)$$

Where h_x, h_y are smoothing coefficients. For our analysis, since the cluster does have radially weighted features, the second option is chosen. As mentioned previously, the choice of $K(x', y')$ depends on the scenario. A number of different ones were used to map the galaxy mass distribution but the two most effective choices were a circular aperture (a radial extension of the boxcar kernel function) and a Gaussian kernel aperture.

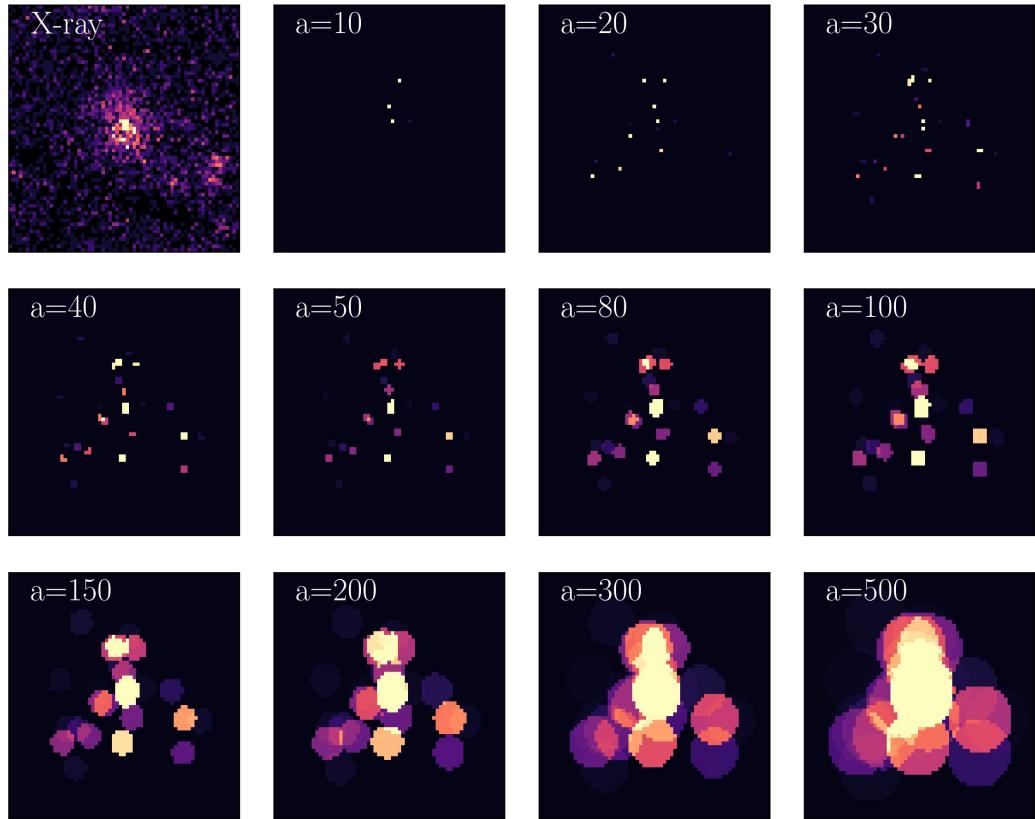


Figure 5.1: Spatial mass densities generated using a circular aperture KDE with varying radii (in kpc). Sample shown here is the 26 member (all) sample of galaxies. The X-Ray flux distribution is included in the top left for comparison. The generated density map has been pixel matched to the coordinate frame of the X-Ray image.

We can write the analytic form of the circular kernel as follows:

$$K_{circ}(r') = \begin{cases} 1 & \text{if } |r'| \leq \frac{a}{2} \\ 0 & \text{otherwise} \end{cases} \quad (5.4)$$

Where $r' = \sqrt{\left(\frac{x_i - x'}{h_x}\right)^2 + \left(\frac{y_i - y'}{h_y}\right)^2}$ and a is the aperture diameter of the circular kernel. h_x and h_y have been chosen to be = 1. The effects of changing the a parameter can be seen in Fig. 5.1 where a is in units of kpc . Much of the same features are present regardless of a change in this parameter but the individual galaxy mass contribution begins to merge at large a .

Now that we have an estimate of the background mass distribution that the galaxies position and mass could be drawn from, we compare it with that of the X-Ray flux for all sub-samples. This is done once again by using a simple Pearson-r correlation between maps in 2D. The results of this can be seen in Fig. 5.2

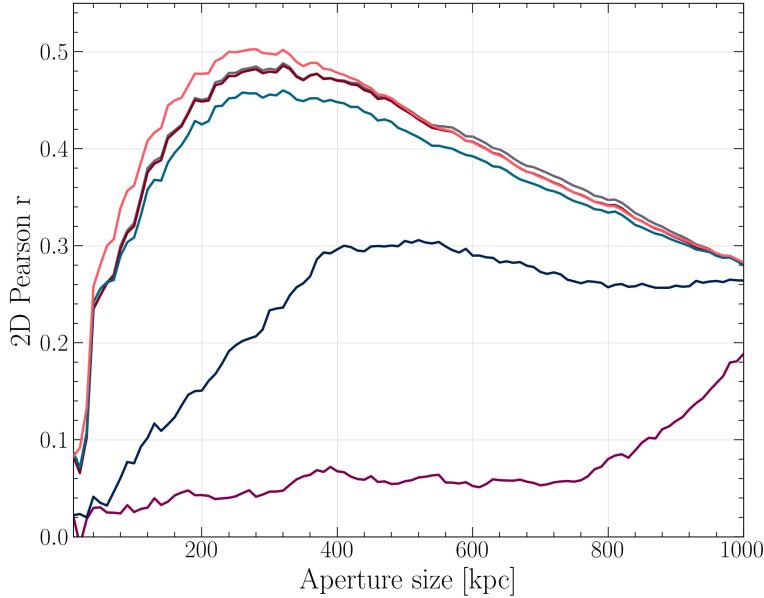


Figure 5.2: 2D Pearson r coefficient as a function of aperture size using a circular aperture with all 6 previously defined samples. Plot gives the scale with which the two distribution types are correlated. Red extended sample is most correlated followed by the red and old samples. Blue becomes moderately correlated at high apertures while the young galaxies are uncorrelated.

Not only does this give us an idea of how each sub-sample are correlated with the X-Ray

emission, it also gives an estimate on which radial length scales we are most correlated. The red extended sample is seemingly most correlated with the gas emission which is likely from the addition of the faint large radius galaxies in this sample which allow for density matches in the outer regions. Following close behind are the red and old samples as usual. These samples all peak around 250 kpc and then decrease. In contrast, the blue galaxies peak much later at a much more moderate correlation coefficient and the young galaxies exhibit almost no correlation until the entire central region is lit up at very large apertures.

It is productive to compare these results with another aperture to see the behaviour of the correlation. We use a Gaussian aperture of the form:

$$K_{Gau}(r') = \frac{1}{\sqrt{2\pi a^2}} e^{-\frac{r'^2}{2a^2}} \quad (5.5)$$

Where a is a comparable aperture size parameter as before but instead of the diameter of the aperture it is the deviation of the Gaussian. We can see how this kernel differs from the original one in Fig. 5.3. There will be considerable spreading out of the mass density with a Gaussian kernel with respect to the circular one.

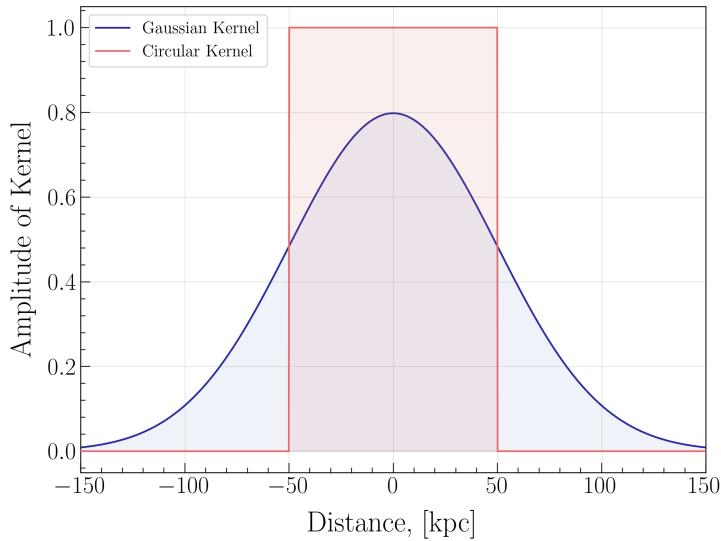


Figure 5.3: Gaussian and Circular kernel with a parameter $a = 50$ kpc. The area under each curve remains the same, but the mass distribution after the application of the kernel for a Gaussian kernel is considerably more spread out.

The application of this aperture can be seen in Fig. 5.4 with a much smoother profile than before. At sufficiently large values of a the mass profile becomes spherically symmetric as all mass is included in a single aperture.

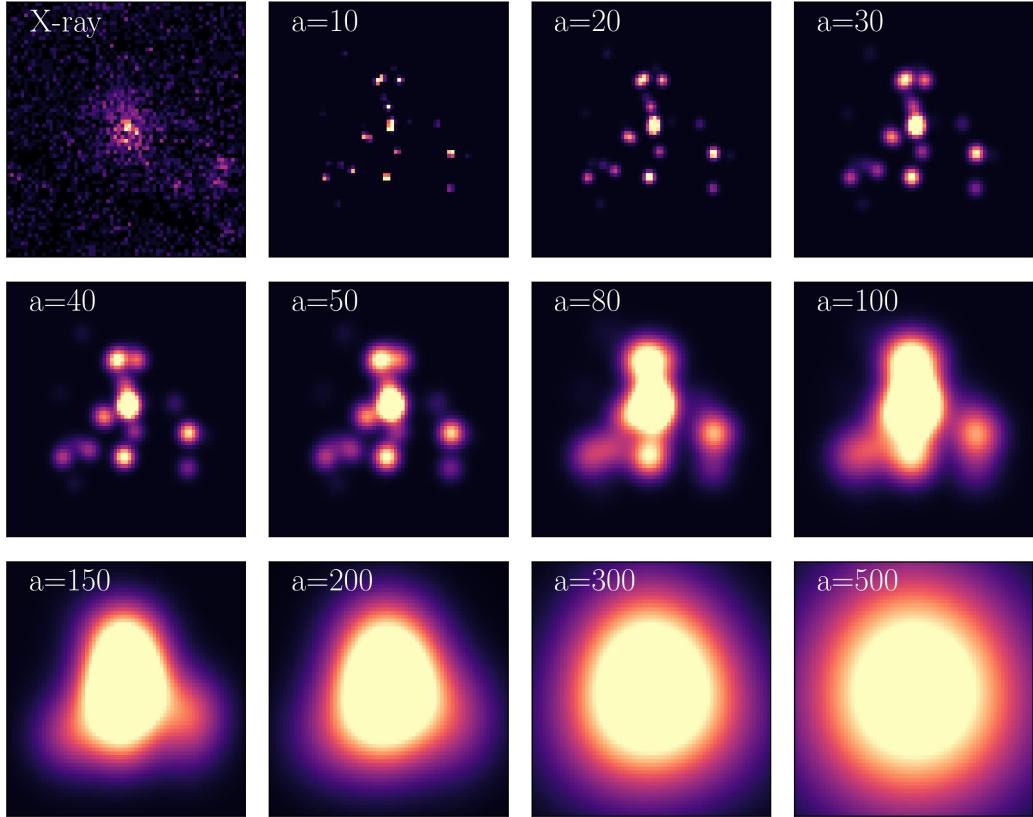


Figure 5.4: Spatial mass densities generated using a Gaussian aperture KDE with varying 'a' parameter (in kpc). Sample shown here is the 26 member (all) sample of galaxies. The X-Ray flux distribution is included in the top left for comparison. The generated density map has been pixel matched to the coordinate frame of the X-Ray image.

Once again, the generated density maps are correlated to the X-Ray emission map. The results of this can be seen in Fig. 5.5. Many of the features are largely the same as for the circular kernel with the red extended sample leading in correlation followed closely by red, old and all members with blue showing moderate correlation lagging behind in aperture size to the other four. Again, the young galaxies do not follow the emission of X-Rays at all. Of note here is the scale that they correlate maximally. Instead of 250 kpc, we see a maximum at ≈ 80 kpc. Of course this is not a one-to-one comparison, but for a Gaussian aperture,

the maps match significantly better for the same value of the a parameter at lower a s and vice-versa.

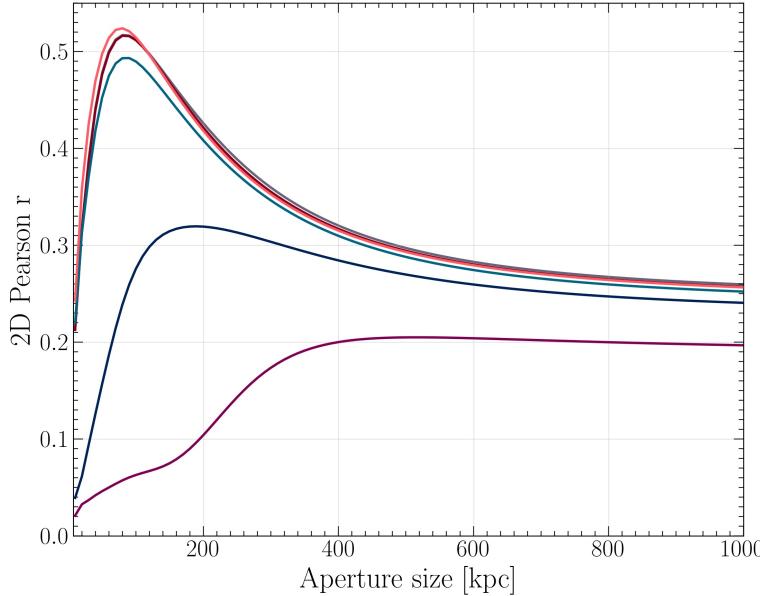


Figure 5.5: 2D Pearson r coefficient as a function of 'a' parameter using a Gaussian aperture with all 6 previously defined samples. Plot gives the scale with which the two distribution types are correlated for a Gaussian kernel. Red extended sample is most correlated followed by the red and old samples. Blue becomes moderately correlated at high apertures while the young galaxies are uncorrelated. When compared to a circular aperture, the Gaussian becomes correlated at smaller length scales.

To demonstrate the stark contrasts between the generated density maps, we may look to Fig. 5.6. The contributions from each galaxy can be seen, showing that each sub-sample vary significantly in their spatial distribution on the sky. The redder/older samples include the BCG galaxy and so much of the density is centralized. The extended red, of course, has density at further outskirts of the cluster, while the oldest of the galaxies seem to be closer to the centre.

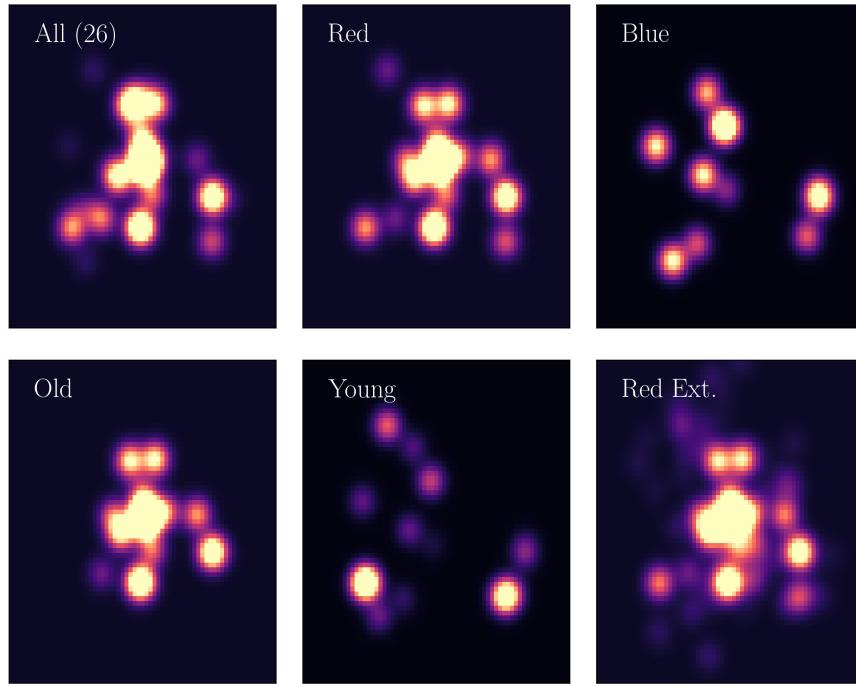


Figure 5.6: KDE generated maps of XLSSC122 using a Gaussian aperture with $a = 50$ kpc for all six samples. The aperture parameter corresponds to roughly the most correlated of the maps with the X-Ray image.

In addition to simple Pearson-r correlation, we may also generate density maps of the *correlation* using a technique implemented here (Virtanen et al., 2020). This is done by computing the normalized cross-correlation between two data sets and mapping the strength of correlation at that place in the images. We can see the result of this for all six samples and the X-Ray image in Fig. 5.7. As we have mentioned in Ch. 1, the majority of correlation in both discrete correlation routines and here is dominated by the central most massive, red galaxies. For sub-samples that do not include this, the cross-correlation mapping is significantly asymmetric in the case of the blue and young sub-samples.

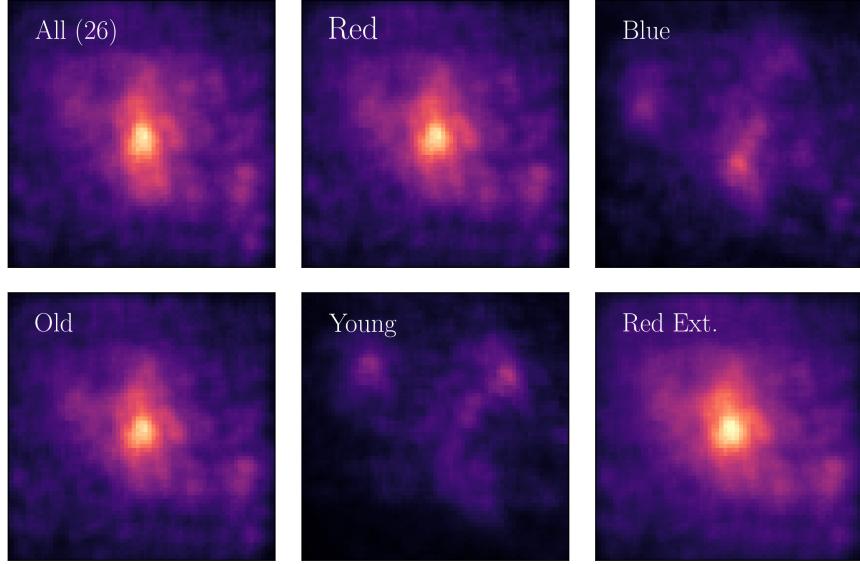


Figure 5.7: Maps of normalized cross-correlation between X-Ray image and kernel density estimators for all six samples for the case of a circular aperture with radius $a = 100$ kpc. Shows the asymmetric correlation density between the two maps. There is a clear difference in the old/red samples vs. the blue and old.

The four correlated sub-samples demonstrate very similar cross-correlation density maps which is expected as they largely consist of similar galaxies in each sub-sample. The inclusion of the central brightest cluster galaxies dominate the maps. The extended red has density extending further out from the BCG as expected which is likely the case of higher correlation in the above analysis. The density present in the young and blue sub-samples is more diffuse and not centralized. The strength of the cross-correlation across the sub-samples differs significantly. The older and redder samples have a clear central high-density of correlation while the young and blue samples lack any spatial correlation.

We have now performed three different statistical techniques to analyze the nature of correlation between the gas mass emission and spatial distribution of galactic mass in galaxy cluster XLSSC 122. Being able to do so is important in describing the population of galaxies in the cluster as well as confirming the maturity of many of these galaxies with respect to the overall cluster. We began with a one-to-one discrete Pearson-r correlation between galactic mass and the X-Ray flux found under each particular galaxy. This was followed by a

radially symmetric CDF comparison using the KS-test where each sub-sample's mass profile was computed in conjunction with the profile of X-Ray flux as a function of radius from the BCG. Our final test involved computing mass density maps using a kernel density approach with multiple kernels. These maps were then correlated with the spatially distributed X-Ray flux. With the results of each of these tests calculated, we may now look to piece together the puzzle pieces to make conclusions about the properties of the member galaxies in the cluster.

Chapter 6

Analysis and Findings

Our original hypothesis regarding the nature of galaxy cluster XLSSC 122 proceeding the analysis undertaken here was that the cluster was indeed remarkably mature for such an early place in the universe. To investigate this claim, we were fortunate to assemble various puzzle pieces of data together to illustrate a statistical narrative of what it means to be mature in the context of individual galaxies in the confines of an underlying dark matter potential. As a reminder, the purpose of the analysis here is to test the claim that the population of galaxies found in this cluster follow what would be expected of a grouping that has had sufficient cosmic time to form, fall in to the dark matter halos, and become virialized with the surrounding environment. To validate this, positions and masses of confirmed member galaxies of the cluster were compared to the magnitude and spatial distribution of the intra-cluster medium gas plasma. This hot gas should, in principle, already have had ample time to distribute according to the potential surfaces of the dark matter profile allowing us to use it as a diagnostic tool for the state of equilibrium of the galaxies within it.

The testing of our hypothesis was done in three separate statistical paradigms. The first test was designed to investigate if the flux of the X-Rays in the region of each galaxy within the cluster is related in any way to the mass of the galaxies. This is a useful test to see if larger mass galaxies are more likely to follow areas of higher density flux. If the cluster were indeed in equilibrium with its surroundings, we would expect this to be the case. To bolster the ability of this technique to validate our claim, the member galaxies were split into sub-samples defined by their relative age and colour with respect to the membership as a whole. A secondary extension to our hypothesis is such that we would expect galaxies that have had longer to form and exist in the halo to better match it. We should see, then, that the older and by extension the redder galaxies more accurately represent a virialized population with the cluster. The splitting of the sample allows us to then test the claim that

the presence of distinct sub-populations of galaxies within the members behave differently in the dark matter potential.

Although we used a number of ways to obtain the flux in the region near each member galaxy, it was concluded that the trends were largely agnostic to the method used. Furthermore, these trends demonstrated a clear difference in the way that each sub-sample's mass was linearly correlated with the resulting X-Ray flux. For the technique outlined in 3, there are four major conclusions that we may draw from the method. The primary result of note is that the redder and older galaxies are in fact correlated with the X-Ray flux to both a high degree and of one of statistical significance. Conversely, the younger and bluer populations of galaxies are either uncorrelated, or were so as to be not of statistical significance. Our second takeaway from this test is that samples that include the brightest cluster galaxies at the center of the cluster have correlations dominated by these galaxies. This is not unexpected as the central galaxies find themselves also at the centre of the X-Ray profile and as such at the centre of the dark matter over-density. Our third conclusion is that the behaviour of blue galaxies is not identical to that of the young galaxies, with the young galaxies being the least correlated. This is somewhat important to note as, although 'redness' and age are highly correlated, they do not quite describe the same thing. This means that the youngest population is the least correlated, suggesting that there is a connection between the absolute age of a galaxy and their relationship with the underlying dark matter. Lastly, when extending the red sub-sample to outer regions of the cluster, we see a moderate correlation when compared to the old and red samples. Although the correlation is of lower magnitude, it is of far greater statistical significance.

Now that we have established that our initial technique proves a bimodality within the sub-samples with respect to the relationship between themselves and the surrounding gas, we may look to gain further insight from the second statistical method analyzing the radial profiles. This test assumes a radial symmetry present within the cluster that is nonphysical but with sufficient members aids in our understanding of the nature of the cluster. The cumulative density functions were generated by summing up the mass of all galaxies present in a given sub-sample. These can be seen in 4.1. Looking at this figure, a number of important features are immediately noticeable. The first is that the older and redder galaxies qualitatively follow the yellow X-Ra flux profile. Within these profiles, we can see that there is a markedly pronounced knee in the curves at $R > 250$ kpc. This is important for either one of two reasons. The first is that the galaxy truly lacks old and red galaxies at these coordinates in the sky. This is unlikely but not impossible and might mean that there is similar deficiency in dark matter in this annulus as well.

Alternatively, we may notice that this falls nearby the radius of r_{500} . This is the region where the field of view of the grism (a combination of a diffraction grating and spectral prism) used by Willis et al. [2020] is truncated, meaning that the spectra of galaxies within these regions may be incomplete. This suggests that the deficit in galaxies here may in fact be an artifact of the data when confirming the membership of cluster galaxies. Regardless of the reason why this deficiency exist, it brings up a question of normalization. Since we are dealing with normalized CDFs, the choice of where the X-Ray flux profile is normalized is dependent on what we compare it to. In our analysis, it has been normalized to the radial coordinate of the last member of a sub-sample. Further analysis using this technique may include a more robust or intricate way to deal with this problem, since the choice of normalization impacts the quantitative results considerably. Furthermore, with the realization that there may be an artifact of the grism size present in our profile, it may be advantageous to only perform this normalization out the coordinate of the knee.

The conclusions to be drawn from this method mirror those from method one. As we can see in 4.2, the same samples exhibit the positive result of the Kolmogorov-Smirnov test and vice-versa. We see once again that the youngest of the galaxies positions and masses are drawn the furthest from the X-Ray profile followed by the blue galaxies. This is supported by the small p-values for these two samples in the KS-test. Conversely, the redder and younger samples have much smaller D-statistics backed up by our inability to reject the null hypothesis for this test for these four samples. Of note here is that the smallest D-statistic results from the extended red sample. We may conclude again that the bimodal nature of the galaxy cluster members persist with the sub-set of old and red galaxies following a statistically similar distribution of cumulative radial mass with that of the X-Ray profile.

As seen prior, we have dealt with the cluster members in both a simple mass-intensity linear correlation routine as well as a radially symmetric cumulative density profile method involving the Kolmogorov-Smirnov test. The last necessary test is to account for the asymmetries present in the way that both the galaxies and the gas are distributed in space. To do this, maps of galactic mass density were generated using a kernel density estimator approach. This was done with both a circular aperture kernel and a Gaussian kernel with somewhat differing results. When performing correlation on either map with the X-Ray data, we see that the trends hold regardless of aperture but the length scales of the features are different. As seen in Figs. 5.2 and 5.5, the redder, older samples experience a much higher correlation independent of the aperture size. We also see that the blue sample has some intrinsic correlation at larger apertures, likely due to the increasing radius; the young sample is uncorrelated once more. For both the circular and Gaussian aperture, the extended

red leads the correlation rate. What is of particular note is both the shape and the location of the maximum for the two methods. Both curves increase with aperture followed by a tail of decrease. This is not unexpected as the scale of the apertures become larger than the feature in the X-Ray image, however the peak of this maximum might also be due to the artifact caused by the grism discussed above. If we again are missing member galaxies at the boundary of the grism field of view, we would see a marked decrease in the correlation at these scales. The size of the aperture where we begin to see this decrease for a circular aperture is Uncoincidentally right near where r_{500} is in the cluster. We do not see quite the same behaviour from the Gaussian aperture, but this can be explained by the discussion surrounding Fig. 5.3.

Alongside the conclusion that we may draw from the r-coefficient, a more qualitatively description of the matching of the density maps with the X-Ray image can be illustrated by combining the results of Figs. 5.6 and 5.7. From the former, we see the way that the delineations of the sub-samples effect their respective spatial distributions. We can see from the X-Ray flux image that it is largely localized in the centre with a first-order symmetry along the radial axis. It is then obvious that any sub-samples without considerable density in the middle will be naturally uncorrelated. Further, we may see that the radial extent of the sub-samples is different even if their statistical results appear similar. The oldest galaxies are localized near the centre; younger and bluer galaxies seem to avoid this region. Similarly, the extended red samples does as its namesake and extends the density profile to outer regions of the cluster. With the morphology of each sub-samples density profile somewhat understood, we may then look to the 2-dimensional cross-correlation function in Fig. 5.7 to wrap up our final conclusions. All of the features mentioned previously are hidden within this plot. We may clearly see that there is a bimodal distribution within the galaxies with respect to how they correlate with the X-Rays with the red and young contributing much more. The central region also dominates the correlation as mentioned before. Once again, the extended red sample brings the density of correlation out to larger radii. We also see that the blue sample has a larger density than the young one, with both lacking the central dominating region but with a surrounding diffuse cloud of low-density correlation.

In taking all of our statistical techniques in aggregate, we may now relate our conclusions with the underlying mechanisms in the cluster. The domination by the central region in all three techniques is due to both the high mass of the BCGs in the cluster combined with the (likely) highest density of gas in this region. The dark matter profile is also the highest density here as well. As for the low correlated sub-samples, we may draw two conclusions from our results here. The first is that both the young and blue galaxies in the cluster

do not follow the underlying dark matter profile up to the limits of our methods. A more thorough analysis would further split the members and investigate which of each type are best described with the gas flux. The second conclusion is that these populations of galaxies are largely found outside the central region of the cluster. This means that it is likely that the galaxies that do not already follow the dark matter profile are in the process of infalling on the cluster. Due to the hierarchical nature of large scale structure formation, it is not unlikely that there is a relatively continuous streaming of matter in the form of galaxies and mass onto the cluster, where these younger and bluer galaxies are still in the process of doing so.

On the other side of things, the red and old samples were the highest (and significantly) correlated samples in all three tests. This means that within the entire population of galaxies, we may conclude that there is indeed a sub-set of galaxies that have reached virial equilibrium with the surroundings. This is quite consistent with the original hypothesis, where we expected much of the cluster to follow the dark matter potential. We have then demonstrated that exists both a collapsing population of young and blue galaxies waiting to reach equilibrium, and an embedded group of red and old galaxies that have already established their place within the dark matter distribution. Both of these results are congruent with the idea that XLSSC 122 is indeed a *mature* galaxy cluster demonstrating properties consistent with that of a highly developed large-scale structure within a much younger Universe than previously expected.

Bibliography

- A. Diaferio, S. Schindler, and K. Dolag. Clusters of galaxies: Setting the stage. *Space Science Reviews*, 134(1–4):7–24, February 2008. ISSN 1572-9672. doi: 10.1007/s11214-008-9324-5. URL <http://dx.doi.org/10.1007/s11214-008-9324-5>.
- Yadolah Dodge. *Kolmogorov-Smirnov Test*, pages 283–287. Springer New York, New York, NY, 2008. ISBN 978-0-387-32833-1. doi: 10.1007/978-0-387-32833-1214. URL <https://doi.org/10.1007/978-0-387-32833-1214>.
- Katherine Garrett and Gintaras Duda. Dark matter: A primer. *Advances in Astronomy*, 2011:1–22, 2011. ISSN 1687-7977. doi: 10.1155/2011/968283. URL <http://dx.doi.org/10.1155/2011/968283>.
- Andrey V. Kravtsov and Stefano Borgani. Formation of galaxy clusters. *Annual Review of Astronomy and Astrophysics*, 50(1):353–409, September 2012. ISSN 1545-4282. doi: 10.1146/annurev-astro-081811-125502. URL <http://dx.doi.org/10.1146/annurev-astro-081811-125502>.
- A. B. Mantz, Z. Abdulla, S. W. Allen, J. E. Carlstrom, C. H. A. Logan, D. P. Marrone, B. J. Maughan, J. Willis, F. Pacaud, and M. Pierre. The xxl survey: Xvii. x-ray and sunyaev-zel'dovich properties of the redshift 2.0 galaxy cluster xlssc 122. *Astronomy & Astrophysics*, 620:A2, November 2018. ISSN 1432-0746. doi: 10.1051/0004-6361/201630096. URL <http://dx.doi.org/10.1051/0004-6361/201630096>.
- Luca Mascolo, Alexandro Saro, Tony Mroczkowski, Stefano Borgani, Eugene Churazov, Elena Rasia, Paolo Tozzi, Helmut Dannerbauer, Kaustuv Basu, Christopher Carilli, Michele Ginolfi, George Miley, Mario Nonino, Maurilio Pannella, Laura Pentericci, and Francesca Rizzo. Forming intracluster gas in a galaxy protocluster at a redshift of 2.16. *Nature*, 615:809–812, 03 2023. doi: 10.1038/s41586-023-05761-x.

M. Pierre, F. Pacaud, C. Adami, S. Alis, B. Altieri, N. Baran, C. Benoist, M. Birkinshaw, A. Bongiorno, M. N. Bremer, M. Brusa, A. Butler, P. Ciliegi, L. Chiappetti, N. Clerc, P. S. Corasaniti, J. Coupon, C. De Breuck, J. Democles, S. Desai, J. Delhaize, J. Devriendt, Y. Dubois, D. Eckert, A. Elyiv, S. Ettori, A. Evrard, L. Faccioli, A. Farahi, C. Ferrari, F. Finet, S. Fotopoulou, N. Fourmanoit, P. Gandhi, F. Gastaldello, R. Gastaud, I. Georgantopoulos, P. Giles, L. Guennou, V. Guglielmo, C. Horellou, K. Husband, M. Huynh, A. Iovino, M. Kilbinger, E. Koulouridis, S. Lavoie, A. M. C. Le Brun, J. P. Le Fevre, C. Lidman, M. Lieu, C. A. Lin, A. Mantz, B. J. Maughan, S. Maurogordato, I. G. McCarthy, S. McGee, J. B. Melin, O. Melnyk, F. Menanteau, M. Novak, S. Paltani, M. Plionis, B. M. Poggianti, D. Pomarede, E. Pompei, T. J. Ponman, M. E. Ramos-Ceja, P. Ranalli, D. Rapetti, S. Raychaudury, T. H. Reiprich, H. Rottgering, E. Rozo, E. Rykoff, T. Sadibekova, J. Santos, J. L. Sauvageot, C. Schimd, M. Sereno, G. P. Smith, V. Smolčić, S. Snowden, D. Spergel, S. Stanford, J. Surdej, P. Valageas, A. Valotti, I. Valtchanov, C. Vignali, J. Willis, and F. Ziparo. The xxl survey: I. scientific motivations xmm newtonobserving plan follow up observations and simulation programme. *Astronomy; Astrophysics*, 592:A1, June 2016. ISSN 1432-0746. doi: 10.1051/0004-6361/201526766. URL <http://dx.doi.org/10.1051/0004-6361/201526766>.

H. I. Ringermacher and L. R. Mead. Model-independent plotting of the cosmological scale factor as a function of lookback time. *The Astronomical Journal*, 148(5):94, oct 2014. doi: 10.1088/0004-6256/148/5/94. URL <https://dx.doi.org/10.1088/0004-6256/148/5/94>.

Thomas Robitaille, Adam Ginsburg, Stuart Mumford, Brigitta Sipőcz, Christoph Deil, Sam Van Kooten, P. L. Lim, Francesco Biscani, Peter Williams, jimboH, Matt Craig, David Stansby, Geert Barentsen, AlistairSymonds, Leo Singer, and Erik Tollerud. astropy/reproject: v0.10.0, January 2023a. URL <https://doi.org/10.5281/zenodo.7584411>.

Thomas Robitaille, Adam Ginsburg, Stuart Mumford, Brigitta Sipőcz, Sam Van Kooten, Christoph Deil, P. L. Lim, Francesco Biscani, David Stansby, Peter Williams, jimboH, Matt Craig, Geert Barentsen, Leo Singer, AlistairSymonds, Erik Tollerud, Fabian Jankowski, James Davies, Ole Streicher, and Sébastien Maret. astropy/reproject: v0.13.0, October 2023b. URL <https://doi.org/10.5281/zenodo.10037214>.

R. J. Rummel. *Understanding Correlation*. University of Hawaii, 1976.

Mauro Sciarratta, Cesare Chiosi, Mauro D’Onofrio, and Stefano Cariddi. Cosmological interpretation of the color–magnitude diagrams of galaxy clusters. *The Astrophysical Journal*, 870(2):70, January 2019. ISSN 1538-4357. doi: 10.3847/1538-4357/aaf00d. URL <http://dx.doi.org/10.3847/1538-4357/aaf00d>.

A Trudeau, J P Willis, D Rennehan, R E A Canning, A C Carnall, B Poggianti, E Noordeh, and M Pierre. The xxl survey. xl ix. linking the members star formation histories to the cluster mass assembly in the $z = 1.98$ galaxy cluster xlssc 122. *Monthly Notices of the Royal Astronomical Society*, 515(2):2529–2547, June 2022. ISSN 1365-2966. doi: 10.1093/mnras/stac1760. URL <http://dx.doi.org/10.1093/mnras/stac1760>.

Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, Stéfan J. van der Walt, Matthew Brett, Joshua Wilson, K. Jarrod Millman, Nikolay Mayorov, Andrew R. J. Nelson, Eric Jones, Robert Kern, Eric Larson, C J Carey, İlhan Polat, Yu Feng, Eric W. Moore, Jake VanderPlas, Denis Laxalde, Josef Perktold, Robert Cimrman, Ian Henriksen, E. A. Quintero, Charles R. Harris, Anne M. Archibald, Antônio H. Ribeiro, Fabian Pedregosa, Paul van Mulbregt, and SciPy 1.0 Contributors. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17:261–272, 2020. doi: 10.1038/s41592-019-0686-2.

Stanislaw Weglarczyk. Kernel density estimation and its application. *ITM Web of Conferences*, 23:00037, 01 2018. doi: 10.1051/itmconf/20182300037.

J. P. Willis, R. E. A. Canning, E. S. Noordeh, S. W. Allen, A. L. King, A. Mantz, R. G. Morris, S. A. Stanford, and G. Brammer. Spectroscopic confirmation of a mature galaxy cluster at a redshift of 2. *Nature*, 577(7788):39–41, January 2020. ISSN 1476-4687. doi: 10.1038/s41586-019-1829-4. URL <http://dx.doi.org/10.1038/s41586-019-1829-4>.

Jon P. Willis. A mathematical description of the universe, January 2022.

E. L. Wright. A cosmology calculator for the world wide web. *Publications of the Astronomical Society of the Pacific*, 118(850):1711–1715, December 2006. ISSN 1538-3873. doi: 10.1086/510102. URL <http://dx.doi.org/10.1086/510102>.

Appendix A

Entire Field of View Images

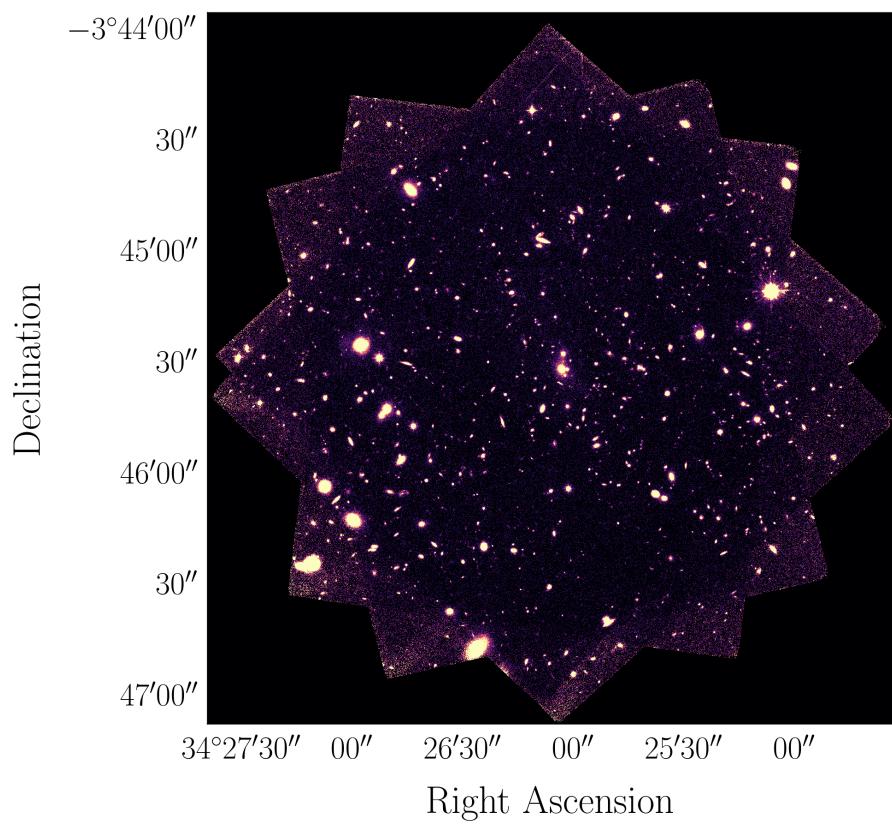


Figure A.1: Image of the Cluster

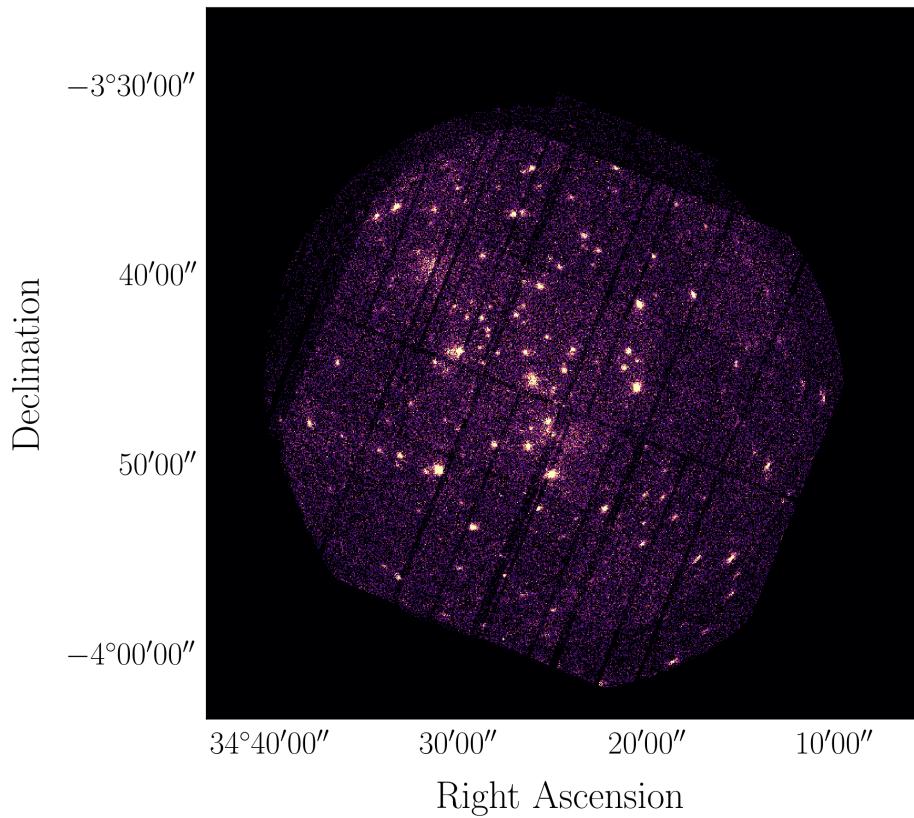


Figure A.2: Xray Image of the Cluster and Surrounding Region

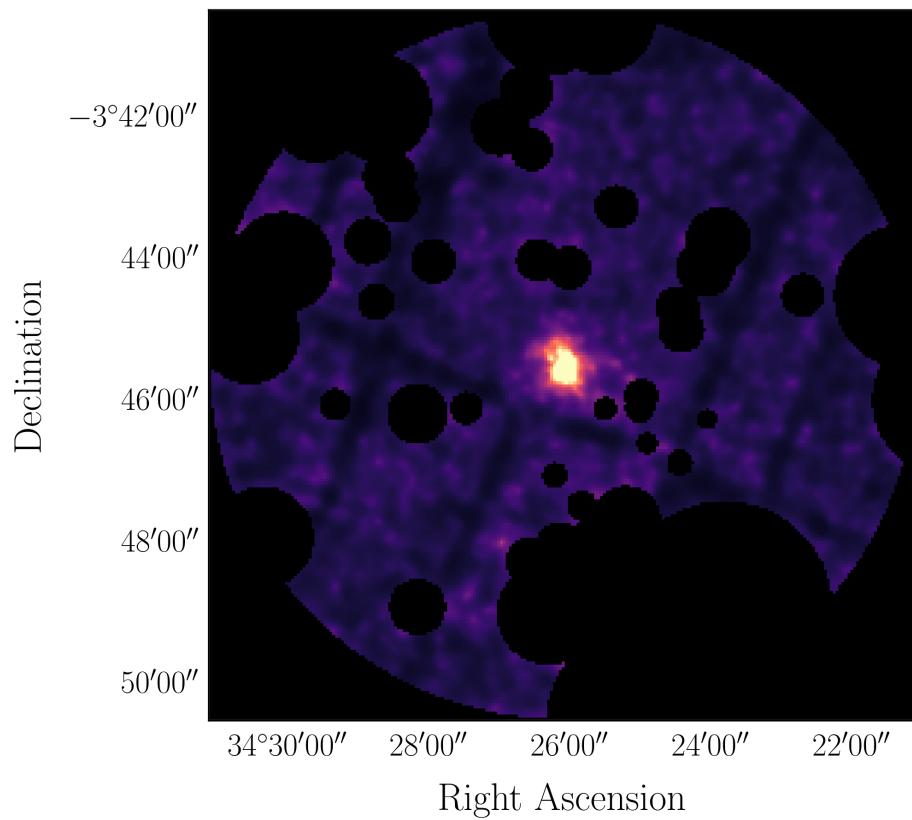


Figure A.3: Smoothed X-ray Image of the Cluster and Surrounding Region