MULTI-WAVELENGTH STRONG LENSING STUDY OF LOW REDSHIFT CLUSTER CORES

by

PAUL EDWARD MAY

A thesis submitted to
The University of Birmingham
for the degree of
DOCTOR OF PHILOSOPHY

Astrophysics and Space Research Group
School of Physics & Astronomy
The University of Birmingham
Sept 2012

Abstract

This thesis covers initial work to increase the total number of galaxy clusters with strong lensing features and, therefore, lensing derived masses. These clusters also possess other multi-wavelength measurements that can be, and were, used for comparison. The work included an investigation into what additional information Dressler Shectman analysis of the incomplete redshift cluster member galaxy samples produced as a result of prioritising mask slit placement on potential strong lensing images. With the test blind on areas centred around the BCG, only a qualitative link was found, providing possible assistance with mass model parametisation.

From the sample of eight clusters reduced, five had strong lensing features, of which, A 3084 had an unusual and rare lensing configuration. Analysis of A 3084 with relatively shallow survey data revealed a cluster with the largest cluster-scale halo centred on the intra-cluster gas and not the BCG: these two were offset from one another. The BCG had a compact DM halo coincident with it, yielding a high substructure fraction of $f_{\rm sub}= 0.73 \pm 0.13$ which, along with smooth X-ray surface brightness contours, bi-modal cluster galaxy redshift histogram and luminosity map, provided evidence that the interpretation of the cluster had suffered a cluster-cluster merger.

From the redshift reduction, five new strong lensing models were produced following similar methods to the ones used for A 3084. With the addition of a cluster from Paraficz et al. (2012), this increased the initial sample size from 17 to 23, providing a $\sim 35\%$ increase, with all clusters having other multi-wavelength measurements. Following previous literature comparisons a fit of $M_{2D}^{SL}(R < 250$ kpc) against T_X was carried out, but large scatter was observed in this. A possible link was found between the residuals of the fit with that of the clusters' BCG ellipticity after a cut was made of $m_{12} \geq 1$ for luminosity gap to remove potentially disturbed clusters.

To my Mother and Father.

ACKNOWLEDGEMENTS

I would like to gratefully acknowledge the help of my fellow PhD students Tim James and Victoria Hamilton-Morris for their help and company while learning the ropes. I would also like to thank the members of the Extragalactic Research Group at Birmingham and the members of the LoCuSS collaboration for their help over the years, with additional thanks to David Stops for his assistance with computing matters.

I would also like to thank my parents, friends and family for their help and support all through the years.

Finally thanking my supervisor, Graham Smith for his support and guidance through my PhD whilst I have been at Birmingham.

STATEMENT OF ORIGINALITY

Both chapters 1 & 2 are based on existing work and literature and are not my own work.

Reduction of *HST*, colour catalogues and J and K band luminosity maps (in chapter 4) were carried out by Victoria Hamilton-Morris and the X-ray reduction and analysis by Alaistair Sanderson.

In chapter 3, collaboration with faint lensed object spectrum extraction and interpretation was provided by Johan Richard (with A 2537 & A 368 redshifts used in strong lensing model constraints derived from his work) and by Graham Smith. In chapter 5, model selection and interpretation was aided by insights from Johan Richard and Graham Smith. The multiple image positions from Johan Richard for A 2537 were used in my strong lens models.

Chapter 4 is based upon a paper that I have under preparation for publication, with the current list of co-authors listed at the beginning of that chapter. The chapter has also benefited from feedback from and discussion with the co-authors (as well as local astronomers).

All redshifts (apart from those already mentioned) and Dressler Shectman statistics were reduced and derived by myself. All strong lensing models and their derived values and analysis were also produced by myself. This was my own work, under the supervision of Graham Smith and with aid provided by relevant colleagues in the LoCuSS collaboration and the Extragalactic Research Group based at the University of Birmingham Physics & Astronomy Department.

Section § 2.2.1 uses equations and information from the book titled 'Gravitational Lensing: Strong, Weak and Micro' by Schneider, P., Kochanek, C. and Wambsganss, J.

(Meylan et al., 2006).

The basic equations and workings in section § 2.4 & § 2.4.1 were taken from "Data Analysis A Bayesian Tutorial" by D. S. Sivia.

The basic equations and workings in this section § 2.4.2 are based on "Bayesian Logical Data Analysis for the Physical Sciences" by Phil Gregory (Gregory, 2005).

Any non-original work is referenced. This work has not been previously submitted for a degree.

CONTENTS

1 Introduction **1**

1.1 Dark Matter . . . 3

1.2 Λ-CDM Cosmology . . . 7

1.3 Galaxy Clusters . . . 10

1.4 Cluster Observables - Global . . . 15

1.4.1 Optical . . . 15

1.4.2 X-ray . . . 18

1.4.3 Millimetre . . . 20

1.5 Cluster Observables - Structural . . . 22

1.5.1 Optical . . . 22

1.5.2 X-ray . . . 24

1.5.3 Millimetre . . . 25

1.6 Gravitational lensing . . . 25

1.6.1 Microlensing . . . 26

1.6.2 Weak Lensing . . . 27

1.6.3 A Short History of Gravitational Lensing by Galaxy Clusters . . . 28

1.6.4 Cluster Lensing . . . 40

1.7 LoCuSS . . . 42

1.8 Thesis Overview . . . 43

2 Theory **46**

2.1 Lensing Theory . . . 46

2.2 Strong Lensing . . . 47
2.2.1 Lens Equations . . . 47
2.2.2 Magnification and Distortion . . . 55
2.2.3 Image Numbers . . . 56
2.2.4 Critical Curves and Caustics . . . 57
2.2.5 Image Classification . . . 57
2.3 Mass Sheet Degeneracy . . . 59
2.4 Bayesian Theory . . . 60
2.4.1 Model Selection . . . 61
2.4.2 Markov Chain Monte Carlo . . . 62
2.4.3 LensTool . . . 66
2.5 Density Profiles of Halos . . . 69
2.5.1 NFW Mass Model . . . 69
2.5.2 PIEMD Mass Model . . . 72

3 Optical Spectroscopy **74**
3.1 Spectroscopic Observations . . . 75
3.1.1 Reduction Method . . . 75
3.1.2 GMOS . . . 76
3.1.3 FORS2 . . . 79
3.1.4 Observations . . . 80
3.2 Redshift Histograms . . . 82
3.2.1 Reflex Additions . . . 84
3.2.2 Velocity dispersion vs T_{X-ray} . . . 86
3.3 Dressler Shectman . . . 88
3.3.1 DS Statistics . . . 89
3.4 Dressler Shectman Analysis . . . 91
3.5 Comparative Observations - *HST* & X-ray . . . 92
3.6 Data Summary . . . 93

3.7 Analysis & Interpretation . . . 94
3.7.1 Abell 2813 . . . 94
3.7.2 Abell 2895 . . . 95
3.7.3 Abell 368 . . . 95
3.7.4 Abell 3084 . . . 101
3.7.5 Abell 3088 . . . 105
3.7.6 Abell 2537 . . . 107
3.7.7 Abell 3364 . . . 107
3.7.8 RXCJ0528 . . . 112
3.8 Dynamical Comparisons . . . 114
3.8.1 Effects of REFLEX additions . . . 116
3.9 Chapter Discussion . . . 118
3.10 Chapter Summary . . . 121

4 Abell 3084 Strong Lensing Analysis 122

4.1 Hyperbolic Umbilic Catastrophe . . . 122
4.2 Observations and Analysis . . . 124
4.2.1 *Hubble Space Telescope* imaging . . . 124
4.2.2 Ground-based near-infrared imaging . . . 126
4.2.3 Optical spectroscopy . . . 129
4.2.4 *Chandra* X-ray observations . . . 134
4.2.5 Summary . . . 140
4.3 Strong-lens Model and Results . . . 141
4.3.1 Model Fitting . . . 143
4.3.2 The Mass and Structure of the Cluster Core . . . 149
4.4 Summary and Discussion . . . 153
4.4.1 Main Results . . . 153
4.4.2 Physical Interpretation . . . 155
4.5 Summary . . . 157

5 Strong Lensing Mass vs X-Ray Temperatures **159**

5.1 Strong Lensing Models . . . 160

5.1.1 Abell 2813 . . . 160

5.1.2 Abell 368 . . . 162

5.1.3 Abell 2895 . . . 164

5.1.4 Abell 2537 . . . 167

5.1.5 Results . . . 167

5.2 Expanded Sample . . . 167

5.3 Data Summary . . . 170

5.4 Analysis . . . 170

5.4.1 Mass Observables . . . 172

5.4.2 Mass vs Temperature . . . 173

5.5 Residuals . . . 176

5.5.1 Substructure Fraction . . . 177

5.5.2 Luminosity Gap . . . 177

5.5.3 BCG Ellipticity . . . 178

5.6 Residual Discussion . . . 180

5.7 Chapter Discussion . . . 182

5.8 Chapter Summary . . . 183

6 Discussion of Work **185**

6.1 Future Work . . . 193

Appendix: Addendum **196**

Appendix: List of References **197**

LIST OF FIGURES

1.1 1E 0657 − 56 (the "Bullet Cluster"). A *HST* colour image of galaxies, with dark matter (blue) and x-ray emitting gas (pink) overlayed, showing the separation of the last two components from one another. Image Credit: X-ray: NASA/CXC/M.Markevitch et al. Optical: NASA/STScI; Magellan/U.Arizona/D.Clowe et al. Lensing Map: NASA/STScI; ESO WFI; Magellan/U.Arizona/D.Clowe et al. Picture taken from: `http://www.nasa.gov/vision/universe/starsgalaxies/dark_matter_proven.html` 5

1.2 Millenium simulation showing the filamentary structure formation in the Universe, with brighter areas signifying larger congregations of mass. Picture taken from "The Millenium Simulation Project" at: `http://www.mpa-garching.mpg.de/galform/virgo/millennium/` 11

1.3 Plot showing the differing shape of $N(M, z)$ (predicted number of clusters on the sky with changing redshift) for differing cosmologies and the resulting change between the slopes of the tail of each function. Figure taken from Voit (2005). 13

1.4 An example of weak shear, with the tick marks showing the distortion direction of background sources. Figure of A383 taken from Smith et al. (2005). 28

1.5 A typical simple arrangement of the source (star), lens (sun) and observer "a". "c" shows the actual star position, "b" the position of the star that the observer ("a") sees, ξ is the impact parameter and α is the deflection angle produced by the lens. 29

1.6 G2237 + 0305, otherwise known as the "Einstein Cross", showing the four multiple images of a single quasar. Picture taken from "First ESA Faint Object Camera Science Images the Gravitational Lens G2237 + 0305" at: `http://hubblesite.org/newscenter/archive/releases/1990/20/image/a/` 30

1.7 LEFT - Figure 7 from Kneib et al. (1995), showing r band imaging with X-ray contour overlay. RIGHT - Figure 1b from Kneib et al. (1996) of the *HST* imaging of A 2218 (zoomed and rotated compared to the left). The improvement in optical imaging is apparent between the two frames. . . . 33

1.8 A true colour reconstruction of a source plane galaxy if observed unlensed by *HST*. The inset image is the same reconstruction after a 0.8 arcsec smoothing has been applied (taken from Swinbank et al., 2007) 34

1.9 Figure taken from Mahdavi et al. (2007, Fig. 8) showing how the dominance of the intra-cluster gas, BCG and DM change depending on the distance from the cluster centre. 36

2.1 Strong lensed multiple images in the system MACS0257. Images of the same background sources are circled in matching colours. Images identified and figure created by author. 48

2.2 A typical simple arrangement of the source, lens and observer showing the deflection produced (as used in equation 2.1 onwards). S is the source image lying on the source plane, S_1 & S_2 are the apparent positions on the source plane of the source after the light rays had been deflected. L is the lens plane and O the observer. Figure taken from Wambsganss (1998) . . . 49

2.3 Critical lines are dotted and caustics are solid lines in this figure. Figure (a) is for a singular isothermal sphere and Figure (b) is for a singular isothermal ellipse. It should be noted that just by changing the symmetry (breaking it in this case), the caustic is now changed from a point to a star shape. Figure taken from Dr. Graham Smith's 2002 thesis - "Gravitational Lensing by X-ray Luminous Galaxy Clusters". 58

3.1 The galaxy cluster member derived velocity dispersion σ_{gal} plotted against the velocity σ_{X-ray} derived X-ray temperature relation for Xue & Wu (2000). The dotted line shows the one to one relation between the two axes, highlighting the systematically higher velocity dispersions on the y-axis. 87

3.2 Top - Abell 2813 *HST* frame with X-rays overlaid in blue contours. Any arcs with measured redshifts are marked. Bottom Left - The resulting cluster member DS analysis. Bottom Right - Redshift histogram with the cluster members highlighted in red. 96

3.3 Top - Abell 2895 *HST* frame with X-rays overlaid in blue contours. Any arcs with measured redshifts are marked. Bottom Left - The resulting cluster member DS analysis. Bottom Right - Redshift histogram with the cluster members highlighted in red. 98

3.4 Abell 2895 multiple images. Top Left - The Lyman-Alpha emission line for the multiple image A1.1 (z=3.39). Top Right - The Lyman-Alpha emission line for the multiple image A1.2 (z=3.39). Bottom - The Lyman-Alpha emission line for the multiple image A2.1 (z=3.72). 99

3.5 Top - Abell 368 *HST* frame with X-rays overlaid in blue contours. Any arcs with measured redshifts are marked. Bottom Left - The resulting cluster member DS analysis. Bottom Right - Redshift histogram with the cluster members highlighted in red. 102

3.6 Top - Abell 3084 *HST* frame with X-rays overlaid in blue contours. Any arcs with measured redshifts are marked. Bottom Left - The resulting cluster member DS analysis. Bottom Right - Redshift histogram with the cluster members highlighted in red. 104

3.7 Top - Abell 3088 *HST* frame with X-rays overlaid in blue contours. Any arcs with measured redshifts are marked. Bottom Left - The resulting cluster member DS analysis. Bottom Right - Redshift histogram with the cluster members highlighted in red. 106

3.8 Top - Abell 2537 *HST* frame with X-rays overlaid in blue contours. Any arcs with measured redshifts are marked. Bottom Left - The resulting cluster member DS analysis. Bottom Right - Redshift histogram with the cluster members highlighted in red. 108

3.9 Top - Abell 3364 *HST* frame with X-rays overlaid in blue contours. Any arcs with measured redshifts are marked. Bottom Left - The resulting cluster member DS analysis. Bottom Right - Redshift histogram with the cluster members highlighted in red. 110

3.10 Top - RXCJ0528 *HST* frame with X-rays overlaid in blue contours. Any arcs with measured redshifts are marked. Bottom Left - The resulting cluster member DS analysis. Bottom Right - Redshift histogram with the cluster members highlighted in red. 113

3.11 Top – Δ/N plotted against $P - Value$ to compare the two DS statistical tests of substructure evidence.Bottom Left – Δ/N plotted against X-ray centroid offset to compare the two values for dynamical evidence. Bottom Right – Δ/N plotted against number of DM halos from SL mass reconstructions. In each plot a Pearson correlation value and null hypothesis probability is given. 117

4.1 Figure 29 from Orban de Xivry & Marshall (2009) showing how the caustics for an elliptical cluster lens change as a function of source redshift. 124

4.2 TOP – The central 30 × 45 arcsec of A 3084, taken from the *HST*/ACS snapshot observations. Three of the five images comprising the hyperbolic umbilic catastrophe image configuration of the galaxy A1 were located adjacent to the BCG. The fourth image, A1.4, was detected after subtraction of the BCG (lower left panel), and the fifth image lay 10 arcsec South of the BCG. A second multiply-imaged galaxy laid ∼ 22 arcsec West of the BCG, marked A2.1, A2.2, and A2.3. LOWER LEFT – Zoom into the central 10 × 10 arcsec, after subtraction of the BCG. The brighter and fainter component (a, and b respectively) of A1.1, A1.2, and A1.3 were clearly visible. A candidate fourth image (A1.4) was detected following the subtraction of the BCG. LOWER RIGHT – Zoom into the 10 × 10 arcsec region South of the BCG where the fifth image, A1.5, was found (two components visible). Note that all four of the securely detected images obeyed the mirror symmetry expected of strongly-lensed images. The dotted lines are centred on the BCG in all panels. Blue (red) lines are the tangential (radial) critical lines. The caustic lines for the mass distribution are shown in light blue (pink) for the tangential (radial) lines. North is up and East is left in all panels. 125

4.3 $(J-K)/K$ colour-magnitude diagram of galaxies within a 10 × 10arcmin field of view centred on the BCG in A 3084, revealing a tight red sequence ridge-line of likely cluster galaxies. The dotted lines mark the selection criteria applied to construct a catalog of cluster members for inclusion in the model of the cluster mass distribution in §4.3 and near-infrared luminosity density map, at right. The galaxies that satisfied this selection function, $1.0 < (J-K) < 1.6$ and $K < K^{\star} + 1 = 16.25$, were plotted as filled symbols. The error bars at the top of the figure show the mean error on $(J-K)$ as a function of K-band magnitude. 127

4.4 The rest-frame K-band luminosity density map of the core of A 3084 (grey-scale and green contours), based on the galaxies selected at left. The map is centred on the BCG, and the dashed box marks the ACS field of view. This panel shows the central $\sim$ 1Mpc $\times$ 1Mpc of the cluster in projection on the sky. North is up and East is left. Figure produced by Victoria Hamilton-Morris. 128

4.5 The reduced one-dimensional spectrum of A1.2, smoothed to the instrumental resolution of FORS2. Oxygen and hydrogen emission lines redshifted to $z = 0.764$ were detected at strong significance. 130

4.6 Redshift distribution of the 37 galaxies identified at $0.2 < z < 0.23$ from the multi-slit observations, plus the two galaxies from Guzzo et al. (2009). The red dashed line shows the mean redshift of $z = 0.2162 \pm 0.0044$; the black dashed line the measured BCG redshift of $z = 0.2177 \pm 0.0002$. The blue histogram shows the total number distribution, the red and black histograms the distribution of galaxies with low and high (respectively) local departure from the global cluster kinematics (see Fig. 4.7). 135

4.7 Spectroscopically confirmed cluster member galaxies plotted as circles of radius proportional to the departure of the local kinematics from global cluster kinematics, following Dressler & Shectman (1988). The figure is centred on the BCG and the plus symbol marks the X-ray centroid. Galaxies that satisfied the $(J - K)/K$ photometric selection described in §4.2.2 that were not spectroscopically confirmed as members are marked as small green crosses. Red and black circles are plotted separately in the redshift histogram shown in Fig. 4.6. The dashed box shows the ACS field of view. North is up and East is left. 136

4.8 Smoothed X-ray surface brightness map based on the *Chandra*/ACIS-I observations (red contours) overlaid on the central 100 × 100 arcsec of the *HST*/ACS snapshot observation. The X-ray emission peak was within ∼ 1 arcsec of the optical centroid of the BCG, whilst the X-ray centroid lay 11.1 arcsec West and 4.8 arcsec South of the BCG (black "plus" symbol). The dashed black box delineates the region shown in Fig. 4.2. Contours were spaced equally in the log, with each contour separated by 0.25dex. Each tick mark on the axes represents 5 arcsec. North is up and East is left. 137

4.9 X-ray temperature profile of the intra-cluster medium, based on the *Chandra*/ACIS-I observations. The profile is centred on the X-ray centroid. The grey shaded area shows the 68% confidence interval around the best-fit model (dashed curve) described in §4.2.4. Figure produced by Alaistair Sanderson. 138

4.10 The central 2 × 2arcmin of A 3084, centred on the BCG, and shows the rest-frame K-band luminosity density of cluster members (as defined in §4.2.2; grey-scale), X-ray flux from *Chandra* (§4.2.4; red contours), and the projected mass density map (§4.3.1; blue contours). The large white plus marks the position of the X-ray centroid. The mass and X-ray contours are centred on the X-ray centroid, reflecting the large-scale structure of the gravitational potential. In contrast, the BCG is off-set from the centre of the potential. All contours and the grey-scale were plotted in the log. North is up and East is left. 146

4.11 Posterior probability density of the 3-halo model discussed in §4.3.1. All model parameters were marginalised over except the x, y co-ordinates of the third cluster-scale halo. Thick, medium, and thin curves showed 68%, 95%, and 99% confidence levels respectively. The figure was centred on the BCG, and the black "plus" shows the location of the X-ray centroid. Despite the very low probability of the data, it revealed that the 3rd halo tended to lay to the South-West, following the K band luminosity map (Fig. 4.4) and the X-ray contours (Fig. 4.8). 148

4.12 Posterior probability density in the $f_{\rm sub} - z_{\rm A2}$ plane for Model D, marginalizing over all other free parameters in the model. The full range of the prior on $z_{\rm A2}$ was plotted on the y-axis – $z_{\rm A2} < 4.8$ because it was detected through the F606W filter with ACS. The degeneracy between $f_{\rm sub}$ and $z_{\rm A2}$ was weak. Thick, medium, and thin curves show 68%, 95%, and 99% confidence levels respectively. 150

4.13 Joint confidence intervals on the core radius of DM1 and $z_{\rm A2}$ from Model D (Table 4.3). Thick, medium, and thin curves show 68%, 95%, and 99% confidence levels respectively. 152

5.1 Abell 2813 with its BCG to the North and the second brightest cluster member situated to its South. The two multiple image systems between the two are labelled A1 and A2. The cross denotes the X-ray centroid position. North is up and East is to the left. 161

5.2 A2.3 spectrum shows the 4000 Å break and several other lines. These gave a redshift of $z = 0.8209 \pm 0.0006$. 162

5.3 Abell 368 with its BCG and one set of multiple images labelled A1 to the West. The cross denotes the X-ray centroid position. North is up and East is to the left. 163

5.4 Abell 2895 with its BCG and the two multiple image systems A1 and A2 arrayed around it. The cross denotes the X-ray centroid position. North is up and East is to the left. 164
5.5 Abell 2895 multiple images. Top Left - The Lyman-Alpha emission line for the multiple image A1.1 (z=3.39). Top Right - The Lyman-Alpha emission line for the multiple image A1.2 (z=3.39). Bottom - The Lyman-Alpha emission line for the multiple image A2.1 (z=3.72). 165
5.6 Abell 2537 with its BCG to the North and the second brightest cluster member situated to its South. The three multiple image systems are labelled A1, A2 and A3. The cross denotes the X-ray centroid position. North is up and East is to the left. 166
5.7 Top – Lensing mass plotted against substructure fraction, $f_{\rm sub}$. Middle – Mass plotted against luminosity gap, m_{12}. Bottom – Mass plotted against BCG ellipticity. 174
5.8 M_{SL} versus T_{X-Ray} in log space. Blue line shows the best fit line and the green lines the error lines after scaling the error bars. 175
5.9 Mass difference from fitted mass as a fraction of fitted mass versus substructure fraction. 177
5.10 Mass difference from fitted mass as a fraction of fitted mass vs luminosity gap. 178
5.11 Mass difference from fitted mass as a fraction of fitted mass vs BCG ellipticity. Those clusters labelled in blue had a luminosity gap magnitude value greater than 1. The Pearson value of -0.18 was calculated for all points (blue and black). 180
5.12 Luminosity gap versus substructure fraction. 181

LIST OF TABLES

1.1 Early Universe History 2

3.1 FORS2 Observing Data 81

3.2 GMOS Observing Data 83

3.3 New object redshift yields 84

3.4 Redshift Galaxy Cluster Sample 85

3.5 Dressler Shectman Results 92

3.6 HST Observing Data 93

3.7 Cluster Members for Abell 2813 97

3.8 Cluster Members for Abell 2895 100

3.9 Cluster Members for Abell 368 103

3.10 Cluster Members for Abell 3088 105

3.11 Cluster Members for Abell 2537 109

3.12 Cluster Members for Abell 3364 111

3.13 Cluster Members for RXCJ 0528.2 − 2942 112

3.14 Strength of Pearson Correlation 114

3.15 DS results with/without REFLEX Additions 119

4.1 Spectroscopically Confirmed Galaxies at $0.2 < z < 0.23$ – "Members" . . . 131

4.2 Other Spectroscopically Confirmed Galaxies – "Non-members" 132

4.3 Gravitational Lens Model Parameters[a] 142

5.1 Strong Lensing Image Constraints 168

5.2 Strong Lensing Modelling Results, with $1\sigma^a$ Errors 169
5.3 Galaxy Clusters 171
5.4 Strength of Pearson Correlation 172

CHAPTER 1

INTRODUCTION

Carl Sagan once wrote:

> There are four Cosmic questions:
>
> 1) Was there ever life on Mars?
> 2) Is Titan a laboratory for the origin of life?
> 3) Is there intelligent life elsewhere?
> 4) What is the origin and fate of the Universe?

Carl Sagan - "Billions and Billions"

Many of these questions have been asked for as long as humanity has been civilized and our understanding of some of them has changed drastically in only the last hundred or so years (Mars was believed to show evidence of canal systems and therefore of civilised life up until the advent of more powerful instrumentation in the early 20^{th} Century). The understanding of each of these questions are several decades of work and this work concentrates on the last question.

Current scientific understanding is that the Universe is expanding. This expansion occurred after an event called the "Big Bang", the results of which can be seen today in the fact that most large extra-galactic bodies are moving away from one another as the Universe expands. There is also a permeating remnant residual radiation from this event called the "Cosmic Microwave Background" Radiation (CMB). CMB has a thermal temperature of $\sim$ 3 Kelvin, but has an uneven distribution due to the fact that, in its

Table 1.1: Early Universe History

Cosmic time	Temperature (MeV)	Events
$t \approx 10^{-4}$ s	$kT \approx 10^{2}$	Quarks form neutrons and protons
$t \approx 1$ s	$kT \approx 1$	Neutrinos decouple
$t \approx 4$ s	$kT \approx 0.5$	Electron-positron annihilation
$t \approx 3$ min	$kT \approx 0.1$	Helium and other light nuclei formed
$t \approx 4 \times 10^{5}$ years	$kT \approx 3 \times 10^{-7}$	Atoms formed and photons decouple

make up, it reflects the "surface of last scattering" as well as the evolution of the Universe over time. Surface of last scattering is the last pattern of scattered photons created before the Universe cooled enough so that photon scattering no longer occurred and, as a result, shows the structure at very early times as a result. Extrapolated back through time, this puts the Universe at an age of over 10 billion years (Phillips, 2003). The current accepted model of the very early history of the Universe is set out in Table 1.1 (taken from Phillips, 2003).

Seeing the Universe at any time when photons had yet to decouple (Table 1.1 sets out the time it took for this to occur) is very difficult for astronomers as the Universe was opaque during this period. Understanding how the Universe began is a field of study using large super-colliders that essentially try to recreate some of the early Universe conditions in order to study the physics that occur. Therefore when understanding how the Universe formed, both astronomy and particle physics overlap. One of the largest multi-discipline areas of study in modern astronomy and cosmology is how, once atoms started to form, they interacted and merged to form the modern day Universe. This thesis specifically deals with one of the largest structures to yet form in the Universe, "Galaxy Clusters" (described in more detail in §1.3) and what information they can provide about the Universe.

Baryonic matter is made up of atoms described by the standard and with which people are familiar with and interact with on a daily basis, it only makes up a small fraction of the matter of the Universe (in the order of $15 - 16\%$). This means that something

else is dominating the matter of the Universe gravitationally, but it cannot be seen. As a consequence this has been called "dark matter" and is something that particle physics is also looking to explain and find a candidate for.

This thesis focuses on these two quantities, dark matter and galaxy clusters; what can be learnt about the two and where and how dark matter relates to galaxy clusters and the formation of the latter. Can studying galaxy clusters provide information about the evolution and ultimate fate of the Universe?

1.1 Dark Matter

Dark matter (DM) is the "glue" that holds the Universe together and has helped mould it into what can be seen today. Without its gravitational interactions the Universe would be very different. Science's understanding of the make up of the Universe has changed considerably over the last 100 or so years. The first signs of the existence of dark matter occurred with the discovery that the galaxies inside galaxy clusters were traveling at velocities too great for them to be contained by just the gravitational potential generated by (all) the observable mass (Zwicky, 1933). Simply put, galaxy clusters should not be able to hold onto their galaxies, but they are observed to do so. Therefore, something must be holding them.

This was then observed on a smaller scale, within individual galaxies, when Babcock (1939) found that the rotational velocity curve for the Andromeda Spiral Galaxy did not behave as expected for the observable matter present. As the distance from the galactic centre increased the rotational velocities did not decrease as they should, indicating a large amount of mass that pervaded and extended beyond the radii of the galaxy. These are commonly called DM halos. This was found in many galaxies and even the Milky Way is not exempt from possessing a DM halo that adds to its total gravitational forces. The shape of the halo enveloping the Milky Way has been studied (Olling & Merrifield, 1998) by examining the thickness and distribution of the gas layers inside it. The pervasive nature

of this DM, that exists in both large cluster and small galaxy scales of the Universe, both controls and shaped our Universe into how it now appears, meaning that understanding DM is vital in order to know how the Universe evolved and how it will die.

Following the discovery of the first signs that the Universe may not be made up purely of the particles from the "Standard Model" of particle physics and that these make up a small fraction of the Universe's matter content, current estimates state that for every unit mass of baryonic matter, there exist almost five unit masses of DM. DM must, because of the fact that it cannot be seen in any wavelength of light, interact very weakly with electromagnetic forces, and appears only to interact with both itself and baryonic matter via gravitational forces (as implied in the preceding paragraphs). DM could also interact via the "Weak" force if found to be made up of "Weakly Interacting Massive Particles" (WIMP, one of the leading candidates so far), though they would still not interact strongly electromagnetically and remain "dark". Many experiments are under way to try to detect them, for example the Super-Kamiokande that utilises neutrino detection and Imaging Atmospheric Cherenkov Telescopes that utilises Cherenkov light emitted in atmospheric particle showers (Rott, 2012).

That DM only interacts gravitationally with baryonic matter is shown in merging clusters such as the Bullet Cluster (1E 0657-56, see Fig. 1.1 Bradač et al. 2006) and Abell 1758 (Ragozzine et al., 2012), which both have the intra-cluster gas (after DM, the largest contributor of the cluster's mass) decoupled from the galaxies (which can be considered collision-less as they interact so weakly), due to the gas interacting very strongly electromagnetically. The DM in these bullet-like clusters is well modelled by dynamically "cold" DM (Massey et al., 2011), which is travelling at speeds that are non-relativistic. Relativistic DM results in a different DM profile and also affects the formation of the Universe and the resulting structures seen today (Combes et al., 1995). For the bullet-like clusters, the DM was found to lie very near the galaxies and not the intra-cluster gas, which shows that it has a low interaction cross section with itself (Markevitch et al., 2002, galaxies are almost collisionless with one another) and baryonic matter. In

contrast the hot intra-cluster gas interacts strongly with itself. Studies of the Bullet Cluster (Randall et al., 2008) and other galaxy clusters such as MACS J0025.4-1222 (Bradač et al., 2008) put upper limits on the possible self-interaction cross section and found it to be very small (Randall et al. found a value of $\sigma/m < 1.25$ cm^{2}g^{-1}).

DM does not appear to form structures smaller than the galaxy halos observed. Attempts to use microlensing to detect small structure DM have resulted in non-detections (Ackermann et al., 2011 & Griest et al., 2011).

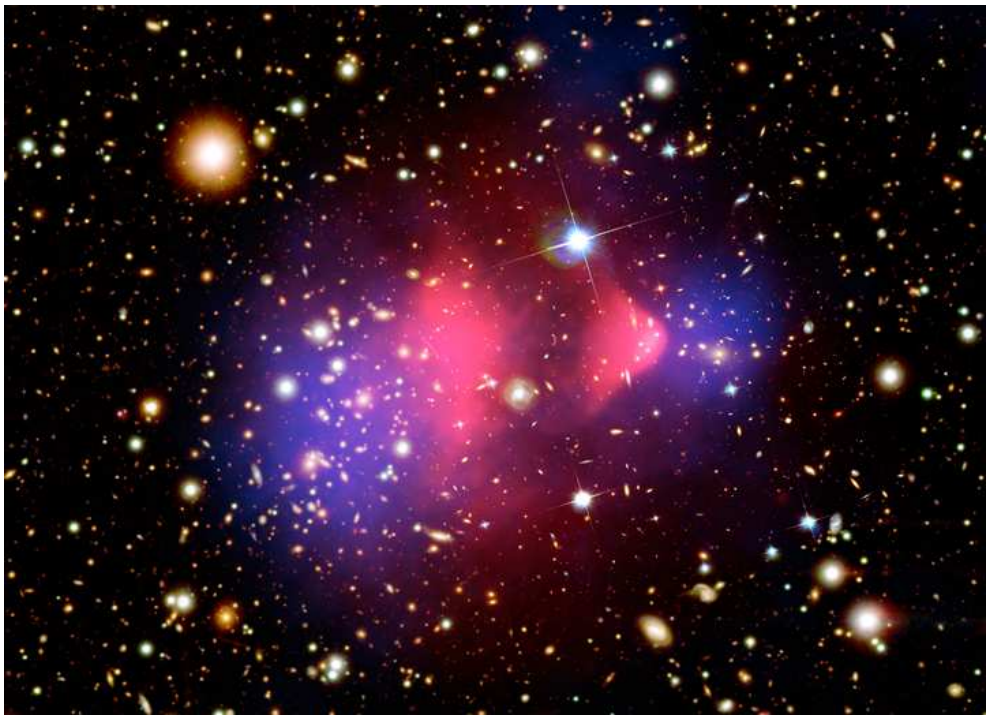


Figure 1.1: 1E 0657−56 (the "Bullet Cluster"). A *HST* colour image of galaxies, with dark matter (blue) and x-ray emitting gas (pink) overlayed, showing the separation of the last two components from one another. Image Credit: X-ray: NASA/CXC/M.Markevitch et al. Optical: NASA/STScI; Magellan/U.Arizona/D.Clowe et al. Lensing Map: NASA/STScI; ESO WFI; Magellan/U.Arizona/D.Clowe et al. Picture taken from: `http://www.nasa.gov/vision/universe/starsgalaxies/dark_matter_proven.html`

What has been learnt so far is that DM performs an important role in the shaping, structure and evolution of the Universe. The mass of DM will have slowed the expansion of the Universe; its low self-interaction rate will have meant it coalesced gravitationally before the baryonic component could do so. This is due to the baryons in the hot early Universe strongly interacting with photons (as a result of ionisation), causing photon

radiation pressure to dominate over gravitational forces (any gravitationally collapsing baryonic matter would be pushed apart by photon interactions). When the baryonic component is cool enough (allowing baryons to form neutral atoms, which do not interact strongly with photons) due to the expansion of the Universe, they begin to accrete under gravity, being drawn to these seeds of DM. Therefore the amount, and also composition ("cold" or "hot" for example), of DM in the Universe has a strong effect on cosmology.

The following[a] summary can be made of the current understanding of DM (to which gravitational lensing measurements have made a large contribution):

(i) The Universe contains about five times more dark matter than baryonic matter.
(ii) Dark matter interacts approximately normally via gravity.
(iii) Dark matter has a very small electroweak and self-interaction cross-section.
(iv) Dark matter is not in the form of dense, planet-sized bodies.
(v) Dark matter is dynamically cold.

[a] - List summary taken from Massey et al. (2010)

Discussing each of these points in turn, it is found that weak lensing enables the probing of galaxies far beyond the area that the visible mass component traces (the visible mass component, such as the velocity rotation curve also provides an indicator of the mass enclosed at key radii). When analysing the stacked lensing signal around galaxies in SDSS survey imaging, it is possible to find a typical total mass from the lensing signal and compare it to a measurement of the stellar mass component and independent radio observations of the gas component (Massey et al., 2010). Combined, this matches the predictions of five times more DM than baryonic matter in the Universe. The fact that DM generates lensing effects similar and in line with that observed by baryonically generated lensing signals is a strong indicator of DM interacting gravitationally in a similar fashion. The fact that in merging galaxy cluster systems (such as with the strong lensing analysis of the Bullet Cluster), the DM component has been decoupled from the strongly self-interacting intra-cluster gas and instead lies more coincident to the (essentially collisionless) galaxies, provides evidence of the low self-interaction cross section (and small

electroweak interaction if formed of WIMPs) of DM itself. Microlensing studies of the Milky Way and Magellanic Cloud have been performed that searched for events caused not only by luminous components of the galaxies, but possibly also by Massive Compact Halo Objects (MACHO) formed of either baryonic or DM. So far evidence of lensing events caused by small mass objects of DM are very low (events are instead dominated by baryonic objects as their cause) and only upper limits calculated for the fraction of the DM halo comprising these MACHOs (Calchi Novati et al., 2009). 'Instead, the results indicate that DM is predominantly in the form of much larger and more massive objects than planetary sized ones. The *WMAP* (Wilkinson Microwave Anisotropy Probe) analysis of the cosmic microwave background over three years produced strong constraints on cosmological parameters and found that cold DM was the best fit in order to reproduce the structures seen today (Spergel et al., 2007). Hot (relativistic) DM produces a different large scale structure (if seen today) than cold DM, due to its ability to free stream away from any areas of increasing clumping of matter in the early Universe. This retards the growth and formation of structure in the Universe, over that of cold DM. This also affects the numbers of strong lenses expected to be observed as the structure of the Universe between the two differing types of DM is different. Strong lensing is sensitive to the underlying cosmology of the Universe as a result of this (Turner, 1990).

1.2 Λ-CDM Cosmology

This is the most accepted theory for describing why the Universe is like it is. It is also one of the simplest, as it is based on the fewest parameters of other competing theories. It also fits very well with observations. A good example is the *WMAP* survey, which was an all sky survey of the CMB (cosmic microwave background). The fit (to a flat Λ-CDM cosmology) to seven years of data had very small errors, as analysed by Larson et al. (2011). Their results showed that the Universe was within a very small error "flat" (all of the local space and geometry had no intrinsic curvature and could be described by

Euclidean geometry) and the age of the Universe was constrained to 13.75 ± 0.13 Gyr. Studying the CMB with *COBE* (COsmic Background Explorer) and comparing it to a black body has shown that its variation is only $\left\langle \frac{\Delta T}{T} \right\rangle \approx 10^{-5}$, determining it to be very smooth (Bennett et al., 1993).

Λ-CDM relies on the Universe being "homogeneous" and "isotropic". Homogeneous in that the Earth is not located in a special place in the Universe, and isotropic in that the Universe looks the same in all directions. It explains the Universe well down to small scales (groups and clusters of galaxies), though on smaller scales (smaller than galaxies for example) where the Universe does not look homogeneous and isotropic, comparisons with Λ-CDM are difficult due to the resolution of current large scale N-body simulations (Springel et al., 2005 had a mass resolution of approximately 10^9 $M_\odot$ per particle).

The Λ in Λ-CDM is derived from the existence of dark energy (which makes up more than three quarters of the Universe). The Λ comes historically from Einstein's theory of general relativity (Einstein, 1915), in which he introduced a "cosmological constant" so that his equations would yield a static Universe. Hubble (1929) showed that the Universe was not static and was in fact expanding; Einstein lost the opportunity to realize his equations predicted this, and this was described as his "biggest blunder" (in an article written by George Gamov in 1980 from "My world line", Gamow 1970).

The relative abundances of DM, baryonic matter and dark energy, will both shape and determine the future of the Universe. Taking these abundances as densities and dividing them by the critical density ($\rho_{crit} = 3H_0^2/8\pi G$) of what is required for the Universe to be "flat" will yield Ω_m (the sum of baryonic and DM) and Ω_Λ (dark energy), and are summed together. Then the curvature (described as $K \propto (\Omega_m + \Omega_\Lambda - 1)$) of the Universe can be described by whether K is < 0 (open), ~ 0 (flat, with $\Omega_m + \Omega_\Lambda = 1$) or > 0 (closed). In a closed Universe, the geometry of the Universe is like a sphere and, if the dark energy content is low enough, the expansion of the Universe will eventually slow and then begin contracting under gravity (causing the Big Crunch). In a flat Universe (the geometry of the Universe is flat to an infinite extent), it just expands forever. Finally in an

open Universe (the geometry of the Universe is negatively curved), it will again continue to expand. As stated previously, the latest measurements of the CMB, as published in Larson et al. (2011), hold that the Universe is flat to within an error of 0.5%, (with a Λ-CDM cosmology) in order to fit the results obtained.

With this in mind, current understanding of structure formation in the Universe is that since DM does not interact electromagnetically it will have started gravitationally contracting before neutral baryonic atoms could form (sometime before the end of the timetable set out in Table 1.1); DM would therefore have provided the nucleating seed for baryonic matter to collapse onto (if this was not the case the Universe would not be as far evolved as it is currently observed). The reasoning for this is that the early Universe consisted of a photon-baryon plasma. Only when the Universe had expanded sufficiently that the energy of the photons was low enough to enable neutrons and protons to form followed by neutral atoms (this period is called recombination), off which photons could no longer interact electromagnetically and scatter (this last photon scattering is seen as the Cosmic Microwave Background) and so the baryons became decoupled from the photons (Combes et al., 1995). At this time baryons could begin to collapse under gravity, this occurred at points of over density caused by small thermal fluctuations in the gas at earlier times. This had been magnified by the expansion of the Universe.

If the universe was dominated by only baryonic matter, gravitational collapse would only occur after neutral atom formation and the shape of the Universe would not be as developed as currently observed (Combes et al., 1995). The Universe is observed to be more developed than a purely baryonic dominated one and since observations show our Universe is DM dominated, implies decoupling and gravitational collapsing of DM was occurring before baryonic matter decoupled from photons. The reason DM (in the form of WIMPs) decoupled before baryonic matter was due to their very weak interactions (only via gravity and the weak force), which meant they only stayed thermally coupled to the photo-baryonic plasma due to elastic scattering with relativistic particles. When the density of these relativistic particles fell below a certain threshold (dependent on the

exact nature of the DM particle) the DM particles would have kinetically decoupled from the baryonic-photon plasma. This kinetic decoupling would have occured at much earlier times than the baryonic decoupling, due to the latter's requirements that the photon energy falls below levels enabling neutral atoms, which do not interact so strongly with the photons electromagnetically any more (Combes et al., 1995 and Bertschinger, 2006).

These structures would have resisted the expansion of the Universe more than other areas. On local scales these would have been stronger gravitationally than the expansion forces; this would have caused larger massed objects to accrete and merge with smaller massed objects. This process is known as "hierarchical" merging and thus developed into stars, then galaxies and then into grouping of gravitationally bound galaxies and eventually lead to the development of clusters (Combes et al., 1995). This is the reverse of relativistic (hot) DM, which would actually require the formation of some of the largest structures in the Universe, which then fracture into the smallest structures that are seen today in a "top-down" scenario (Combes et al., 1995). This "hierarchical" merging would have led to the development of filamentary structure, with the largest and most massive objects forming at the junctures of these filaments as mass is funneled down the filaments. This is seen in the Millennium Simulation by Springel et al. (2005), where their modelling of the Universe leads to filamentary structure and the development of galaxy clusters in the junctions of filaments as seen in Fig.1.2. This is a bottom up approach to the Universe's development (basic structures appear first, followed by more complex ones).

1.3 Galaxy Clusters

As stated in § 1.2 the current model of formation in the Universe (based on Λ-CDM cosmology) is that small structures accrete and merge with smaller structures and become larger via "hierarchical" merging (Combes et al., 1995 and Kravtsov & Borgani, 2012). In this way galaxy clusters form by merging with smaller clusters or groups of galaxies and their associated dark matter halos. There is not a distinct divide between groups

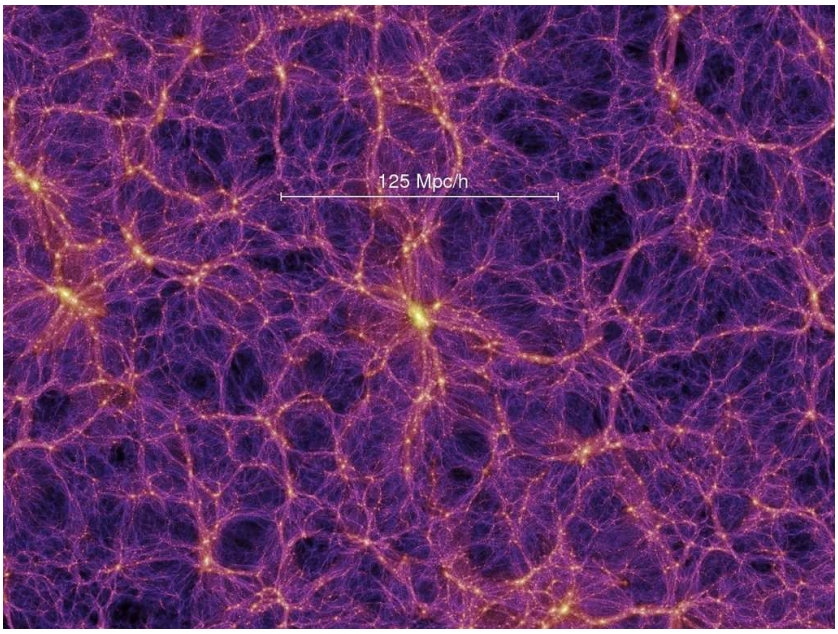


Figure 1.2: Millenium simulation showing the filamentary structure formation in the Universe, with brighter areas signifying larger congregations of mass. Picture taken from "The Millenium Simulation Project" at: `http://www.mpa-garching.mpg.de/galform/virgo/millennium/`

(tens of galaxies held together gravitationally by dark matter) and clusters (hundreds to thousands of member galaxies), but clusters are some of the most massive collapsed objects in the Universe (clusters tend to be around $\sim 10^{14} - 10^{15}$ $M_\odot$) to date and are in fact a very recent occurrence in the history of the Universe. As a result these would only be expected to be observed at low redshift, with very few observed beyond a redshift of one (Jee et al., 2009). The most distant confirmed galaxy was at a redshift of $z \sim 1.6$ (Santos et al., 2011) with the most distant candidate galaxy cluster at $z \sim 2.2$, although it may be between the stages of a proto-cluster and a galaxy cluster (Spitler et al., 2012).

Objects weighing more than 10^{14} $M_\odot$ are also measured to contain $\sim 10\%$ of all the mass in the Universe (Massey et al., 2010), because they exist in the tail of the mass function $N(M, z)$ (i.e. the number of clusters of each mass expected for a given redshift); differing cosmologies affect this distribution as can be seen in Fig. 1.3. This makes galaxy clusters very rare. The fact they appear in this tail makes them interesting to study cosmologically, as any model of galaxy clusters has to reproduce the steepness of this tail or would result in an over or under abundance of clusters compared to what is actually

observed. This tail of the function can probe dark energy via the growth of structure over changing redshift.

Because of a galaxy cluster's age and mass it can contain a record of everything they have merged or accreted with over their lifetime. Going back to the earliest times of their existence, nothing has escaped the gravitational well of the large DM halo that holds clusters together. As such, they have some of the oldest galaxies contained within them. This makes them useful in the study of these oldest galaxies and also makes clusters easier to detect. Plotting the colours of these galaxies against those of the surrounding non-cluster galaxies produces a prominent line of red galaxies (the galaxies also tend to be some of the brightest). Clusters also generally contain a large galaxy (usually located at the bottom of the cluster gravitational potential well), that tends to be significantly brighter than any other member galaxies; this is called the "Brightest Cluster Galaxy" (BCG). The current theory on the development of BCGs is that, as they live at the centre of the potential, they merge with smaller galaxies in the cluster, becoming the dominant galaxy (Smith et al., 2010) through cannibalism of infalling galaxies (Ostriker & Hausman, 1977 & Rasmussen et al., 2010).

Cluster galaxies are predominantly ellipticals because of the small volume and high number of galaxies they live with, when compared to the Universe as a whole. As such they are put under immense gravitational stress and the effects of mergers; this disrupts star formation in the galaxies, but also means that spirals tend to exist on the periphery of the cluster where they are less likely to be disrupted and evolve into ellipticals. Gravitational forces also tidally strip cluster galaxies of their DM halos (Natarajan et al., 2002) and also of their gas. These two components accrete to the bottom of the potential well, with the DM losing energy to the gas to enable it to be compressed and concentrated (Barkana & Loeb, 2010). Significant DM substructure or offsets of this "intra-cluster" gas indicates that the cluster may not be settled and may be suffering the effects of a merger with another object, smaller galaxy cluster or group of galaxies. The intra-cluster gas is shock heated to temperatures of ~ 10 keV and higher as the cluster forms, making them very

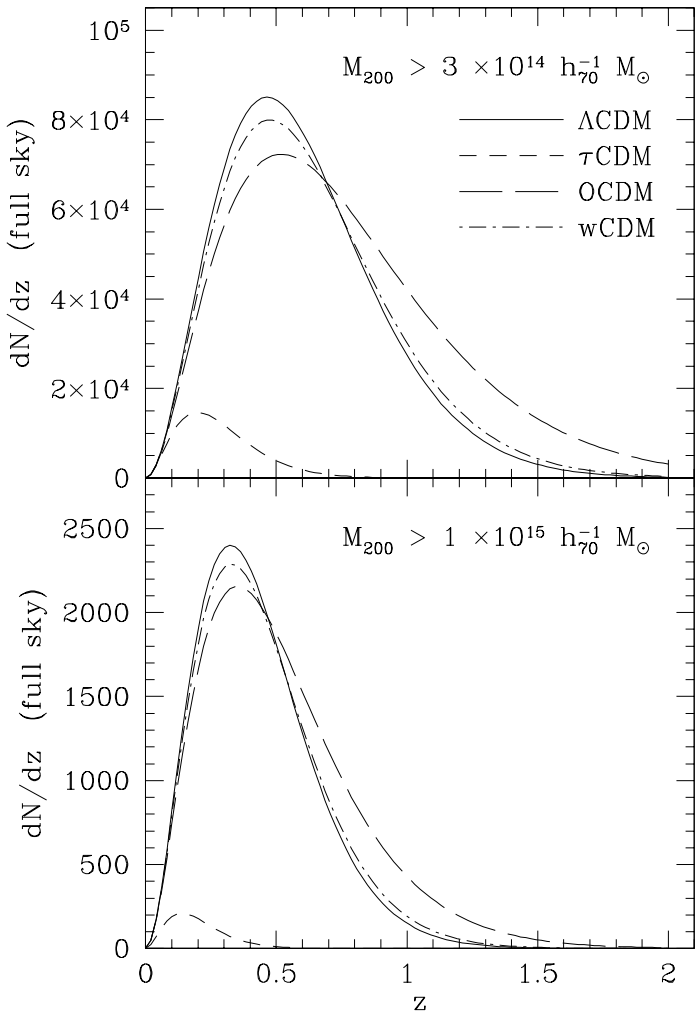


Figure 1.3: Plot showing the differing shape of $N(M, z)$ (predicted number of clusters on the sky with changing redshift) for differing cosmologies and the resulting change between the slopes of the tail of each function. Figure taken from Voit (2005).

X-ray luminous; this aids their observation with X-ray telescopes.

This unique environment can provide interesting insights into the interaction between galaxies, intra-cluster gas and DM, but will also contain a record of its evolution. 3% of the general make up of the mass of a cluster is in the form of the stars (estimates can be made by looking at the spectrum of emitted light against theoretical models, Massey et al., 2010), that make up the galaxies, 13% is in the hot intra-cluster gas and the rest is DM (X-ray luminosity changes with the gas density and so can produce estimates of the total amount of gas present, Peterson & Fabian, 2006; additionally radio telescopes can be used to measure the amount of atomic hydrogen, Read & Trentham, 2005). Observations of the cores of clusters can enable their masses to be determined. This is vital for studying elements of cosmology (Voit, 2005), such as:

(i) Measuring the mass function and its evolution with redshift.

(ii) Studying if mass follows light. Is a DM halo seen where the concentrations of luminosity are highest?

(iii) Studying brightest cluster galaxies (BCGs) and their roles in clusters. Does every cluster have a BCG? How and when do BCGs form?

(iv) Studying the interaction of these three components in mergers.

(v) What is the extent and density slope of DM? Is it axi-symmetric or tri-axial? Does this match numerical body simulations?

(vi) Looking at the cluster observables (see following sections), what is the relationship between the following quantities?

a) Mass and luminosity

b) Mass and X-ray temperature

c) Mass and cluster richness

d) Galaxy velocities and cluster mass

Some of these points had more in-depth explanations in previous paragraphs, but their overall importance is that they are observable quantities of the Universe, which current theories (and also numerical models) of Universe evolution or composition must be able to match or confirm predicted relations between them. The more observable quantities there are for comparison, the better the constraints that can be put on the possible explanations of the Universe. Using this, measurements of the mass and internal structure of clusters is capable of providing good constraints on cosmology (is structure as developed as predicted for example) and relations between other cluster observables. The next two sections outline the methods of viewing cluster observables, so that both global and structural measurements can be taken.

1.4 Cluster Observables - Global

The following section is based upon Voit (2005) and the equations are taken from this source.

1.4.1 Optical

Luminosity

This was one of the first ways that galaxy clusters were seen, with both the Coma and Virgo Clusters being identified by observing the clustering of the bright galaxies (Herschel and Messier both identified these two clusters in the 18^{th} Century). This has evolved to the advanced state where the luminosity of the galaxies in a cluster is measured in several optical wavelength bands and these values hold important information from the clusters. By plotting the colours of galaxies, the member galaxies form a distinct "red sequence". These are the bright, old red galaxies that make up the cluster; studying these can give insights into the early Universe. As these were some of the earliest structures to form, following hierarchical merging, they will contain some of the earliest structures in our

Universe (Voit, 2005). Additionally, by constructing a census of cluster luminosities, it is then possible to examine how it changes with redshift. This holds important cosmology information as different cosmological theories must be able to reconstruct this function (Harsono & De Propris, 2009). It is impractical to measure every cluster galaxy's luminosity. Since they can extend to hundreds of arcseconds from the cluster centre this would require large amounts of telescope observation time in order to cover such a large area, this is not possible for every cluster (every doubling of the radius of observations from the cluster centre, results in a factor four increase in the area needing to be observed for example). But the core of the cluster exhibits the brightest areas of the luminosity function and so the luminosity within a set area is measured, which can be used as a proxy for mass.

Richness

Richness is a cluster member count. It was first used to identify potential clusters by Abell (1958) and further refined in Abell et al. (1989). They found the 3^{rd} brightest cluster galaxy and then measured any galaxies that were up to two magnitudes fainter than this in a fixed aperture. The 3^{rd} brightest was chosen to reduce count errors caused by confusion between the brightest central galaxy and field galaxies (Abell, 1958). These were calibrated by galaxy counts in cluster-free regions of space; the resulting cluster galaxies that met a minimum of 30 galaxies were included in his catalogues as clusters (Abell et al., 1989) - the numbers define the richness. The initial richness classification started at 30 galaxies or a richness class of zero, but some of the clusters numbered more than 300 galaxies and had the highest richness of five. The richness class of a cluster, and therefore how many galaxies it contains, will be a proxy for mass (Rykoff et al., 2012).

Galaxy Velocities

In keeping with virial theorem (where twice the kinetic energy balances the potential energy of the cluster), the galaxy velocities will be related to the amount of mass in

the cluster. To keep a galaxy travelling at a certain velocity inside the cluster's grip, a certain amount of cluster mass is needed (first measured for clusters by Zwicky 1933). To first constrain the cluster members, spectroscopic measurements are required because photometric redshifts are not sufficient to do this. By then examining the distribution of redshifts in redshift space and for a relaxed cluster, the velocity dispersions will approximate a Gaussian distribution around the cluster's redshift (statistical tests have shown this to be a good fit, Yahil & Vidal, 1977). Selecting members can then be as simple as fitting a Gaussian distribution and excluding any redshifts outside a suitable sigma cut off (Yahil & Vidal, 1977). With sufficient redshifts it should be possible to perform the mass estimations at several radii from the cluster centre, providing an estimation of the cluster's mass.

The fit of the Gaussian will also yield a standard deviation value for the cluster velocity dispersion along the line of sight. This can also be estimated via bootstrapping, where N synthetic samples are constructed by randomly re-sampling (with replacement) from the actual galaxy members' distribution (as determined by fitting a Gaussian as defined above), the resulting N samples will give an estimation of the distribution. From virial theorem the average kinetic energy of the cluster will be twice that of the average gravitational potential energy, yielding the equation:

$$2KE = PE \tag{1.1}$$

Where

$$KE = \frac{3M\sigma_{1D}^2}{2} \tag{1.2}$$

and

$$PE = \frac{GM^2}{r} \tag{1.3}$$

Where σ_{1D} is the line of sight velocity dispersion, M is the cluster mass contained

within a radius r and G is the gravitational constant. Substituting in the equations for Kinetic Energy and Potential energy gives[1]:

$$M = \frac{3\sigma_{1D}^2 r}{G} \tag{1.4}$$

This links the relationship between the cluster's mass and the velocity dispersion of observed cluster galaxies, providing an estimation of the clusters mass from observed quantities. This was one of the first indications of the need for dark matter (Zwicky, 1937). It should be noted that equations work under the assumption of a spherical cluster that is relaxed, which is not always the case. In a disturbed cluster or one where the redshifts are poorly sampled, fitting a Gaussian distribution to constrain cluster member galaxies may not be possible and the resulting mass estimation will also be incorrect. An incorrect cut in redshift space will affect the resulting velocity dispersion calculation by outlier contamination. So a non-Gaussian velocity dispersion gives a possible indication of a cluster being disturbed.

1.4.2 X-ray

X-rays are emitted from clusters as a result of compression or shock heat of the intra-cluster gas. The reason gas large amounts of gas is adrift in the Universe is due to galaxy formation being inefficient, with only 10% of the total Universe baryons residing in the stars of galaxies (Voit, 2005). The gas itself is only visible because it is inside the cluster's potential well, otherwise it tends to be difficult to observe when lying outside of gravitational potential wells.

Luminosity

X-ray emission from clusters results from thermal Bremsstrahlung radiation (from observations of the Coma Cluster by Felten et al., 1966); by looking at the emission rate

[1]Equations sourced from Voit (2005)

of energy, a cooling function can be calculated as a luminosity per unit volume. From an assumption of hydrostatic equilibrium, this can be used to create an X-ray surface brightness model called a "beta" model (Cavaliere & Fusco-Femiano, 1976). It is named β from its relation to the density of the gas. The surface brightness profiles of clusters show that the majority of the X-ray luminosity emanates from a relatively small fraction of the intra-cluster gas (that near the cluster core), meaning they are very bright and concentrated on the sky. This makes using X-ray luminosity a good method of picking clusters out due their deep potential wells (from their large masses), that compresses the intra-cluster gas to high temperatures, making them potentially much more X-ray luminous than other extra galactic objects.

Temperature

The temperature of the intra-cluster gas should correlate to the mass that is constraining it by hydrostatic equilibrium (see end of subsection). The number of photons recorded, as well as the model used to produce a 3D mass from the projected data, will determine the accuracy of the mass measurement. The more photons observed the smaller the binning that can be applied, and this will affect the accuracy that a temperature gradient can be measured. Without temperature measurements, a scaling relation between luminosity and temperature can be used but is undesirable as two scaling relations are in use (the other is the relation between mass and temperature).

A simple scaling relation between temperature of intra-cluster gas and cluster mass can be constructed starting from the virial theorem. Taking equation 1.4 from §1.4.1 shows $M \propto \sigma_{1D}^2 r$. The velocity dispersion of the galaxies is related to the temperature of the gas with $kT_X \propto \sigma_{1D}^2$ (where k is Boltzmanns constant) via their specific energies matching. Combined, this yields a relation between the temperature of the intra-cluster gas and the cluster mass via $M \propto T_X r$. Despite the problems with dealing with systematic errors, the X-ray temperatures are found to correlate well with velocity dispersion (Xue & Wu, 2000) and therefore mass.

1.4.3 Millimetre

The Sunyaev-Zel'dovich effect (SZE) (see Sunyaev & Zeldovich, 1970 for a full derivation) provides information on density perturbations of the Universe and some very dense galaxy clusters have been discovered using this. Thermal SZE is the distortion on the radiation from the Big Bang (this radiation is the CMB as discussed in § 1.2), caused by its passage through ionised gas (inverse Compton scattering). One of the first measurements of this was carried out by Birkinshaw et al. (1991), who measured the SZE effect in Abell 665 and compared it with its X-ray properties. The CMB has very little disturbance in it (with a temperature of 2.72548 ± 0.00057 Kelvin it is exceptionally smooth (Fixsen, 2009)), so these effects are very small, but detectable using specialized and cooled telescopes. Examples of the instruments used so far for SZE measurements are the South Pole Telescope (SPT, a 10m telescope in Antarctica), Atacama Cosmology Telescope (ACT, a 6m telescope located on Cerro Toco, Chile) and the Planck spacecraft. The large concentration of intra-cluster gas in a cluster means they have a "larger" ($<$ 1mK, Carlstrom et al., 2002) effect on the CMB than other sources, as they provide a larger volume for CMB photons to scatter through (still a photon only has approximately a 1% chance to interact with an ICM electron, Carlstrom et al., 2002). A common value cited and calculated in SZE research is the "Compton-Y" parameter (y in equation 1.5). The following equations were taken from Carlstrom et al. (2002):

$$\frac{\Delta T_{SZE}}{T_{CMB}} = f(x)y = f(x) \int n_e \frac{k_b T_e}{m_e c^2} \sigma_T dl \tag{1.5}$$

Where ΔT_{SZE} is the temperature change caused by SZE, n_e the electron number density, σ_T the Thomson cross-section, $m_e c^2$ the electron rest mass energy, T_{CMB} is the CMB temperature and y is the Compton-Y parameter. The frequency dependency of SZE is $f(x)$:

$$f(x) = \left(x \frac{e^x + 1}{e^x - 1} - 4 \right) (1 + \delta_{SZE}(x, T_e)) \tag{1.6}$$

Where δ_{SZE} is a relativistic correction to the frequency dependence and T_e is the electron temperature. Equation 1.5 expresses the SZE distortion on the CMB as a temperature change at a dimensionless frequency x. With x being defined as:

$$x \equiv \frac{h\nu}{k_B T_{CMB}} \tag{1.7}$$

Where h is the Planck constant, k_B is the Boltzmann constant and T_{CMB} the CMB temperature. For an isothermal sphere, Compton-Y can be considered as the optical depth. The other values are σ_T for the energy gain per inverse Compton scattering, n_e the electron density in the ICM and T_e the electron temperature in the ICM. A very useful characteristic of this equation is that it is distance (redshift) independent, allowing its use to high redshifts.

There is another method of effecting the CMB, the "Kinematic SZE" effect, which is the energy gained by the CMB by the bulk motion of electrons in the gas. An example is if the intra-cluster gas from a cluster is moving along the line of sight (or a component of it) at a velocity ν_{los}, the bulk motion of the gas will impart additional distortion to the CMB (from the inverse Compton scatterings of the CMB photons from the gas) due to the Doppler effect. This distortion is shown in the following equation taken from (Carlstrom et al., 2002):

$$\frac{\Delta T_{SZE}}{T_{CMB}} = -\tau_e \left(\frac{\nu_{los}}{c}\right) \tag{1.8}$$

Where τ_e is the optical depth (for an isothermal cluster this is equal to y in equation 1.5). This results in a shift of the CMB spectrum in temperature, with the direction of the temperature shift dependent on the peculiar velocities of the gas (Carlstrom et al., 2002).

The disadvantage to SZE is that it does not directly measure the total mass of the cluster, but rather inverse comptonization from the gas the cluster contains and, because of the low temperature values that the cosmic microwave background has, requires very specialised and cooled telescopes used for long observation times in order to measure it.

1.5 Cluster Observables - Structural

The previous sections considered globally derived values for clusters and the next will consider what these areas can provide in the area of structural information. The following section is based upon Voit (2005). The equations are taken from this source.

1.5.1 Optical

BCG Dominance

Examination of the dominance of the BCG in the cluster should give indications of the dynamical status of the cluster, as this arises through the accretion of other galaxies. This is seen in fossil group simulations, which suggest the merging of close galaxies to form one much larger elliptical galaxy. This will be a lot larger and brighter and still possess the DM halo (Ponman et al., 1994). BCG formations is still an active area of study, with simulations attempting to match observations (Martizzi et al., 2012).

Studying the BCG may also give insights into the structure and history of the cluster as a whole. Smith et al. (2010) looked at the luminosity gap Δm_{12}. This is the magnitude difference between the brightest cluster galaxy and the next brightest one. They found that for low values, the morphology of the BCG covered the range from spiral to elliptical and the clusters themselves could possess high substructure fractions. Larger values gave a more general elliptical shape and the clusters generally had low substructure fractions. This indicates a possible link between recent merger history because, as the cluster settles, the BCG will absorb material, becoming larger and more luminous and increasing the luminosity gap. Smith et al. (2010) also point to the inability of current cluster merging models to reproduce the range of Δm_{12} that the authors observed, citing it as an important cosmological model tester.

Merger / Non-Merger

Bautz & Morgan (1970) developed a classification system (named "Bautz-Morgan", BM) which is based on the relative degree to which the brightest member galaxy (BCG) stands out against the rest of the cluster population. The classification runs from "I" (distinctly brighter galaxy that is centrally located) to "III" (no dominant galaxies can be picked from the cluster galaxy population) and can be used to provide information as to whether the cluster is regular, intermediate or irregular morphologically. However there will be some overlap (Bahcall, 1977).

Dressler Shectman

The redshifts of galaxy cluster members from spectroscopic observations can also be combined with their positions on the sky. Dressler & Shectman (1988) used this information to produce a Dressler Shectman (DS) test, which is discussed in more detail and examples of resulting plots given in chapter 3. The tests produce a plot that shows possible substructure location and size as a plotted circle for each galaxy. The circle sizes relate to the local galaxy velocity dispersion mean and standard deviation as it relates to the global total sample values. This visual depiction of structure is very useful, but can be misleading if the cluster member numbers are low. This is why a DS statistical test can also be made to yield a numerical value of the level of significant substructure within the cluster.

It should be noted that in this method redshifts are needed in order for strong gravitational lensing analysis to be performed. As such this can be a useful extra tool to examine the dynamics of the clusters and give information when performing gravitational lensing mass modelling, as settling on the priors can be a laborious and time consuming process in order to achieve a good fit to observations. A more rigorous explanation and methodology for DS analysis is given in § 3.3.

1.5.2 X-ray

Merger / Non-Merger

X-ray luminosity can constrain the X-ray centroid of emission; where this is situated in the cluster in relation to the BCG position can give insight into whether the cluster is under a merger or not. If the cluster is disrupted, displacement from the cluster potential well is expected, especially for gas which self-interacts strongly, while the galaxies interact very little with one another.

It is also possible to use the "power ratio" as a method of measuring the dynamical status of clusters. This ratio is computed from a multi-pole expansion of the two dimensional potential well of the cluster (Tsai & Buote, 1996). These ratios are sensitive to substructure in clusters, with a bimodal cluster having a lower ratio than just an ellipsoidal one. They can, therefore, provide an indication of whether a cluster is relaxed (virialized) or not (Buote & Tsai, 1996).

Cool Core / Non-Cool Core

A calculation for the time taken for the emission area to cool can be made by examining the energy emitted in a cluster. For many clusters this is much shorter than the lifetime of the cluster (Peterson & Fabian, 2006). This suggests that there is infall of material to maintain this cooling core and that the cluster is relaxed in order for this infall to occur (mergers would disrupt the cooling flow for example). Clusters showing this effect are called "Cool Cores" (CC), those that do not are "Non-Cool Cores" (NCC). The ability to measure a CC in a cluster is a good indicator of the structure in that cluster. It also has interesting cosmology effects, with both Burns et al. (2008) and Poole et al. (2008) presenting theories on the formation and evolution of cool cores in clusters. Burns et al. stated that once cool cores are destroyed by mergers they cannot reform and Poole et al. that if disrupted and destroyed they will settle and reform. This stems from Burns et al. (2008) finding in their statistical simulations that galaxies form as NCC with an embryonic

cool core at their centre and if mass accreted slowly enough onto them this turns into a full CC. If, however, they are disrupted by a merger, it destroys the embryonic CC and set up conditions that prevent it reforming. This is in contrast to Poole et al. (2008), who found that only direct collisions destroy CC, but it does not set up conditions that prevent it reforming, this results in the CC re-establishing itself when the cluster has settled down to a relaxed state again. Evidence of a cluster that appears to be settling after a merger and exhibits a CC would prove interesting as an indicator of Poole et al. being correct.

1.5.3 Millimetre

SZE cannot provide substructure information on its own as its resolution is simply too large (in the order of hundreds of arcseconds), but it is complementary to X-ray measurements. The X-ray measurements can give the integrated line of sight gas density and the temperature. By combining this with the Compton-Y parameter from SZE (assuming a spherical object) yields the cluster's gas density profile and can derive a line-of-sight thickness for the cluster (Carlstrom et al., 2002). For a spherical object, this thickness can provide a distance by comparison with its angular size (Birkinshaw et al., 1991).

1.6 Gravitational lensing

The most basic explanation of the principle of gravitational lensing is to compare it to how an optical lens works, in that a "lens" is placed between the observer and the source to be viewed and the object is magnified. This analogy is not perfect for several reasons. For example, using an optical lens the deflection angle increases with increasing distance from the centre of the lens whilst in a gravitational lens the reverse is true. The shape of the gravitational lens can cause distortions in the images seen on the sky and can even create multiple images of the object (explained further in § 2.1). The major advantage this method has over the others already mentioned is that it directly measures the total mass

within a certain radii, although it is data and constraint dependent and does not rely on any assumptions about the cluster's dynamical state in order to give a mass estimation. The disadvantage is that complex clusters can be difficult to model and there can be a need for data with low atmospheric distortion or seeing (for example, there is a need for low atmospheric distortion of the shapes of background galaxies over a potentially wide area for weak lensing studies) and long enough integration times that distant background objects are discernible in order to make good mass estimations. However, the magnification of background sources by strong lensing can reduce the observation times required. This means that not all telescopes, either by their position on the earth, size, weather conditions or field of view are necessarily suitable for these searches. For example, at sea level they will be more affected by seeing, than telescopes located in dry arid areas at $14,000$ feet. In strong lensing, for example, it is necessary to be able to detect multiple image systems that are sometimes very faint, match them and also measure a redshift. Strong lensing analysis without a multiple image redshift is possible, but increases the parameter estimation errors. The mass estimation made is also a total mass integrated along the line of sight and so projection effects can become an issue if not carefully accounted for.

The main focus of the work in this thesis is on strong lensing but short descriptions of the other two methods are included in the following sections. The descriptions of both microlensing and weak lensing utilise information taken from Massey et al. (2010). Section 1.6.3 provides a short history of the use of strong lensing with galaxy clusters and includes the basic principles of measuring cluster properties with this method.

1.6.1 Microlensing

Using this method, it is possible to detect small scale objects (such as dim white dwarfs or exoplanets) that might be too dim to be detected by normal means. It works by a bright object (such as a star acting as the source) being in direct alignment with the observer and the object to be detected (this latter acts as the lens). This magnifies the source and so the observer sees a brightening of that source. The length of time and the increase of

brightness can give details about the object causing the lensing event. These events tend to be chance observations and most easily observed within our own galaxy (Massey et al., 2010). This method has also been used to try to detect DM in small structures (discussed in § 1.1).

1.6.2 Weak Lensing

This process will not cause any multiple images or obvious distortions (from any natural random shape) to a single background source and so requires a large sample statistical approach to analyse. To find any general distortion from background source shapes in one direction, all the sources in that area must be recorded and analysed to remove the effects of the random orientation of the galaxies' natural shape. An example of weak shear analysis can be seen in Fig. 1.4 (the shear produced is tangential to the potential well). The advantage of this is that it can be carried out to much larger radii from the cluster centre than for strong lensing and also for lower cluster masses, but is not as well constrained near the cluster core. This can also be done for almost any large object if there are sufficient objects that can be statistically combined to retrieve a lensing signal (i.e. it is possible to measure a shear) and does not require strong lensing images; this analysis can be performed with ground-based telescopes if the seeing can be kept small enough that shapes of the galaxies can be measured. A disadvantage of this method is its averaging nature; while it can extend to large radii, the resolution is quite large and so cannot resolve structure at smaller radii inside the cluster core (the resolution starts to approach this value). It will also not work well in the strong lensing regime (where background images are distorted or multiplied). This resolution comes from the assumption that although all galaxies have some intrinsic elliptical shape on the sky (which introduces shape noise in the measurements), when averaged over a sufficiently large sample (≈ 100, Massey et al., 2010) this shape noise is reduced to unity. This requirement for a sufficiently large number of background galaxies limits the resolution (typically around a few square arcminutes, Massey et al., 2010), as it is determined by the density of galaxies whose shape can be

discerned from the observations.

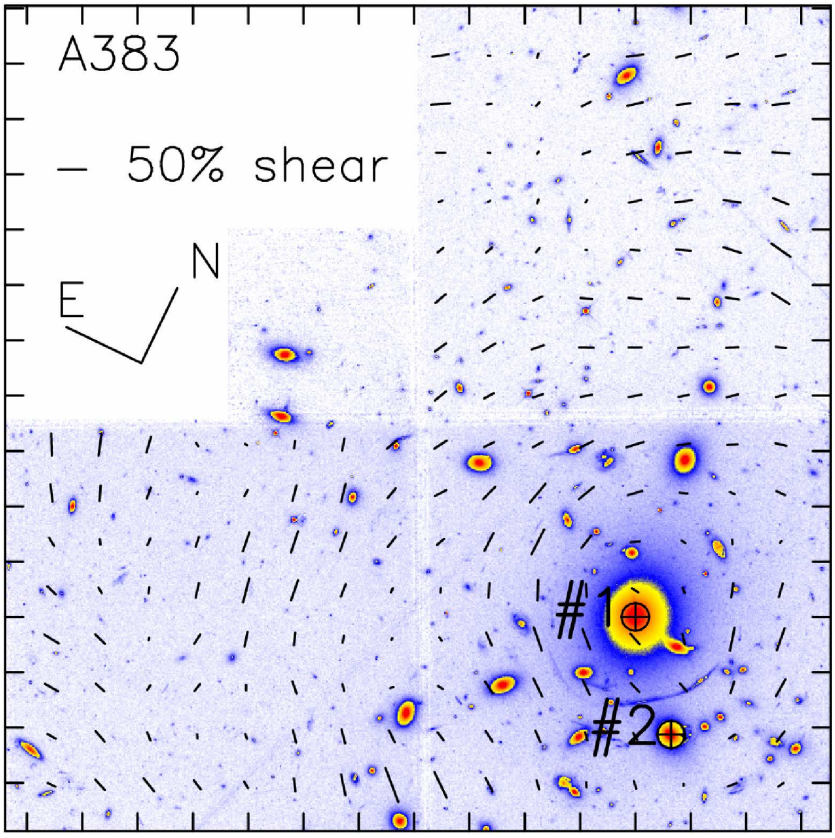


Figure 1.4: An example of weak shear, with the tick marks showing the distortion direction of background sources. Figure of A383 taken from Smith et al. (2005).

1.6.3 A Short History of Gravitational Lensing by Galaxy Clusters

The idea of light being bent by gravity is not a new concept. It has been conjectured for many years, with even Newton having discussed the possibility of light being bent by gravity if it could be treated as a particle. The Newtonian result (both equations taken from Weinberg, 1972) for this deflection was $\alpha = \frac{2GM}{c^2\xi}$, which is a factor of two smaller (this factor was one of the stated reasons for Eddington's 1919 expedition, Eddington 1919) than the result $\alpha = \frac{4GM}{c^2\xi}$ derived from Einsteins theory of General Relativity (Einstein, 1915). Modern gravitational lensing is based on Albert Einstein's Theory of General Relativity (Einstein, 1915). Gravitational lensing was one of the first confirmations of

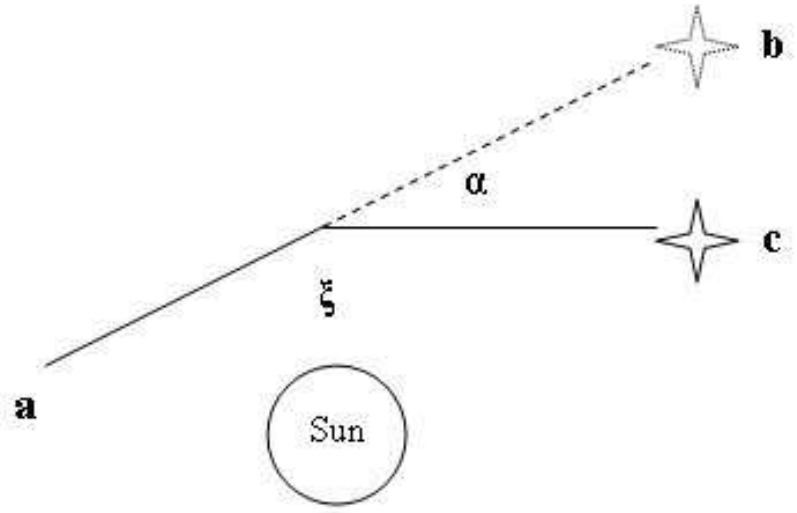


Figure 1.5: A typical simple arrangement of the source (star), lens (sun) and observer "a". "c" shows the actual star position, "b" the position of the star that the observer ("a") sees, ξ is the impact parameter and α is the deflection angle produced by the lens.

his theory, coming from the results of the 1919 Eddington expedition which set out to measure the shift in the position of stars close to the sun during a solar eclipse (Eddington 1919, see Fig. 1.5 where "b" is the observer seen star position and "c" the actual stellar position). This expedition found that the deflection angle matched the predictions of General Relativity and not Newtonian based equations (see equation 2.9) (Eddington et al., 1919).

First Detection from Clusters

Einstein conjectured about Einstein rings (wherein an image is formed into a ring) being possible, but that they would lie within the radius of any star that was being used as a lens to produce them and so he thought they were an effect unlikely to be seen. Fritz Zwicky in Zwicky (1937), later suggested that observing Einstein rings and arcs would be a possibility when using galaxies, galaxy clusters or any high mass objects (or groups of objects) outside our own galaxy as a lens. Some 42 plus years after Zwicky first suggested the possibility, the first gravitational lens was observed in 1979 as a doubly imaged quasar by Walsh et al. (1979). A spectacular example is the four multiple images of the quasar

G2237 + 0305 (Crane et al., 1990), otherwise known as the "Einstein Cross" which can be seen in Fig. 1.6. As such, observational gravitational lensing is still a fairly new and still developing field that has only been around for thirty five years in a practical form.

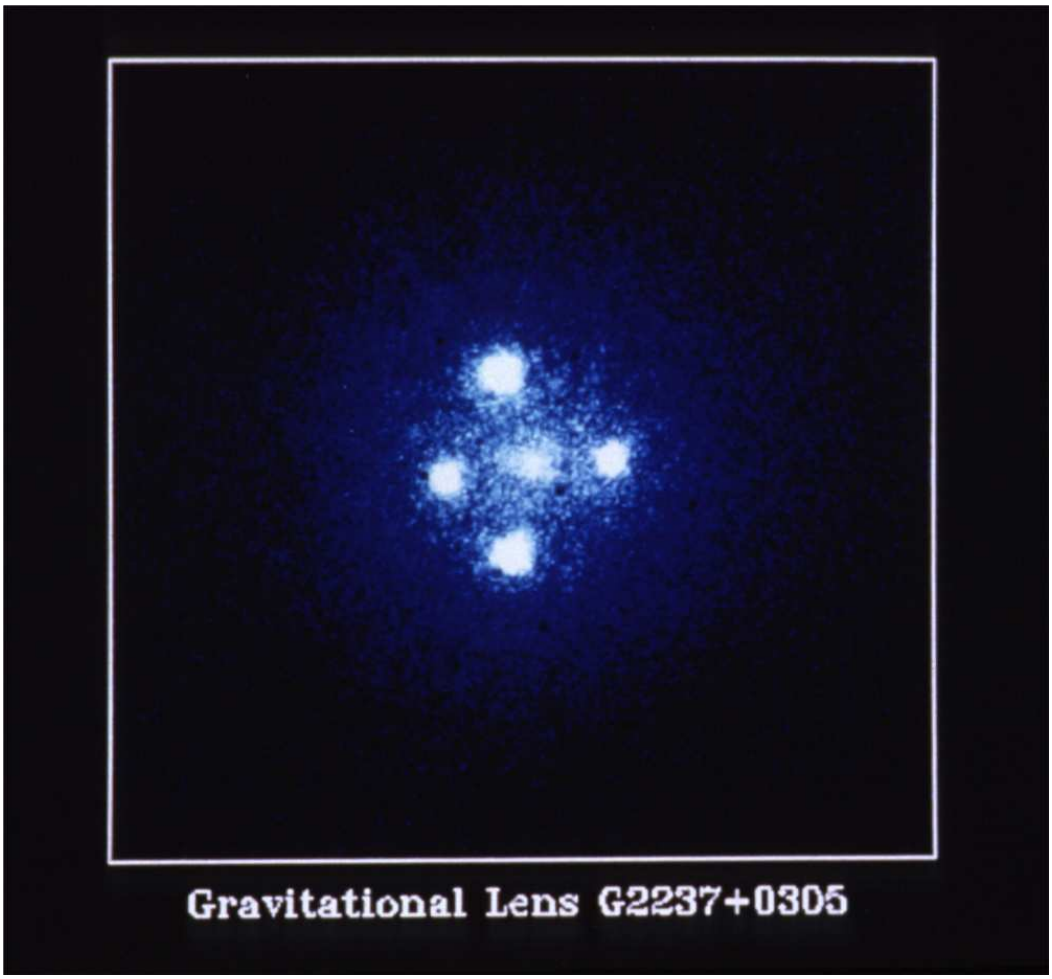


Figure 1.6: G2237 + 0305, otherwise known as the "Einstein Cross", showing the four multiple images of a single quasar. Picture taken from "First ESA Faint Object Camera Science Images the Gravitational Lens G2237 + 0305" at: `http://hubblesite.org/newscenter/archive/releases/1990/20/image/a/`

Einstein Rings

Although Zwicky postulated about the likelihood of seeing arcs or Einstein rings in 1937 (Zwicky, 1937), an actual Einstein ring was not finally observed until September 1985. This was in the galaxy cluster Abell 370 (at a redshift of z=0.373), observed by Soucail et al. (1987a). The ring was a serendipitous discovery as the cluster was originally cho-

sen due to it being very bright in the X-ray region and for its bright blue arc colour. They found the cluster possessed a ring-like structure inside it, which after much detailed analysis to ensure it was not an artefact in the data or caused by some other event (e.g. quasar, gas or star formation processes), was eventually narrowed down to the most likely candidate (due to its spectrum) of being a spiral galaxy at a redshift of z=0.59. This spiral galaxy was being gravitationally lensed by the cluster (Soucail et al., 1987b).

This discovery was quickly followed by another strong lensing discovery made by Lynds and Petrosian from 1976 test data taken with the Kitt Peak National Observatory video camera. Only in 1986 with new observational data were they able to confirm the existence of arcs in three of the clusters they had surveyed (A370, A2218 and CL 2244-02). Observational data had improved by this point, providing more convincing evidence that gravitational lensing events were not just a theoretical possibility (Lynds & Petrosian, 1989).

Mass Comparisons

Galaxy clusters causing gravitational lensing arcs enabled mass estimations of the lensing clusters (see Einstein ring equation 2.12). These were found to have different values to existing X-ray mass estimates. Miralda-Escude & Babul (1995) found, for example, that in two of the clusters analysed (Abell 2218 and Abell 1689), the gravitational lensing derived value was a factor of 2 or more higher than that for the X-ray estimation. This discrepancy seemed to be caused by the lensing estimates having a more concentrated mass distribution around the core than the X-ray estimations. This suggested that one or both estimations were wrong. Various explanations for the discrepancy were put forward, such as the cluster's mass being extended along the line of sight, or one cluster being in front of another (so the gas temperature would be measured for only one cluster, but both clusters would contribute to the lensing estimation). The problem with such explanations is that they either do not compensate enough for the mass discrepancy (only by a reduction factor of 1.04 to 1.15 for these two clusters) or need such a good alignment that they

should be rare (such as for the mass distribution being extended or a superposition of two clusters one on top of the other).

This discrepancy between gravitational lensing and X-ray mass estimates was ultimately explained by the assumptions of virialisation made about the cluster for the X-ray estimations (Allen, 1998). Allen found that for the 13 clusters he analysed, those clusters possessing cooling flows, which should be the most dynamically relaxed, agreed very well with both mass estimates. Clusters with cooling flows have a slow accretion of gas from the outside of the cluster to the inside. Those clusters that did not show an agreement between mass estimations also showed a discrepancy on the mass centres and their BCGs, a significant offset being observed in some cases. This offset implies that the assumption (required for the X-ray mass estimate) that the cluster is relaxed, not disturbed and therefore virialised is false in these cases and can cause a mass discrepancy of a factor of 2 to 4 as a result (Allen, 1998).

The launch of *HST*

While the discrepancy between X-ray and gravitational lensing mass estimates continued, a further technological revolution in gravitational lensing occurred in the form of data from the Hubble Space Telescope (*HST*). Before the advent of the *HST*, an analysis was performed on galaxy cluster Abell 2218 by Kneib et al. (1995). This analysis used only ground-based observations in the optical and near-infrared parts of the spectrum, along with space based X-ray observations. From this they found a bimodal mass distribution, with the two centres coincident on the two brightest galaxies in the cluster. A subsequent paper by Kneib et al. (1996) utilising *HST*'s higher resolution observations on the same cluster, enabled features in the arcs to be broken into their multiple images. Several new images were also found, enabling the refinement of the mass model originally produced by Kneib et al. (1995). They also stated that they found that *HST* data better constrained the cluster mass profile (compared to the older data) and also allowed mass estimations for the cluster members to be made. Fig. 1.7 shows the improvement of the optical imaging

from ground based observations in Kneib et al. (1995) to *HST* in Kneib et al. (1996).

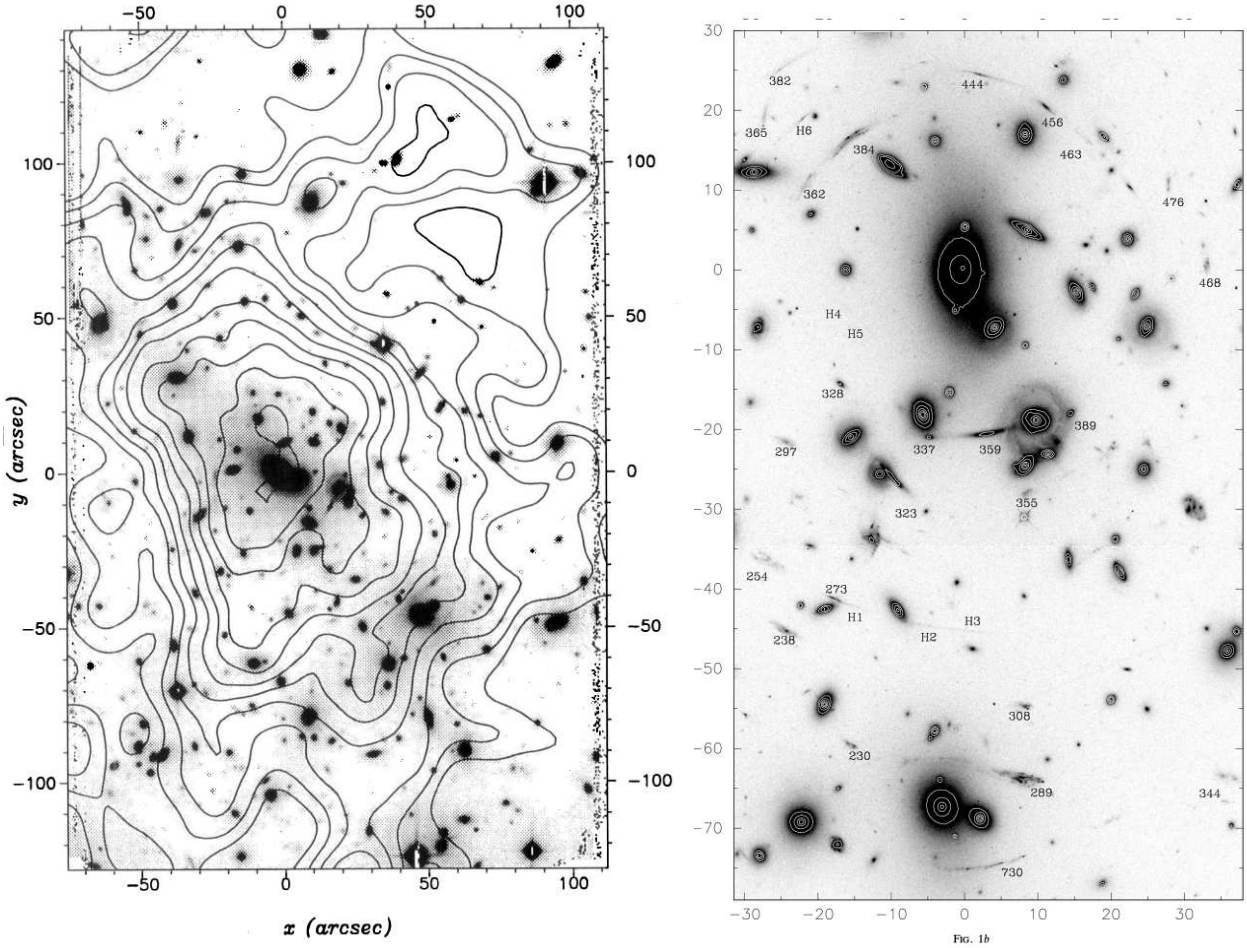


Figure 1.7: Left - Figure 7 from Kneib et al. (1995), showing r band imaging with X-ray contour overlay. Right - Figure 1b from Kneib et al. (1996) of the *HST* imaging of A 2218 (zoomed and rotated compared to the left). The improvement in optical imaging is apparent between the two frames.

Gravitational Telescopes

With the *HST's* deeper and higher resolution images, it became a lot easier to exploit the magnification aspect (see §2.2.2) of gravitational lensing in order to observe high redshift objects. These objects would otherwise be very hard to observe with current technology. Examples of this deep imaging include a serendipitous discovery of a high redshift (z=4.92) galaxy that was being lensed by the cluster CL 1358+62 made by Franx et al. (1997). They also attempted to reconstruct the original source image by inverting

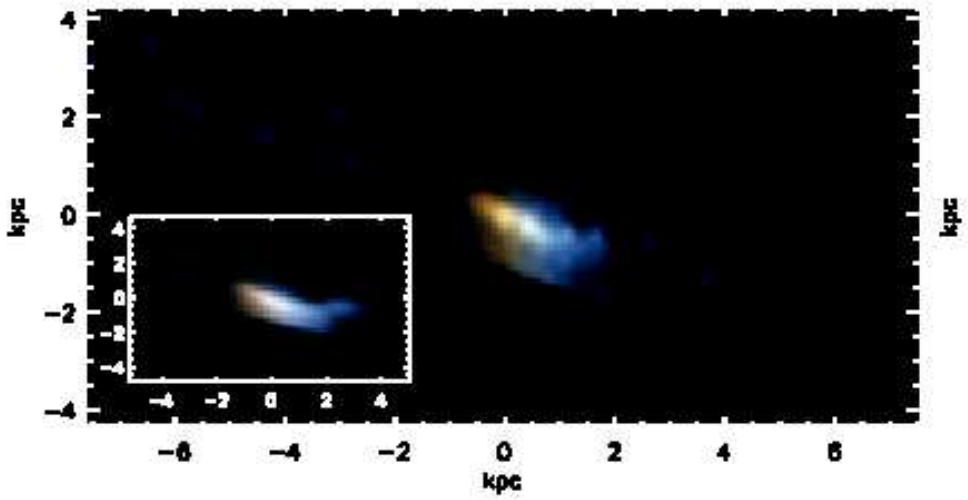


Figure 1.8: A true colour reconstruction of a source plane galaxy if observed unlensed by *HST*. The inset image is the same reconstruction after a 0.8 arcsec smoothing has been applied (taken from Swinbank et al., 2007)

the gravitational lensing model they made for the cluster.

Another example of a "gravitational telescope" is the deep survey of a Lyman-break galaxy at $z \sim 5$ by Swinbank et al. (2007), in which spatial features and details of the source galaxy could be determined. This is important, as the work revolved around the issue of early galaxy formation and whether they possess super winds (some of them appear to, so do they all?), which leads to them losing mass early in their lives. This information is important when investigating how galaxies form and evolve, and for producing simulations of them. Observing a galaxy without the lensing would lead to very poor spatial resolution and an extremely faint magnitude. In this case the object was amplified by a factor of 16, enabling spatial details to be reconstructed when the galaxy cluster gravitational lensing model is inverted (see Fig. 1.8).

The observation of high redshift galaxies with gravitational lensing has its own practical limits, as can be seen in work by Kneib et al. (2004). Here they tried to probe a very high redshift ($6.6 < z < 7.1$) galaxy lensed by the cluster Abell 2218. Their work highlights the difficulty of exploiting gravitational lensing images that lie at such high redshifts. This difficulty is a result of the Lyman-α line for redshifts $z \sim 6.5$ lying in the

region where the quantum efficiency of Charged Coupled Devices (CCD) starts to become very poor. A few other difficulties occur when ground-based spectroscopy becomes very hard at such faint limits (without the magnification in these cases, the source would not be seen) and also when Lyman-α line becomes redshifted into the near infra-red where the sky background is very bright. Bradley et al. (2008) also found a Lyman Break galaxy candidate which, if confirmed, would lie at a redshift of $z \sim 7.6$ and so would put it at the forefront of investigations into the re-ionization of the Universe. In addition they highlight other candidates with redshifts higher than 6 that have also been analysed with the help of gravitational lensing. This indicates that this is a useful technique because although such events are uncommon they are not unique or singular across the whole sky.

Probing the shape of Dark Matter

Gravitational lensing is not just limited to mass estimations and for use as gravitational telescopes, but is able to probe DM itself. Gravitational lensing measures all mass along the line of sight (even DM), giving a total mass estimate. By taking this total mass estimate and combining it with data that only probes the baryonic component mass, it is possible to decouple the DM mass inside cluster cores as demonstrated by Sand et al. (2008). They attempted this by combining gravitational lensing and dynamical data for the clusters Abell 383 and MS2137-23. Although DM dominates over baryonic matter, on small scales such as inside the BCG (Brightest Central Galaxy, which the DM halo is often centred on), this may not hold true. Gas and the BCG may in fact dominate over certain scales as can be seen in Fig. 1.9 from Mahdavi et al. (2007).

When constructing a mass model, different matter density profiles can be used with Navarro et al. (1997) suggesting the possibility of a "Universal Density Profile". The most commonly used profiles are an NFW and PIEMD (see §2.5.1 & §2.5.2 for more information). The results of the work by Sand et al. (2008), showed that a mass model for the cluster could be constructed that reproduced both the velocity dispersion for the BCG and the strong lensing features observed for Abell 383. The only problems were that

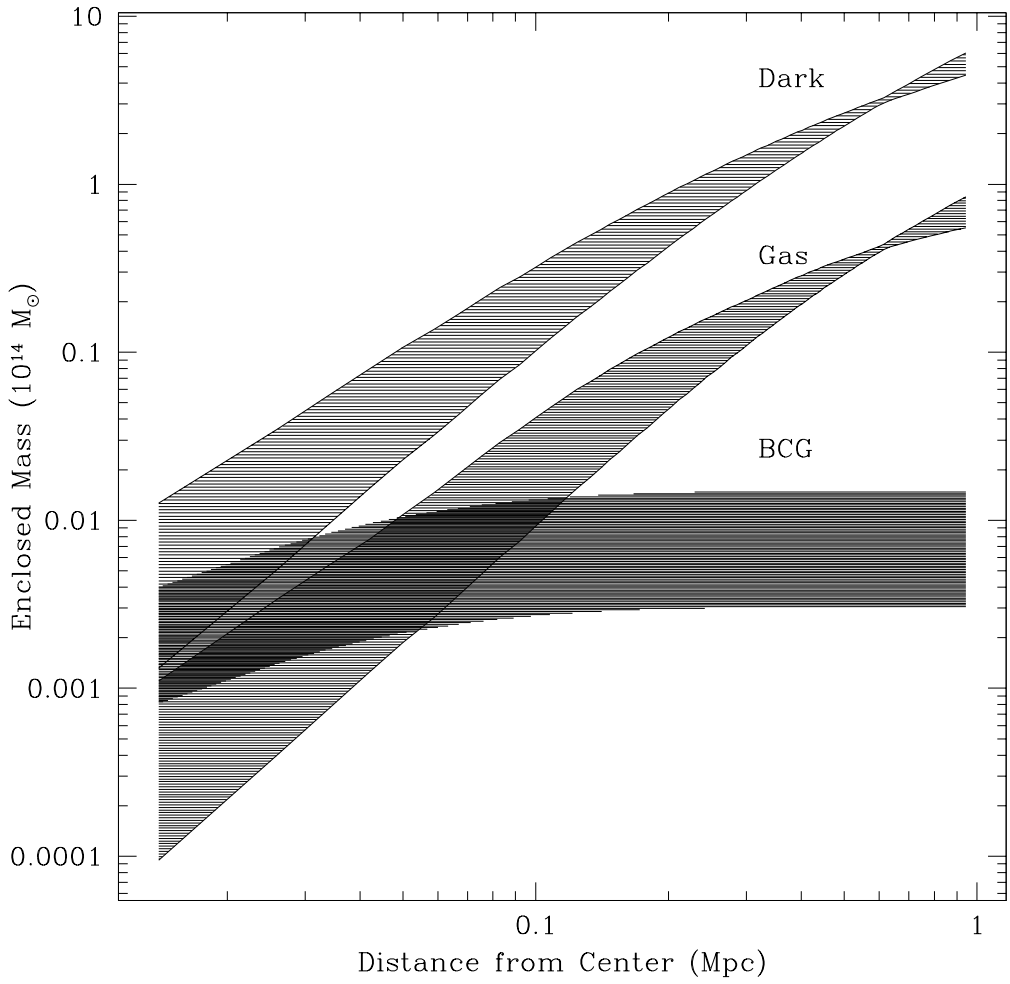


Figure 1.9: Figure taken from Mahdavi et al. (2007, Fig. 8) showing how the dominance of the intra-cluster gas, BCG and DM change depending on the distance from the cluster centre.

the values used were near the limits of acceptable constraints and only worked for very low resolution models. This could possibly indicate that not all clusters fit with universal profiles. The second cluster MS2137-23 could not be modelled at all, possibly indicating that either the values used as constraints were incorrect or the wrong mass model was chosen for this cluster.

This baryonic and dark matter interaction can also be probed both statistically and observationally, with the results giving cosmological constraints on the Universe. Limousin et al. (2008) attempted this with cluster Abell 1703. Abell 1703 was used due to its apparent circular morphology by the multiple image arrangements, and also due to four multiple images lying very close to the BCG, providing a very good constraint on the mass contained within this radii. Their results found that the NFW DM slope was $\alpha \sim 1.09$ (very close to the original NFW slope proposed by Navarro et al., 1996), but that numerical simulations of observation data needed to include both baryonic and DM particles in order to constrain models sufficiently to test possible cosmological values. Baryons and other effects, like feedback, can affect the simulation results (Bartelmann et al., 2013) and hence the resulting NFW DM slope.

Dark Matter Self-Interaction Cross Section

Lensing itself can cross into the realm of particle physics and set an upper limit on the self-interaction cross section of DM and thereby inform about how it would interact. This upper limit narrows down theories that predict it must have higher values. This is potentially very useful information for numerical studies of DM halo shape. Randall et al. (2008) used the Bullet Cluster to constrain DM self interaction to a value of $\sigma/m < 1.25$ $\mathrm{cm^2 g^{-1}}$, which is one of the lower values, with Bradač et al. (2008) performing a similar analysis with MACS J0025.4-1222 $\sigma/m < 4$ $\mathrm{cm^2 g^{-1}}$. While both values are low, and this is still a fairly new technique which, should improve with statistical stacking, this value is still several orders of magnitude higher than particle physicists predict that possible candidates have. Any DM candidate found by particle physics (e.g. by using particle

accelerators like CERN) will have to fit into the abundances and production rates of current cosmological understanding and observables, i.e. can a candidate be produced in the early Universe in sufficient numbers relative to baryonic matter, such that it can account for the dark / baryonic matter mass ratios observed in the Universe?

Testing Cosmology

Deep imaging of individual clusters can also provide support for standard cosmological models and enable highly detailed mass models of the cluster to be recreated, especially data that results in high resolution images that give the ability to pick out multiple images from a wide range of redshift background sources. In work published by Broadhurst et al. (2005), the deep imaging from *HST* enabled large numbers of multiple images over a wide redshift range ($1.0 < z < 5.5$) to be detected, even faint de-magnified ones lying close to the cluster core. The advantages to the mass model reconstructions were that they could trace the mass distribution accurately all the way to the DM halo centre and also measure, for the first time, the change in bending angle of light rays from sources that lie at different distances. This latter measurement is in agreement with a flat Universe cosmological model and is capable of confirming whether a predicted halo density model (such as the PIEMD or NFW density profiles described in § 2.5.2 & § 2.5.1 respectively) fits observational constraints or not across several radii from the cluster centre. Gravitational lensing can, in principle, provide a constraint on cosmological parameters due to the dependency of lensing on distances from cluster to observer and source image to observer. This produces a distance ratio that has a dependency on the cosmology of the Universe (Broadhurst et al., 2005).

Multi-Wavelength Comparisons

As numbers of gravitational lenses observed and the detail with which they can be studied grow, much work has been put into multi-wavelength (X-ray or Sunyaev-Zel'dovich) comparisons of large cluster samples. Smith et al. (2005) performed a comparison of

X-ray data against a combination of both strong (including where possible spectroscopic data on any arcs or multiple images) and weak lensing. This paper compared 10 X-ray luminous clusters, showing that seven of these were unrelaxed and set an estimation of how many clusters at a set redshift could be expected to be relaxed or not. Those that were unrelaxed were found to be significantly hotter than the relaxed clusters, suggesting the possibility that when clusters merge they boost the luminosities of the X-rays emitted from the gas (via the boosting of gas temperatures). This theory was followed up in Smith & Taylor (2008), in which 10 X-ray luminous clusters were studied, comparing theoretical infall cluster history models with the lensing-derived mass substructure fraction calculated inside a radius of R<250h^{-1}kpc. This substructure fraction was calculated using the equation:

$$f_{sub} \equiv \frac{M_{sub}}{M_{tot}} \tag{1.9}$$

Where M_{sub} is the mass contained in non BCG and main halo components and M_{tot} is the total cluster mass. Smith & Taylor (2008) using a semi-analytic approach, looked at the formation of these clusters at the points at which they had acquired 50% and 90% of their current observed masses (at $z \approx 0.2$). By doing this they found that clusters with a high f_{sub} fraction had formed and also acquired the bulk of their mass very recently (400 Myr for 10% or less rough mass growth) compared to clusters with a low f_{sub} (2 Gyr).

The first multi-wavelength comparison using gravitational lensing and the Sunyaev Zel'dovich effect (SZE), was performed in a LoCuSS paper by Marrone et al. (2009). A comparison of the Compton-Y parameter and the total cluster mass derived from gravitational lensing in a radius of 350 kpc from the cluster centre was performed. They found a link between the two values that did not seem to depend on whether the cluster was relaxed or not (a fact that can heavily bias X-ray mass estimates), highlighting the possibility of using SZE measurements to probe astrophysics in clusters, as well as showing promise for probing cosmology. Marrone et al. (2012) followed up this paper by increasing the sample from 14 to 18 x-ray luminous clusters with Compton-Y parameter

measurements and weak lensing mass measurements, producing a tighter $M_{WL} - Y_{sph}$ scaling relation than the earlier paper. They also found that the BCG ellipticity appeared to be a good tracer of under- or over-estimations of mass due to halo tri-axiality (see also results from the work for this thesis in chapter 5).

Another multi-wavelength comparison was carried out by Richard et al. (2010), where strong lensing mass and substructure measurements were taken within a radius of $R < 250kpc$ for 20 *HST* observed strong lensing clusters. When compared to X-ray data, a mass discrepancy of the order of ~ 1.3 was found between the X-ray and strong lensing derived masses. This increased with the mass fraction, indicating that the cluster was becoming more and more disturbed; as such the assumptions of hydrostatic equilibrium that the X-ray derived mass estimates used became more inaccurate. Their overall conclusions were that cluster-cluster mergers played an important role in what occurred in clusters that were being taken further away from hydrostatic equilibrium and possibly also disturbing their cool cores.

This $M_{SL} - MX$ relation should be linked to a $M - T_X$ relation, with predictions that they should be linked with $M \propto T_X^{\alpha}$ and $\alpha = 3/2$. This is due to virial theorem and a self-similar relation between the two values (the working was shown in chapter 5 § 5.7). Using 22 high redshift clusters ($z \gtrsim 1$) and performing a weak lensing analysis Jee et al. (2011) found that $\alpha = 1.54 \pm 0.23$, which is consistent with predictions and is another area of comparison followed in the strong lensing regime in chapter 5; however they did find some evidence that the relations' normalisation might well evolve with redshift and hence time.

1.6.4 Cluster Lensing

Strong lensing is the method that can be most useful for small scale structure and global estimates for the activity in cores of clusters. A cluster is in the strong lensing regime when the mass contained within a certain radius exceeds a certain critical value (see §2.1 for more details). If the alignment between source, lens and observer is correct, strong

lensing images can be observed, but will result in distorted images, multiple images and even a magnification of the source. This latter magnification effect can be useful when using gravitational lensing as a telescope with which to view objects that may be too faint to see by other means.

Gravitational lensing was most often observed using galaxy clusters as the lensing object, due to the fact that they are the most massive relaxed objects in the Universe and are also physically large on the sky (spanning tens of arcseconds in the strong regime in some clusters). This means they are more likely to produce strong lensing images and, if so, constraints can then be put on the mass that is constrained by those images. The disadvantages of this technique is that it is only sensitive to the mass that is causing the strong lensing and so may not be sensitive to any structures outside of the image constraints; it also requires images to be correctly matched and, if possible, their redshifts known (otherwise mass sheet degeneracy can be a problem, see §2.1). It is also only sensitive on small scales in the cluster cores, although conversely this can make measuring the density profile of DM at small radii possible to a high level of accuracy. It also requires very good and deep optical observations (the *HST* is excellent for this, due to it being outside the atmosphere) for detection of faint images, which limits the number of these observations that can be made, with no guarantee that any images will even be detectable, because the sample size will be small in comparison to other methods.

Gravitational lensing mass estimation provides a more direct estimation of the total cluster mass (and so can give insights into the DM content, due to the mass being directly probed, rather than indirectly, as in the other methods). The downside is that this method provides an estimation of the integrated mass along the line of sight, which if coupled with limited distance information, can create further problems, such as mass over estimation or problems with mass sheet degeneracy (see chapter 2). The direct mass estimation also raises problems in that the baryonic mass component cannot be de-coupled using lensing data alone as lensing measures all the mass, not a specific component.

1.7 LoCuSS

The Local Cluster Survey Substructure Survey (LoCuSS, PI: G. P. Smith) was formed to probe and establish relationships related to the structure and formation of galaxy clusters and their components. This came about as a result of work carried out in two papers, Smith et al. (2003) and Smith et al. (2005), with the core of the survey started with the LoCuSS pilot survey, which comprised ten clusters at a redshift of $z = 0.2$ that were studied in Smith et al. (2005). The LoCuSS group comprises of ~ 30 people working to obtain observational data (both ground and spaced-based, from the X-ray to radio spectrum) for around 165 massive X-ray luminous galaxy clusters.

The LoCuSS sample was of 165 cluster galaxies with an X-ray luminosity of $L_X \geq 2 \times 10^{44}$ erg/s (2.4keV). This was in order to select clusters with large masses, within a declination range of -70 to $+70$ degrees and a redshift range of $0.15 \leq z \leq 0.3$ that were indistinguishable from a volume limited sample (Okabe et al., 2010). The LoCuSS sample was selected from the *ROSAT* All Sky Survey (RASS, Ebeling et al., 1998, Ebeling et al., 2000 and Böhringer et al., 2004) and this choice influenced the selection due to the completeness of the RASS catalogues deteriorating beyond $z \geq 3.0$ (Smith et al., 2005). The selection via X-ray luminosity was deliberate due to the ready availability of X-ray catalogues (compared to ground-based optical ones) when *HST* time was applied for (Smith et al., 2005). This would also blind the selection function to the physical properties of the clusters (Okabe et al., 2010), which should then span the full range of cluster morphologies in order to reduce biases. The X-ray selection was thus used as a proxy for mass and will incur some bias, but this will be limited by the high cut applied as it should reduce contamination by non-clusters. The full sample was also lacking in complete spectroscopy, which limited some of the investigations that could be undertaken with some of the clusters (strong lensing was not possible for example).

The main goal of the LoCuSS survey was to calibrate mass scaling relations with other cluster observable properties. Identifying any systematic uncertainties and quantifying scatter in expected relations. Those clusters that had strong lensing features identified by

HST allowed the investigation of the mass and structure (along with X-ray temperatures and thermodynamics) of the cluster cores. Richard et al. (2010) produced the largest sample of LoCuSS strong lensing clusters for investigation, but with this work the sample size was further increased, making it the largest LoCuSS strong lensing sample to date. This larger sample should allow enhanced tracing of the underlying population, better identification of outliers, a decrease in noise that and, as a result, provide improved calibration of mass scaling relations with temperature and quantification of the resulting large scatter.

The sample in this work was taken from this larger LoCuSS sample and contributed to its completeness, forming the largest sample (23) to date of strong lensing clusters offering multi-wavelength data taken in a consistent way. At the time of this report not all 165 clusters had been fully observed or analysed.

1.8 Thesis Overview

This thesis used several inter-linked methods to gain a deeper understanding of clusters and to determine how multi-wavelength data could be used to understand what was occurring within the cores of clusters. Without redshifts to spectroscopically confirm multiple lensing images (if morphology or colour selection cannot be carried out either), strong lensing analysis cannot be performed accurately as large errors will result. Therefore the large amounts of spectroscopic data for each cluster (chapter 3) were reduced. The methods used, results and any spectroscopic confirmation of potentially lensed image candidates are given in each chapter as they are used. Additional information about the cluster (either structurally or globally) obtained by selecting potential cluster galaxy members from redshift space and links to other observables are discussed. Potential disturbances were shown by producing DS plots and calculating DS statistics.

Utilising redshifts obtained in chapter 3, Chapter 4 describes the method for modelling mass distributions from strong lensing constraints using the most interesting cluster Abell

3084 from the sample (indicated by its image configuration). The methods and reasoning for selecting sensible priors for the lens models are also described. These priors include selecting how many mass components (DM and baryonic) to include and what the values for those components should be (are they free to move within a parameter space or fixed?). Using the program LensTool, parametrised results were produced as a result of fitting models to the strong lensing image constraints. Other multi-wavelength derived measurements were compared and discussed to build a picture of what is occurring in the cluster core, along with information on its recent history.

In chapter 5 a similar methodology was followed as for Abell 3084, parametrising the remaining clusters possessing strong lensing imaging. The results of this produced mass models that were added onto an existing LoCuSS cluster sample (Richard et al., 2010). The increased sample size was then used for comparison with other cluster observables, derived from multi-wavelength observations, to construct a mass versus X-ray temperature relation. Examining the large intrinsic scatter evident in this fit, residuals were compared to cluster observables to explain the deviation from M_{SL} (the measured mass) and M_{fit} (predicted mass).

The final chapter contains a summary of the work performed and discussion of future tasks that could build upon these results.

The areas this thesis intended to investigate were:

1) What does a multi-wavelength approach yield over just a purely strong lensing one?

2) The physics of cluster cores.

3) What can be inferred about clusters as a whole and any cluster mass relations to other cluster observables?

The goals set above were intended to determine whether or not multi-wavelength data was important in determining cluster history and formation or if a purely strong lensing analysis was sufficient. It was also important to see whether this additional information changed the interpretation of a purely strong lensing analysis. For instance, were some clusters being misidentified as relaxed whilst in reality they were still under the effects of

a merger? Naturally, this would affect the interpretation of the cluster population as a whole.

The physics that occurs in cluster cores is important for understanding the merging of clusters. N-body simulations suggest a hierarchical merging with BCG's forming by accreting smaller galaxies. Would examination of the lensing configurations of unusual clusters provide unique insight into certain epochs of a cluster's merger history? If so by examining what occured in the cores of these clusters, would this agree with current cosmology models? Observationally clusters are seen to undergo mergers or are in the process of settling down soon after, but little is known about the settling of clusters beyond this. Other recent works showed a non-zero offset between BCG's and their X-ray peaks and hint at the possibility of this offset characterising merger and relaxation histories (Zitrin et al., 2012). The ability to confirm this observationally would be advantageous as it would allow the calculation of merger probabilities and duration. For instance, N clusters are predicted to be in the late stages of a merger at any one time. A useful N-body simulation comparison, though with only one rare lensing configuration in this sample, would be required for a firm link to be established.

Can a larger sample of strong lensing clusters provide information on their dynamical state or mass distribution? The final goal was to determine what a sample of strong lensing clusters revealed about clusters as a whole. Would strong lensing mass correlate with predictions from cluster observables and would these agree with similar comparisons but utilising weak lensing measurements? Both weak and strong lensing measurements have to agree (as both are the result of the deflection of light by a mass) for any correlation to be confirmed. The ability to quantify the scatter in an mass-observable relation would allow for corrections of this effect. Overall for astronomy, the ability to associate a more easily observed quantity and to tie it to a cluster's mass would enable mass determinations for clusters that do not possess strong lensing features or deep enough imaging to take weak lensing measurements. This would allow a mass function to be constructed (at least for the high mass tail end of the function), which was a constraint on cosmology.

CHAPTER 2

THEORY

This chapter explains and conveys the basic concepts, knowledge and techniques used in this thesis. In preparation for the chapters that follow a description of lensing theory is detailed first to explain the mechanism by which strong, and weak, lensing images occur. Bayesian statistics and model selection follows, ready for a description of the working mechanism for LensTool (based on Bayesian statistics) and how it parametrises the values of these DM density model values. The chapter finishes by describing the two most popular (NFW and PIEMD) descriptions for the density profiles of DM halos that can be used in LensTool.

2.1 Lensing Theory

Gravitational lensing is simply the application of Fermat's Principle to a light ray passing near a massive object and its associated gravitational field. The light ray will take the route that results in its earliest arrival (generally a straight line) when travelling between two points. Because mass distorts space-time around it, the light ray takes the shortest route in four dimensional space, which to the observer appears to be a curved path and not a straight line. Therefore the more massive the object or the closer the light ray approaches the object, the more curved a path the light ray will take.

Gravitational lensing can be broken down into three distinct categories: strong, weak

and micro lensing. Strong lensing analysis was the method used to determine cluster masses in this work; the next sections outline its workings.

2.2 Strong Lensing

This was the most visually obvious form of lensing, due to its characteristic production of multiple images of the same background source and, if the alignment was sufficient, with the mass distribution itself also being circularly symmetric, Einstein rings. An example of this multiple image behaviour is shown in Fig. 2.1. The advantage of strong lensing was that the arcs gave a good constraint on the mass inside the cluster centric distance from the arc, as the density must exceed a critical value within the area swept by this radius. The cluster centric distance is that from the cluster centre to the arc, or if not cluster centred, the distance from the arc to the centre of the circle the arcs appear to lie on. The need for a strong gravitational potential limits the distance inside which strong lensing effects (rings, multiple images and arcs) are expected to occur from the cluster centre. This limit in radius only enables estimation of the mass contained in the cluster within this limited radius, although this can be overcome using weak lensing, which can provide mass measurements out to much further cluster centric radii. Strong lensing analysis is the main gravitational lensing method used in this thesis.

2.2.1 Lens Equations

From Einstein's theory of General Relativity, any mass concentration will cause a deflection in the path of passing photons resulting in a deflection angle per unit length, $\delta\alpha/\delta l$, (Rey, 1999) given by:

$$\frac{\delta\vec{\tilde{\alpha}}}{\delta l} = -2\vec{\bigtriangledown}_{perp}\frac{\phi}{c^2} \tag{2.1}$$

Where l is along the line of sight from the source to the observer, c is the speed of light, ϕ is the Newtonian potential (which is located at the lens position) and the derivative

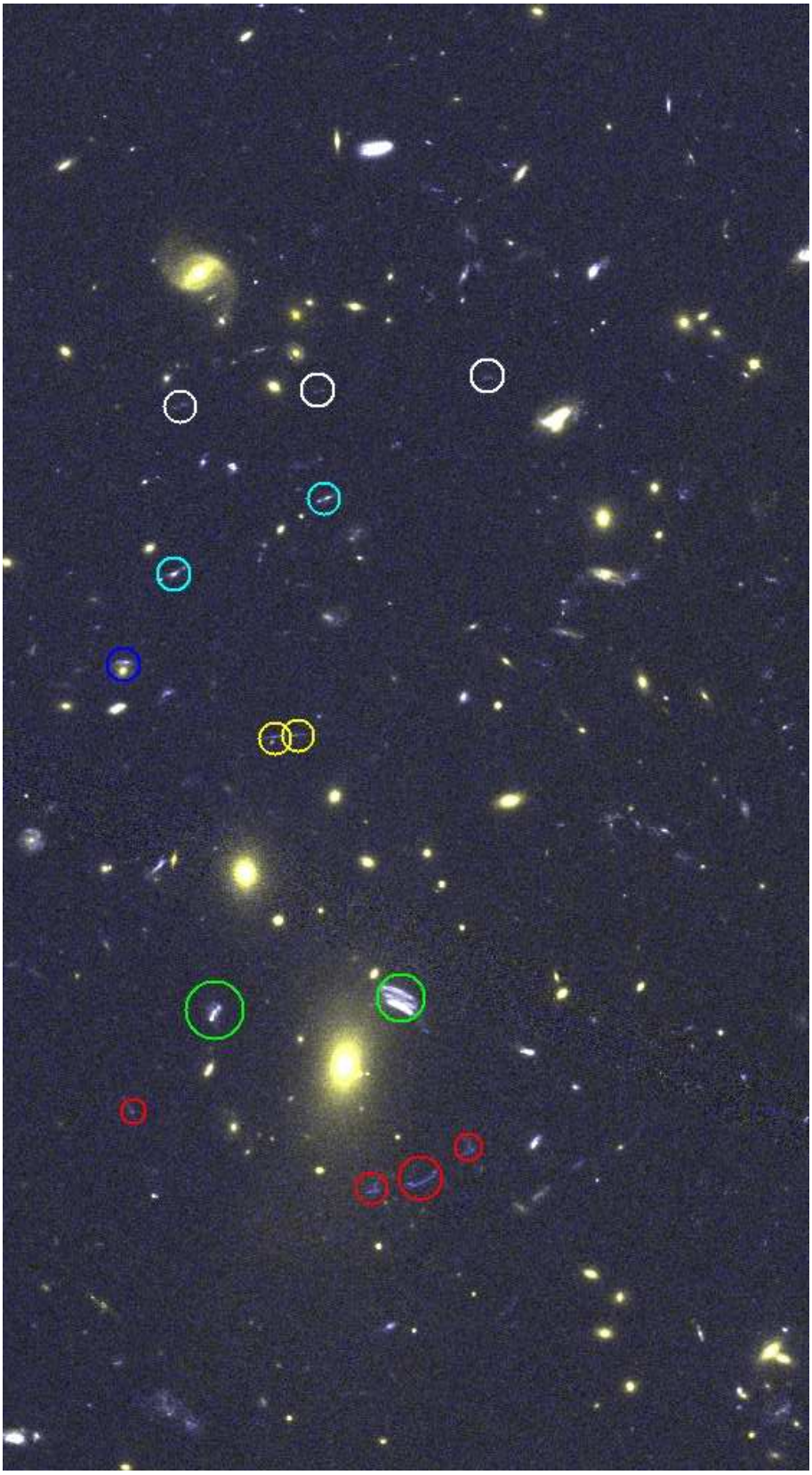


Figure 2.1: Strong lensed multiple images in the system MACS0257. Images of the same background sources are circled in matching colours. Images identified and figure created by author.

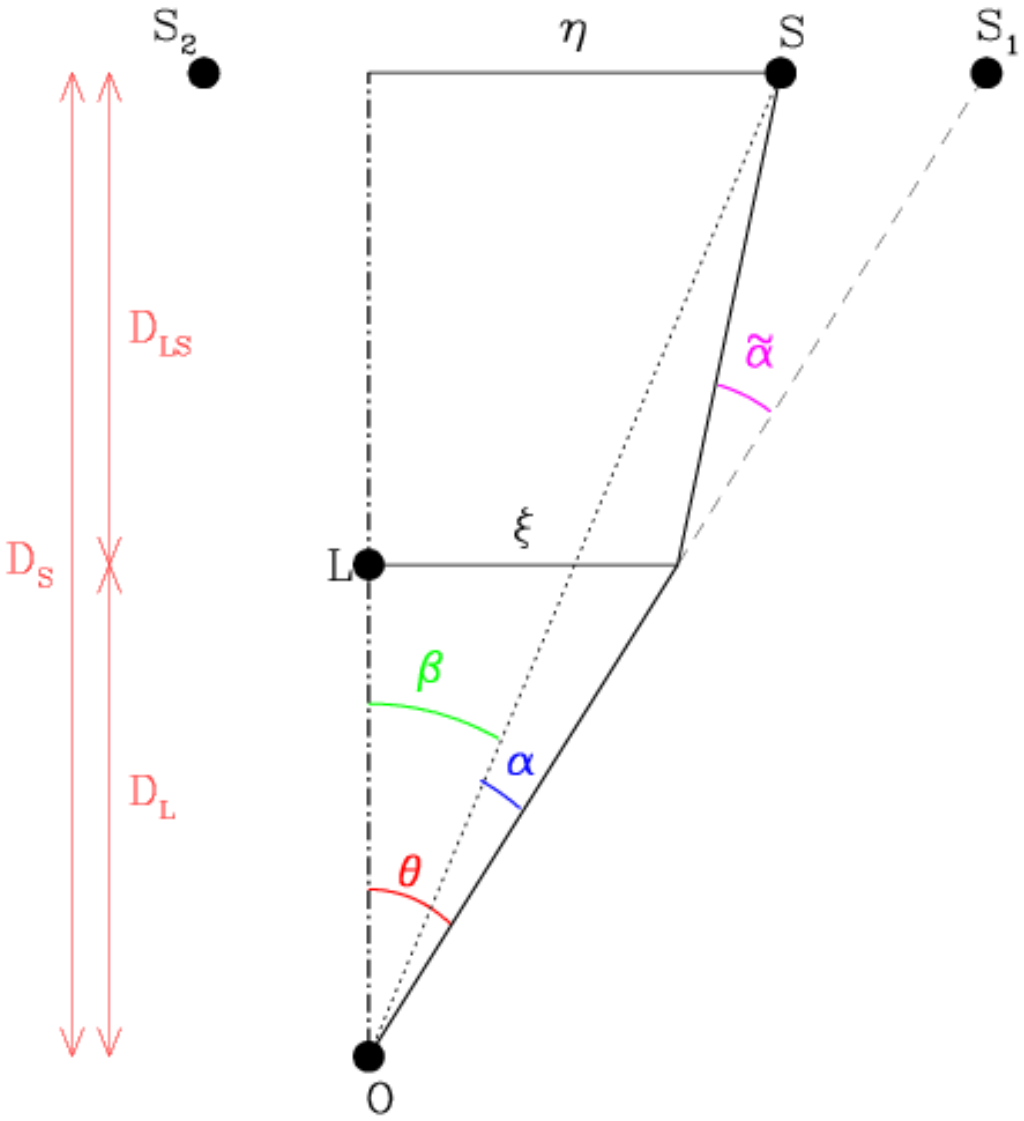


Figure 2.2: A typical simple arrangement of the source, lens and observer showing the deflection produced (as used in equation 2.1 onwards). S is the source image lying on the source plane, S_1 & S_2 are the apparent positions on the source plane of the source after the light rays had been deflected. L is the lens plane and O the observer. Figure taken from Wambsganss (1998)

is taken perpendicular to the photons direction of travel. Note all angles and distances can be seen on Fig. 2.2 for convenient comprehension. Integrating along the l returns the total deflection $\vec{\tilde{\alpha}}$ (see Fig. 2.2) as a result of the deflection of the Newtonian potential at the lens position. Giving (Narayan & Bartelmann, 1996):

$$\vec{\tilde{\alpha}} = \frac{2}{c^2} \int \vec{\nabla}_{perp} \phi \; dl \tag{2.2}$$

The deflection angle $\vec{\tilde{\alpha}}$ can now be seen to be as a result of the Newtonian potentials gradient perpendicular to the line of sight, l.

Combining equations 2.2 with the lens equation derived result of $\alpha = \tilde{\alpha} D_{ls}/D_s$ (equation 2.7) gives:

$$\vec{\alpha} = \frac{D_{ls}}{D_s} \frac{2}{c^2} \int \vec{\nabla}_{perp} \phi \; dl \tag{2.3}$$

By defining ψ (Narayan & Bartelmann, 1996), the lensing potential as:

$$\vec{\nabla}_{\theta} \psi = \frac{D_{ls}}{D_s} \frac{2}{c^2} \int \vec{\nabla}_{perp} \phi \; dl \tag{2.4}$$

Results in a relation between the scaled deflection angle and the gradient of the lensing potential:

$$\vec{\alpha} = \vec{\nabla}_{\theta} \psi \tag{2.5}$$

Equation 2.5 demonstrates how the gradient of the lensing potential results in a deflection of a photon's path. This illustrates that once the distances of the source and the lens are known, solving the lens equation (see equation 2.8) allows the mass distribution (that generates the Newtonian potential) of the lens to be constrained. In practice the total deflection angle is small (generally no more than an arcminute, Rey 1999), and so the light ray paths can be modelled as straight lines (the Born approximation) that only bend as they cross a thin plane where the lensing mass lies (the lens plane), with the

bending being instantaneous and not gradual. The size of the largest clusters (the largest massed objects) when compared to cosmological distances is still very small, which also lends to this thin lens approximation being applicable. This approximation was used above in generating this result.

Fig. 2.2 shows a light ray being bent under the thin lens approximation with the bending of the light ray's path only occuring as it crosses the lens plane. Examination of the relationship between the angular diameter distances in the source plane allows the following relation between the distances and angles to be made:

$$D_s\vec{\beta} = D_s\vec{\theta} - D_{ls}\vec{\tilde{\alpha}}(\vec{\xi}) \tag{2.6}$$

Where D_s is the distance from the source plane to the observer, D_{ls} the distance from the lens plane to the source plane, D_l the distance from the lens plane to the observer and $\vec{\xi} = D_{ol}\vec{\theta}$. This utilizes the small angle approximation where $cos\alpha \approx tan\alpha \approx sin\alpha \approx \alpha$. Please note that in general $D_s \neq D_{ls} + D_l$.

It can be seen from Fig. 2.2 that:

$$\vec{\alpha}(\vec{\theta}) = \frac{D_{ls}}{D_s}\vec{\tilde{\alpha}}(\vec{\xi}) \tag{2.7}$$

Where $\vec{\alpha}(\vec{\theta})$ is the scaled deflection angle. Combining equations 2.6 and 2.7 results in the lens equation:

$$\vec{\beta} = \vec{\theta} - \vec{\alpha}(\vec{\theta}) \tag{2.8}$$

This result shows that if the source (for extragalactic objects this is often not observable) and image positions are known, constraints can be put on the deflection angle and hence the Newtonian potential. This equation is generally non-linear in that multiple θ can exist for a given β.

For a point-like mass distribution acting as the lensing potential:

$$\vec{\hat{\alpha}}(\vec{\xi}) = \frac{4GM}{c^2\vec{\xi}} \tag{2.9}$$

Where G is the gravitational constant, M is the mass of the lens, and ξ the impact parameter. This equation is also dependent on the impact parameter being much larger than the Schwarzchild radius, R_s, (for a given mass, it is the radius within which if the mass was compressed the escape velocity would equal the speed of light), with $\xi >> R_s \equiv 2GMc^{-2}$, otherwise light will not escape from the latter radius. The Schwarzchild equation governs the radius inside of which mass has to be compressed in order for it to collapse into a black hole. Equation 2.9 is also the well-known that served as one of the first tests of General Relativity (Eddington, 1919) due to its factor of 2 difference from Newtonian derived deflection. By changing M (the total lens mass) to be $M(\xi)$ (the lens mass contained in a radius ξ), equation 2.9 becomes the expression for a circularly symmetric lens.

Equation 2.9 can be put into equation 2.8 in the format:

$$\tilde{\alpha}(\xi) = \frac{4GM}{c^2\theta D_l} \tag{2.10}$$

Where it can be seen that $\xi = D_l\theta$ (see Fig. 2.2). Using the scaled deflection angle relation in equation 2.7 gives the lens equation as:

$$\beta = \theta - \frac{D_{ls}}{D_s D_l}\frac{4GM}{c^2\theta} \tag{2.11}$$

From this it can be clearly seen that if $\beta = 0$, implying a perfect alignment of source, lens and observer generates an Einstein Ring:

$$\theta_E = \sqrt{\frac{4GM}{c^2}\frac{D_{ls}}{D_l D_s}} \tag{2.12}$$

Where θ_E is the angular radius of the Einstein ring in arcseconds and M is the mass of the lens. This equation gives the angular radius of the ring as it appears on the sky and not its actual radius in distance. This could mean that for some objects that are acting as a lens, the Einstein radius size is such that the resultant ring is masked by the lensing object i.e. the Einstein radius is smaller than the lensing object's physical radius. The number of observed Einstein rings should be low, as a perfect alignment is not expected to occur very often on the sky and the lensing mass distributions tend not to be generally symmetrical either; instead, multiple images and arcs would be more prevalent.

To show this effect, the situation where $\beta \neq 0$ is considered and then equations 2.11 and 2.12 can be used to give:

$$\beta = \theta - \frac{\theta_E^2}{\theta} \tag{2.13}$$

Solving this for θ gives the positions of the multiple images of the source.

$$\theta_{\pm} = \frac{1}{2}\left(\beta \pm \sqrt{\beta^2 + 4\theta_E}\right) \tag{2.14}$$

A point-like mass distribution results in two multiply imaged sources at θ_+ and θ_- for a given β.

In general, lensing masses will not be point-like but have their mass concentrated and spread over an area. The shape and distribution of this mass distribution can cause large numbers of multiple images to be produced (some smeared together), but always the total number of images is an odd number (see § 2.2.3). The occurrence of two images in this example instead of three is due to the singularity at the centre of the point-like mass distribution. If the mass however was not a point mass but was spread out into a concentrated mass distribution, three images would be produced, but the third would lie close to the centre of the mass distribution and with extremely small magnification (Meylan et al., 2006). For a point-like mass distribution, light rays passing close to the top of the mass centre are deflected far below it and *vice versa*. This discontinuity can be shown non-trivially to not violate odd number theory.

For either the Einstein ring and/or multiple images to occur, the density within a certain radius must generally exceed a critical value, which is determined by the critical surface mass density:

$$\sum crit = \frac{c^2}{4\pi G} \frac{D_s}{D_l D_{ls}} \tag{2.15}$$

Where G is the Gravitational constant and σ_{crit} is the critical surface mass density. The value that is compared to equation 2.15 is the surface mass density given in:

$$\Sigma(\boldsymbol{\xi}) = \int dr_3 \rho(\xi_1, \xi_2, r_3) \tag{2.16}$$

Where dr_3 is the direction of travel for the light ray and ρ is the three-dimensional density of the mass distribution. Where the surface mass density is ρ integrated along the line of sight and projected onto the lens plane to obtain a two-dimensional surface mass density. $\boldsymbol{\xi}$ is a two component vector in the lens plane, where $\boldsymbol{\xi} = (\xi_1, \xi_2)$.

"Convergence", κ here, is related to the lensing potential and is defined as:

$$\kappa = \frac{\sum(\boldsymbol{\xi})}{\sum_{crit}} \tag{2.17}$$

When the surface mass density $\sum(\boldsymbol{\xi})$ is greater than $\sum_{crit}$, κ will be ≥ 1. This is normally used as the indicator of the possibility of encountering the effects of strong lensing, with a $\kappa << 1$ being in the weak lensing area. κ categorises the strength of the potential and, therefore, close into the cores of clusters the surface mass density could exceed the critical value and cause the creation of multiple images. κ describes the isotropic focussing of the background light causing magnification (or de-magnification) as the source is mapped onto an image of the shape but of a differing size (larger size magnifies and *vice versa*).

2.2.2 Magnification and Distortion

Lioville's theorem (photons all evolve the same way with time), coupled with the lack of photon absorption or emission due to gravitational lensing means that the density of photons does not change with time, but in fact will conserve surface brightness (Schneider et al., 2006). Gravitational lensing alters all wavelengths of light by the same amount, enabling spectroscopic redshift determinations to be taken of lensed images (this is one of the ways of matching images to the same background source). The lensing effectively focuses the light from the source, meaning that additional photons that would not have been observed, have now been deflected by the lens to be observable, changing the source's solid angle. This means that the larger the area of the lensed image the larger the magnification. In simple terms the magnification μ, will be the ratio of difference between an element of the source image area ($\delta\beta^2$) mapped onto an area of the lensed image ($\delta\theta^2$). Giving:

$$\mu = \frac{image\ area}{source\ area} = \frac{\delta\vec{\theta}^2}{\delta\vec{\beta}^2} = \frac{1}{detA(\vec{\theta})} \tag{2.18}$$

The magnification factor μ can be calculated from the inverse determinant of the Jacobian matrix (Meylan et al., 2006), $A(\theta)$ (see equation 2.19). This introduces the idea of parity in images, as these can have negative or positive magnification, resulting in mirror images of the source (the distortion of the images means this is not always apparent by observation). The Jacobian matrix describes the lensing transformation between the source and image planes, with:

$$A(\vec{\theta}) = \frac{\partial\vec{\beta}}{\partial\vec{\theta}} = (\partial_{ij} - \frac{\partial^2\psi(\vec{\theta})}{\partial\theta_i\partial\theta_j}) = \begin{pmatrix} 1-\kappa-\gamma_1 & -\gamma_2 \\ -\gamma_2 & 1-\kappa+\gamma_1 \end{pmatrix} \tag{2.19}$$

It also shows that it is calculable in terms of the lensing potential ψ, but can be expressed in terms of convergence κ (the volume mass density) and shear into a 2×2 matrix. $\vec{\theta}$ actually being a two dimensional vector in the form $\vec{\theta} = (\theta_1, \theta_2)$. The Jacobian matrix also contains the components of shear γ (see equation 2.20 for the breakdown

of γ into components), κ which is the volume mass density and ψ which relates to the Newtonian potential ϕ (as detailed in equation 2.4). With $\kappa \equiv (\psi_{11} + \psi_{22})/2$ (κ causing isotropic magnification of the source). Magnification can therefore be useful in breaking any degeneracy in the model if the original source dimensions are known and allows lensing objects to act as gravitational telescopes.

The shear terms are defined as:

$$\gamma_1 = \frac{1}{2}(\psi_{11} - \psi_{22}) \quad , \quad \gamma_2 = \psi_{12} \tag{2.20}$$

Where $\psi_{12} = \partial_{\theta_1}\partial_{\theta_2}\psi$ and so on (shear is related to the second derivatives of the lensing potential, Rau et al., 2013). The Jacobian matrix contains terms of γ and κ, which alter the size and shape of the lensed image resulting in distortion from its original non-lensed shape and size. With the components of γ distorting the shape of the source image (anisotropy in the lens mapping). For example, for a circular background image the shear maps it onto an ellipse. These two effects combined can distort a circular source image into a long thin curved arc.

2.2.3 Image Numbers

In multiple image systems involving galaxy lensing, the number of images is always odd (assuming the image does not lie on a caustic) and was shown in Burke (1981). However in actual life observations only even numbers of images may be observable. In this case an image will often lie near (or behind) the lensing object and can be highly de-magnified (the received flux is sensitive to the steepness of the potential), making it hard to pick out against the lens object.

2.2.4 Critical Curves and Caustics

On the lens plane, there exist areas where the Jacobian (equation 2.19) is zero. These tend not to exist as singular points but as curves (often smooth and closed) that lie on the lens plane; they are called "critical curves". These can then be mapped onto the source plane, giving the lines known as "caustics". These possess interesting properties, such as:

- Objects that are near to caustics can produce multiple images that appear close to the corresponding critical curve, but with very high magnifications. This is caused by the fact that images that are near to a critical curve have their magnification μ (see equation 2.18) increased to very high values, as the magnification equation diverges the closer to the caustic they become. μ will never attain infinity, however, as it will reach a point where the assumptions made earlier (enabling the use of geometric optics) are no longer applicable and a wave optics solution will have to be performed instead (the diffraction patterns only give finite magnification values).

- If an object crosses a caustic line, a pair of images is either created or destroyed. The position of the object and, which side of a caustic it lies on, governs how many images appear near the corresponding critical curve. The pair will be opposite in parity (parity is conserved), due to the fact that the magnification μ changes sign as it crosses over a critical curve and the image pairs will be either side of the critical curve.

Figure 2.3 displays two simple examples of critical lines and caustics.

2.2.5 Image Classification

Equation 2.21 describes the light-travel time for a light ray and also enables images to be classified according to whether an image is located at a minimum, maximum or saddle point of τ (described below). Each multiple image will have a different light-travel time, resulting in a time delay between images. If the source has an intrinsic time variability

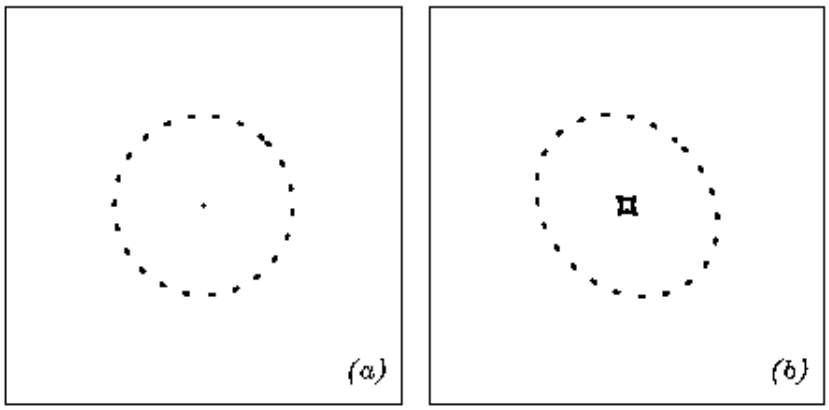


Figure 2.3: Critical lines are dotted and caustics are solid lines in this figure. Figure (a) is for a singular isothermal sphere and Figure (b) is for a singular isothermal ellipse. It should be noted that just by changing the symmetry (breaking it in this case), the caustic is now changed from a point to a star shape. Figure taken from Dr. Graham Smith's 2002 thesis - "Gravitational Lensing by X-ray Luminous Galaxy Clusters".

then the multiple images will display this. However, due to differing light time arrival (at the observers position) between the images, the variability will not occur at the same time for all the images. If variability can be observed it can be utilized as a means of picking out images that are from the same object. The time delay function for gravitationally lensed images (which shows the travel time for a light ray emitted at position β traversing the lens plane (at a redshift of z_{lens}) at position θ and arriving at the observer) can be written as (Wambsganss, 1998):

$$\tau(\theta;\beta) = \tau_{geom} + \tau_{grav} = \frac{1+z_{lens}}{c}\frac{D_l D_s}{D_{ls}}\left(\frac{1}{2}(\theta-\beta)^2 - \psi(\theta)\right) \tag{2.21}$$

This equation for the function $\tau(\theta;\beta)$, illustrates how the time delay is a combination of the geometric time delay τ_{geom}, due to the increased distance the light ray has to travel compared to a straight line, and the gravitational time delay τ_{grav}, caused by the gravitational potential causing the lensing. This gravitational time delay effect was first proposed as a solar system test of General Relativity by Shapiro (1964) and was subsequently confirmed (known as Shapiro delay). Other parameters in equation 2.21 are the same as for equations 2.8 and 2.5. It should be noted that the the lens equation (see

equation 2.8) can be re-written as:

$$\bigtriangledown \left(\frac{1}{2}(\theta - \beta)^2 - \psi(\theta) \right) \tag{2.22}$$

With the terms in brackets being the same as those in brackets on the right hand side of equation 2.21. From this it can be noted that:

$$\bigtriangledown \tau(\theta; \beta) = 0 \tag{2.23}$$

This is equivalent to the lens equation (equation 2.8) and shows that any images lie on points where τ is stationary and subsequently at minimum, saddle or maximum of the light travel time. This also follows Fermat's principle (light travels the path of least travel time).

The multiple images can be classified using the Jacobian matrix (equation 2.19) since it is the Hessian of τ, with the eigenvalues of A being positive for a minimum, negative for a maximum and of different signs for a saddle point.

2.3 Mass Sheet Degeneracy

One problem in lensing is "mass-sheet degeneracy". This occurs when insufficient lensing constraints exist for a given lensing system, such that the mass reconstruction cannot be fully realised. For example, if a sheet of uniform mass was placed in front of a cluster, it could return the same results (for the same observational constraints) as for the cluster without the extra mass-sheet in front of it, but both would possess different cluster properties (and hence mass). This is because the convergence can also be described as (from Bradač et al. 2004):

$$\kappa \rightarrow \kappa' = \lambda\kappa + (1 - \lambda) \tag{2.24}$$

Where κ undergoes a degeneracy transformation and λ is an arbitrary constant. This degeneracy can be broken if there are two or more sources at differing known distances. This is because the critical surface mass density is source distance dependent (see equation 2.15). This degeneracy can also be broken if the value of the magnification can be determined for the images. This is an especially big problem with weak lensing (as multiple images do not exist to use as constraints) and so sets a finite accuracy on mass reconstructions if no further assumptions or constraints to the model are made. Magnification can also be hard to measure as the original dimensions for the source cannot be determined directly.

2.4 Bayesian Theory

The basic principles of Bayesian probability are that a statistical measure can be made that hypothesis H is valid given a set of data D and using background information I. Equation 2.25 shows how these are related:

$$P\left(H \mid D, I\right) \propto P\left(D \mid H, I\right) \times P\left(H \mid I\right) \tag{2.25}$$

Where $P\left(H \mid I\right)$ is the probability the hypothesis is true given prior information and is called the "prior" probability, and $P\left(D \mid H, I\right)$ is the probability that the data is true given the hypothesis and information, known as the "likelihood function". The likelihood function requires all the relevant information be included, as omittances can have a significant effect on the calculation. The resulting probability, $P\left(H \mid D, I\right)$ (the resulting final value being calculated) is the resulting "posterior" probability, which is the probability the hypothesis is true given the data and information. The reason equation 2.25 is proportional is due to the lack of $P\left(D \mid I\right)$ or "evidence", which is the probability the data, D, is true given that the information, I, is true. This is omitted to simplify the equations and regardless of which theory is used; it normalises the equation. $P\left(D \mid I\right)$ cannot always be omitted as standard, however, because in model selection (see §2.4.1) its

inclusion is of vital use to select between models. In its full form equation 2.25 becomes:

$$P(H \mid D, I) = \frac{P(D \mid H, I) \times P(H \mid I)}{P(D \mid I)} \tag{2.26}$$

A useful function of the equation is that as more data is obtained, the posterior probability from the last function can be folded into the new calculation (utilising this new data) as the prior. This could be done for each measurement as it arrives from an experiment or calculated all together at the end; the results are the same. This is possible because Bayesian Evidence depends on the data not relying on what came before it (each point is independent of one another).

2.4.1 Model Selection

Bayesian statistics can help to determine which is the preferred, probability wise, model when there are two competing theories (or models) which both fit the data but which are each inherently different.

With just two differing theories (or models), then in the simplest case, the ratio of the posterior functions are taken:

$$posterior\, ratio = \frac{P(A \mid D, I)}{P(B \mid D, I)} \tag{2.27}$$

$P(A \mid D, I)$ and $P(B \mid D, I)$ are the posterior probabilities of both theories A and B, i.e. the probability that the theory is correct, given the data and prior information is true. The interpretation of this ratio is that a value much greater than one will strongly favour theory A, but a value much less than one will strongly favour theory B. Any value near unity is unable to distinguish between the two models (though the simplest model may be better just from a simplification standpoint). Expanding the posteriors results in:

$$\frac{P(A \mid D, I)}{P(B \mid D, I)} = \frac{P(D \mid A, I)}{P(D \mid B, I)} \times \frac{P(A \mid I)}{P(B \mid I)} \tag{2.28}$$

Because of the expansion of these terms, substituting the priors for equation 2.26 yields an equation with a dependence on the prior information. The problem with this is that the ratio of priors biases the equation towards the original theory at the start of the calculation. For unbiased selection this ratio of priors should be set to one.

If one theory is more complex than the other (B in this simple case) and has an adjustable coefficient (for example, in a lensing model it could be the halo's cut-off radius or velocity dispersion is allowed to vary) then the "likelihood function" requires the integration of this coefficient, C. The "likelihood function" for B then becomes:

$$P\left(B \mid D, I\right) = \int P\left(D \mid C, B, I\right) \times P\left(C \mid B, I\right) \mathrm{d}C \tag{2.29}$$

Where $P\left(D \mid C, B, I\right)$ is now just a "likelihood function" as everything else is given and $P\left(C \mid B, I\right)$ is the "prior" for the coefficient, C.

When carrying this through to its conclusion to produce a ratio of the two theories' posterior term, the resulting equation will contain a bias for the more complicated model and a balancing term to penalise for this (otherwise more complicated models would always be favoured over less simple). The bias comes from the fact that because, theory B has an adjustable coefficient, it will more likely trace the data. Therefore, the "likelihood" function will be favoured towards B and so a balancing term also appears. The balancing term comes from the integration term, coupled with the "prior", forming a penalty term penalising for the use of B. This penalty term is often called the "Ockham Factor"; unless the "likelihood" function very strongly favours B, the penalty function will favour the simpler model.

This can also be applied in a case where each theory has a changeable coefficient.

2.4.2 Markov Chain Monte Carlo

From § 2.4 it is easy to see that the parameters for a model close to the true parameters will return a high posterior probability and parameters far away from the true values

a low posterior probability. Sampling the full volume of parameter space can in many cases be computationally prohibitive such as where there are complex functions with many parameters. Instead only a sampling of this space can be undertaken. Sampling only areas where the posterior probability is high can provide a much more efficient estimation of the true parameter values than by pure random sampling of the parameter space. Markov Chain Monte Carlo is an example of this.

The basic operating principles of Markov Chain Monte Carlo (MCMC) is that it begins by selecting a sample from a sample space (covering all the possible parameters that go into a function) and constructing a Markov Chain. A Markov chain is a random process where the next state only depends on the current state and not any of the previous states. The MCMC takes the parameter sample for this chain and calculates the posterior for it, by then taking another sampling of parameter space and calculating the posterior again, the two posterior values can be compared. If the new samples posterior is preferred (this is described in more detail below) to the old samples, it takes this as the new sampling, taking again another sampling from parameter space and comparing the resulting posterior value, and so on. If the old posterior is preferred, it discards the new sample and keeps the old, beginning again another sampling from parameter space. By working in this way, an estimation of the true parameter values is given (based on samples with high posterior probabilities), without exploring all of parameter space. LensTool is an example of a program that uses this method in order to optimize the parameters required to reconstruct a model that recreates any strong lensing constraints.

As MCMC is still an evolving field and refinements continue to be made, only a simple case of MCMC is considered by using one of the more basic methods, that of the Metropolis-Hastings algorithm.

The expected value, X for a function $f(X)$ is simply given as the integral of the function over X with each value of the function multiplied by its posterior as a weighting factor, as shown in equation 2.30 below.

$$\langle f(X)\rangle = \int f(X) \times P(X \mid D, I) dX = \int g(X) dX \tag{2.30}$$

With $g(X)$ representing these two values multiplied together. As stated above, in order to get an exact value would require exploring all of the possible range that the value X could take. This is very often not feasible, especially for multi-parameter investigations, so an estimation can be made by sampling randomly (and also uniformly) from the possible range of X values N times. This turns equation 2.30 into the form:

$$\langle f(X)\rangle = \int_V g(X) dX \approx V \times \langle g(X)\rangle \pm V \times \sqrt{\frac{\langle g^2(X)\rangle - \langle g(X)\rangle^2}{N}} \tag{2.31}$$

This yields an approximate value for X within the volume V of the range of possible values, along with an error bar, with:

$$\langle g(X)\rangle = \frac{1}{N}\sum_{i=1}^{N} g(X_i) : \langle g^2(X)\rangle = \frac{1}{N}\sum_{i=1}^{N} g^2(X_i) \tag{2.32}$$

This relies on N being independent; additionally as N increases, the error will decrease with the predicted value of X. This method, while useful, wastes time sampling areas where the posterior is very low and so contributes little to the final value. A better method (such as the one used by the Metropolis-Hastings Algorithm) is to only sample those areas where the posterior is high, because few samplings are required and it is therefore computationally less expensive to do.

In simple terms, the Metropolis-Hastings Algorithm starts at an arbitrary point and calculates its posterior, it then selects a new point in parameter space, calculating the new posterior and comparing the two. If the new point is preferential it moves, otherwise it remains where it is (and then loops again). From its initial point sample X_t, it selects for the new sample, X_{t+1} a new position Y, from the "proposal distribution" $q(Y \mid X_t)$. This "proposal distribution" can be modelled as a Gaussian, with a width of σ (the value of this can affect the sampling numbers required and, as a result, the accuracy of the convergence, as discussed later in this section). When modelled like this, Y will be

selected within this range from the X_t value. The algorithm then needs to decide whether to accept this value or to reject it and keep the current value. This is carried out using the "Metropolis Ratio":

$$r = \frac{P(Y \mid D, I)}{P(X_t \mid D, I)} \times \frac{q(X_t \mid Y)}{q(Y \mid X_t)} \tag{2.33}$$

The last fraction at the end of equation 2.33 is equal to one if the "proposal distribution" is symmetric. If $r \geq 1$, then $X_{t+1} = Y$. If not, then it randomly selects a value between 0 and 1. If this value is less than or equal to r then it again selects $X_{t+1} = Y$; otherwise $X_{t+1} = X_t$.

The selection of σ itself is important: If the value is too small then the steps will be very slow and the algorithm takes a much longer time to give a stable result. Too big and the algorithm will select new values of Y that will be further from X_t and may also be outside the area where the posterior is high and so will discard the new value Y, keeping the values for X_t. This results in the algorithm converging quickly, but not accurately. As a consequence there is a trade off as to how many samples are required and the accuracy of the values for the parameters. Consideration is also required as to the "burn-in". From the arbitrary start position, some number of the samples taken will be in low posterior areas as the algorithm searches for and travels towards high value areas. After a certain number of samples, any sampling taking place after that will be in higher value posterior areas. These values are ignored in order to not bias the final result. In an extreme example, if it takes 1000 samples for the parameter values to converge, but 100 samples are needed in order to "burn-in", these values will non-trivially contribute to any analysis of these samples and should be excluded.

In summary, the "Metropolis-Hastings algorithm" procedure is:

a. Initialize X_0; set $t = 0$.

b. Repeat:

(i) Obtain a new sample Y from $q(Y \mid X_t)$

(ii) Compute the Metropolis ratio ratio, r (equation 2.33).

(iii) Sample a uniform(0,1) random variable U

(iv) If $U \leq r$ set $X_{t+1} = Y$ otherwise set $X_{t+1} = X_t$

(v) Increment t

2.4.3 LensTool

LensTool relies on Bayesian statistics and the MCMC in order to optimize the free parameters to best reproduce the lensing signals. The method that LensTool utilises is set out in Jullo et al. (2007) and the following is based on this paper.

Two methods of reconstructing the mass distribution are non-parametric and parametric methods:

The non-parametric method produces a map divided into a grid of cells. As the method produces the final map, the cells will be assigned a certain amount of mass, possibly as a top hat across the entire cell, or a Gaussian for example. This is useful as non-parametric methods can fit asymmetric mass distributions, but the increase in parameters to fit (every cell) can lead to issues in computation as well. This can be offset by dynamically changing the cell size, i.e. dense regions have more cells per unit area than less dense regions. Subsequently it requires processing in order to produce comparable quantities, e.g. DM density slope, from the resulting pixel mass map.

The parametric method uses a parametrised model for the mass distribution and then fits and optimizes for these parameters. Because of the use of physically motivated models, some values such as dark matter density slope are given as inputs and so subsequently are given as direct outputs that do not need post-processing. The use of physically motivated mass models such as PIEMD or NFW (see § 2.5) to fit the lensing constraints also means that some parameters can be set at fixed values, provided they are physically motivated. Therefore fewer parameters need to be fitted with a subsequent reduction in computation time. LensTool uses this method.

Before starting, it is necessary to determine which parameters require parametising. These are generally smoother cluster-scale halos and generally include intra-cluster gas and DM, because the two are often overlayed and hard to separate. The type (PIEMD, NFW etc) and number of halos to select is important, but regardless of the halo choice, some of the basic parameters will be ellipticity, position angle and position on the sky. Then smaller subhalos representing galaxies can be added in. The measured position angle and ellipticity of the galaxies is used to find their subhalo angle and ellipticity. However, leaving the other values as free parameters would quickly provide more free parameters than constraints, whilst also increasing computational time. These galaxy subhalos generally only provide mass contributions towards the total mass as well as substructure fraction, but do not all contribute towards creating the multiple images seen. To reduce this problem, only those subhalos which may affect the strong lensing image constraints (either by altering their position, or even splitting the images) are left to be modelled for the cluster-scale potentials. The other subhalos' parameters are scaled with their luminosities utilising scaling relations.

After deciding on the parameters, constraints are then set. The positions of strong lensing multiple images are the strongest constraints. If more than one identifiable feature can be discerned in the multiple images, these can provide even further constraints (see §4 for an example of this). The redshift of these strong lensing images also provides a constraint (and, as discussed in previous sections, helps to break mass degeneracy problems), though this can be left as a free parameter for LensTool to estimate.

Identification of the location of critical lines will provide a position and an error which can be given as a model input for use as a weight for the models generated. Locations can often be identified as an area of dimming where merging image systems meet at a critical line and interface with one another. That is, if the models generated during the LensTool optimisation do not have critical lines that appear within the user defined area, they suffer a weighting that decreases their favouring over those models that do.

Whenever a model is produced, the model-generated images are compared to the

constraints and a χ^2 value calculated and used to discard or keep the models. The images can be matched in the image or the source plan. In the latter the source positions are matched for minimal scatter. Matching in the image plane is the more accurate of the two but computationally expensive because it requires an inversion of the lensing equation each time it is run. Using the source plane only requires a linearisation of the lensing equation and so is computationally easier to carry out. The lens equation (see § 2.1 equation 2.8) can be utilised to map the image positions onto the source plan so that the difference in the source positions can be calculated. When images are matched in the image-plane they are first mapped to the source plane and then LensTool maps these back to the image plane by inverting and solving the lens equation to return the predicted image positions for that particular source position. For complicated mass distributions this can take time when mapping is non-linear as the solutions cannot be analytically derived. In addition, the number of images is not always known as a prior which makes a numerical inversion non-trivial as well. There is also a delay caused by trying to match image positions to predicted image positions for poorly constrained models. LensTool is reasonably robust to this effect as the observed image is coupled to the predicted one, but where the models produce different configurations of images, as can happen early on in the LensTool optimization process, it can slow the convergence time considerably.

LensTool uses the MCMC method, which both explores the range of values for the parameters and compares the models that are produced as a result. It uses 10 inter-linked Markov chains. These minimize the chances of becoming locked in local minimum as it is unlikely all 10 will fall within the same trap. This method builds on the MCMC set out in §2.4.2 but uses an additional process called "selective annealing". At every repeat, ten new samples are drawn and assigned to one chain, using the generated posterior from the last sample run. The new samples are assessed in a similar way to that set out in §2.4.2, as to whether they should be accepted or discarded (instead of the old value being reused). The best samples are then duplicated to bring the chain number back up to ten. BAYESIS, which is integrated into LensTool, is then used to then randomly move the

samples in parameter space. This is the selective part of "selective annealing".

The annealing part is the addition of a cooling factor that affects the convergence speed. This is controlled by the "inverse" section of the LensTool parameter file. A lower number results in the annealing taking longer to perform, but the parameter space will be better explored after convergence when the "burn-in" phase is done. If the inverse is set high, it will rapidly converge but the posterior may not be well explored. The idea of the annealing is that if the likelihood between subsequent steps widens, then the rate of convergence will slow and *vice versa*. This means that for a slower and smoother convergence a better-characterised (and in an equilibrium state) sampled posterior will be returned for each parameter. This is due to increased sampling of parameter space. This posterior is then used for the real sampling. A lower inverse rate will slightly lower the mean χ^2 and also lower (more than for the χ^2) the error in the Bayesian evidence of the final models (Jullo et al., 2007).

After "burn-in", the MCMC will then draw samples to produce models, taking more samples from where the posterior is higher and less samples from where it is lower. The more samples drawn, the better the estimation of the parameter values.

2.5 Density Profiles of Halos

2.5.1 NFW Mass Model

This model was first proposed by Navarro, Frenk and White (NFW) in 1996 (see Navarro et al. (1996)). The model was based upon dark matter simulations, that suggested the possibility of a universal density profile, which is, in principle, testable. The density profile of a NFW halo is given by equation 2.34.

$$\frac{\rho(r)}{\rho_{crit}} = \frac{\delta_c}{(r/r_s)(1+r/r_s)^2} \tag{2.34}$$

$\rho(r)$ being the density at a radius of size r, r_s is the scale radius and δ_c is a dimensionless

density. ρ_{crit} is the critical density for a closed Universe and is defined in equation 2.35, where H is the Hubble constant and G is the gravitational constant.

$$\rho_{crit} = \frac{3H^2}{8\pi G} \tag{2.35}$$

The use of the NFW profile over other profiles such as an Einasto profile which J. Einasto first introduced at a conference in 1963 (Fesenko, 1963), is still being researched (Broadhurst et al., 2005). Weak lensing is not currently able to distinguish between the two due to degeneracies related to the density slope, making strong lensing a useful tester for profile preference if sufficient lensing constraints are available to constrain the density slope (Clowe et al., 2000). N-body simulations have provided limited ability to distinguish some profiles as preferable over others, but they are not yet at a sufficient resolution to enable one profile to be picked out as superior to all others. There is, however, some debate (see Sand et al. 2008) over whether the NFW density slope is really $\rho_{NFW} \sim r^{-2}$ or whether it varies from r^{-3} (for $r >> r_s$) to r^{-1} (for $r << r_s$), where r_s is a scale radius. To reflect this, many papers use equation 2.36 to describe a NFW and determine a value for α for the best fit slope (see Limousin et al., 2008 for an example). In its original proposed form, a NFW profile needs an $\alpha = 1$, which changes equation 2.34 to (Limousin et al., 2008):

$$\frac{\rho(r)}{\rho_{crit}} = \frac{\delta_c}{(r/r_s)^{\alpha}(1 + (r/r_s))^{3-\alpha}} \tag{2.36}$$

Equation 2.36 now reflects this alterable density slope. Other values of interest are r_{200} (the radius at which, the density inside is 200 times the average Universe density) and the concentration parameter, c, (see equation 2.37 taken from Navarro et al., 1997) which describes the central density of the halo.

$$c = \frac{r_{200}}{r_s} \tag{2.37}$$

The concentration value is of interest as it is anti-correlated to mass. As a result of its

expansion, the Universe is at a much lower density now than it used to be in the past, so massive clusters that formed relatively recently will have a low concentration value. Older clusters formed in a much denser Universe have less mass (high concentration value) and so the concentration values should provide information on the age of a cluster's formation, if not disrupted by later processes.

When the NFW profile was first suggested as a "universal" profile for the fitting of DM halos in a CDM Universe, it was found to provide a much better fit of the density slope from N-body simulations than the previously used isothermal spherical profiles. This was due to the observed gentle change in the density slope with radius, which (with a constant slope) isothermal profiles do not well approximate (Navarro et al., 1996)). This implies that NFW's should be better able to accurately reproduce strong lensing arcs than an isothermal profile. This can be seen in a study by Shu et al. (2008), where they found that for strong lensing arcs that are well produced by both profiles, the magnifications of the arcs can strongly differ. Shu et al. also modelled Abell 370 using both profiles and, while they can both reproduce the arcs, the critical curves and the resulting caustics from mapping onto the source plane, differ strongly as well.

As stated earlier in this section, the NFW profile is a lot flatter near the central core than an isothermal profile and, outside a certain core radius, it is a lot steeper than an isothermal profile. The concentration value for a halo was also found to be a good indicator of the period at which the halo itself had formed in time, with massive halos being more concentrated and less massive ones being the reverse. These comparisons were carried out by Navarro et al. (1996) utilising N-body simulations, whereby halos from different massed clusters were extracted and then re-simulated to the same resolution (to reduce bias, because otherwise larger halos would have more particles and so a higher resolution etc.).

The shape of NFW halos was also suggested by Navarro et al. (1997) from N-body simulations to be universal, despite the mass of the cluster or the cosmology that is being used. This renders it insensitive to these parameters. The phenomenon of concentration

increasing with smaller mass indicates that the low-mass cluster halos merged at a much higher redshift than larger massed ones. This link with the time at which the clusters formed means that different cosmologies can be used when modelling the DM halos of clusters without affecting the result. Thereby the same "proportionality constant" can be utilised between the density of the cluster and the density of the Universe when the cluster first collapsed (Navarro et al., 1997). Navarro et al. also stated that mergers act as a mechanism to produce an equilibrium of the dark matter in the halo that is independent of initial conditions, and that this mechanism must occur rapidly due to the similarity of relaxed NFW halos.

2.5.2 PIEMD Mass Model

The Pseudo-Isothermal Elliptical Mass Distribution (PIEMD) was first proposed in Kassiola & Kovner (1993), but has over time turned into a superposition of two PIEMDs (confusingly also called a PIEMD), resulting in equation 2.38. PIEMDs in their original form were proposed due to their ability to accurately represent objects across a wide range of ellipticities, which profiles such as Pseudoisothermal Elliptic Potentials (PIEP) could not do. The combination of two PIEMDs were found to provide a better fit to observations of cD galaxies than a single PIEMD on its own could (see Natarajan & Kneib, 1996).

$$\rho(r) = \frac{\rho_0 r_{core}^2 r_{cut}^2}{(r_{cut}^2 - r_{core}^2)}\left(\frac{1}{r^2 + r_{core}^2} - \frac{1}{r^2 + r_{cut}^2}\right) \tag{2.38}$$

Where r_{core} is the radius at which the inner slope changes (becomes much steeper), r_{cut} is the truncation radius and ρ_0 is the central density (this will be larger than r_{core}) (taken from Natarajan & Kneib, 1996).

The values of interest when using PIEMD models are r_{core} (the radius above and below which the density slope will be steeper and flatter respectively) and the velocity dispersion σ_{PIEMD} (given in equation 2.39). The velocity dispersion is also related to the

deflection angle α given in § 2.2.1 equation 2.5 (Elíasdóttir et al., 2007), in that a higher velocity dispersion translates to a greater mass in the cluster in order to constrain it and *vice versa*, which is shown below:

$$\sigma^2_{PIEMD} = \frac{4}{3} G\pi\rho_0 \frac{r^2_{core} r^3_{cut}}{(r_{cut} - r_{core})(r_{cut} + r_{core})} \tag{2.39}$$

Equation 2.39 relates the three-dimensional mass of the cluster to the velocity dispersion. PIEMDs can provide a good approximation to the same profile shape that a NFW produces, although there is debate over which is better. The motivation for using PIEMDs came from the fact that X-ray modelling of clusters used to be modelled isothermally; a similar parametrisation was then applied to lens modelling. Other profiles such as a Singular Isothermal Sphere (SIS) are not good fits to observations, however, as the density profile is too steep near the centre and produces too many images.

CHAPTER 3

OPTICAL SPECTROSCOPY

This chapter contains contributed work from other authors as outlined in the statement of originality and is additionally utilised in subsequent chapters. X-ray data reduction was performed by Alaistair Sanderson and the Infra-red and HST data by Victoria Hamilton-Morris, with the information from the reduction method description also being contributed from the same via discussion. For the strong lensing models all redshifts for arcs were determined by the author except for A 2537 (the author reduced and measured more precise redshifts subsequently) and the P-cygni profile fitting for the redshifts of A 368, which were provided by Johan Richard. The image positions for A 2537 were also provided by Johan Richard. Otherwise the redshift determinations and reduction, strong lens modelling, calculations, statistics and plots were the work of the author unless labelled otherwise or as set out in the statement of originality at the beginning of this thesis.

The aim of this chapter was to show the methodology and results for the reduction of spectroscopic data. This was the vital starting block required in order to produce the other chapters which each build on the work described in the previous chapter. This chapter starts with an explanation of the observational details, reduction method, results and then comparisons with other observables; do they agree or disagree? As each cluster is different, for clarity, a discussion is included after each set of observational details. This is followed by a discussion of the results and finally a summary.

From the LoCuSS observation plans, large amounts of optical spectroscopic data were

taken on X-ray selected galaxy clusters. From this data set, 8 clusters were chosen because they offered the possibility of measuring the redshifts of strong lensing arcs for mass determinations and, because the acquisition of unreduced spectroscopic data from the VLT and Gemini-South telescopes, could enable arc redshifts to be determined for use in strong lensing modelling. On average a new object redshift was measured for over two thirds of the slits in each mask used, with faint lensing images overlapped across several of the masks in the hope of increasing the signal to noise of their faint spectrum when combined. In many cases the signal to noise was still too low to make a confident detection.

3.1 Spectroscopic Observations

3.1.1 Reduction Method

With an accurate redshift of a galaxy group and/or of a lensed image it was possible to put constraints on the gravitational potential causing the lensing (see § 2.1). When modelled using LensTool, this resulted in a better estimation of the most likely values for parameters under investigating. Without accurate redshifts, an estimate of redshift could be made using the colour of the object (photometric redshifts, which were not used for this work). This utilised observations of the object object using a range of filters, that are used to look for the drop in its luminosity between filters to detect the 4000 Å break. However, this was not an accurate method and the best way was to use a spectroscopic redshift. The 4000 Å break is usually seen as a sharp dip in the continuum of the spectrum, with the bluer side at a much lower level than the redder one. This break was due to the blanket absorption of high energy photons by metals in the atmospheres of stars and by a lack of hot blue stars. The output of blue wavelengths in the stars making up cluster galaxies was low as they were old red stars. The break has two very strong absorption lines of Calcium CaII, often called H (3969 Å) and K (4303 Å).

A "mask" was required to take spectroscopic observations. This was a very simple metal sheet that had holes cut out of it (referred to as slits). When orientated on the sky correctly the slits were coincident with the objects to be observed. The mask was then placed in front of the telescope used to take the observations. With only the light from the slits entering the telescope, a filter was used to select a wavelength range of interest. This was then passed through a grism, which was a combination of a grating and a prism. The result of which was a spectrum with a central wavelength with the spectrum dispersed around this central point in one axis (this allows observations with a specific wavelength range). The spectrum then fell onto the telescope's charge-coupled devices and could be read. The use of a mask prevented the spectrum of objects close to one another becoming overlapped and the spectrum blended together. If this happened the spectrum's would have been incredibly difficult to separate. A side effect is that the mask slits could not be overlapped in the grism dispersion direction, so consideration had to be made of the priority of slit assignment to objects and orientation of the mask on the sky.

As the raw data contained hot pixels, biased pixels, cosmic rays etc., it had to be reduced, starting with combination of the multiple observation images for each pointing. When combining these images into one main image, additional operations were required. These operations sometimes differed in their reduction depending on the source telescope. The two methods used in this work are listed in the following two sections.

3.1.2 GMOS

For the GMOS data a set of tasks were available that were performed in the IRAF environment. The observatory also provided a standard "cookbook" but, following discussion with the local GMOS experts in Hawaii, this was altered slightly to include image combination and sky ray rejection tasks that were required in this thesis. These tasks were performed in succession with the output of one then used in the subsequent task. A change of order of work can be undertaken with some of these tasks and some tasks required certain actions to be undertaken before they could be used correctly. For example,

the slit positioning information appended to a file before a task could isolate and cut out the object's spectrum.

The reduction started by providing the GPREPARE task with a list of initial files. This task attached the Mask Definition File (mdf) to each of the files. This file contained the information about slit sizes and positions, which aided the GMOS IRAF tasks that were to be undertaken subsequently. Because each of the observations were split into multiple observations (some of the objects were very faint), combining the observations at this point was performed for the ease of the automated tasks. GEMCOMBINE was the task used as this also had the additional advantage of performing sky ray rejections as it looked at differences in each observation and rejected large deviances as it constructed a single combined fits file. An alternative method of cosmic ray rejection could also be performed using GSCRREJ, which looked at each pixel in turn and compared it to a fit of the pixels in a box around the targeted pixel, subtracting this fit and then searched for large residuals. Any such aberrant pixels were substituted with the fitted value.

The flat and bias images needed to be prepared before they could be used. Any bias images were combined into one bias image using GBIAS. GSFLAT taking a gprepared flat file (or list of files) and generating a normalized fits file where only the pixel to pixel variations and fringing were left.

With all the relevant files reduced, GSREDUCE was now called. This subtracts the bias fits file and used the flat fits file to normalize the pixel response across the Charged Coupling Devices (CCD). It also trimmed the images into separate extensions in the fits file using the mdf information on the slit positions. Using the header information on grism and filter selections it also performed a rough wavelength calibration on the data, by examining the grating and central wavelengths used for the observations, and appended this information to the fits file.

For a more detailed wavelength calibration the arc fits images (the observations taken of the arc lamp calibration spectrum using the identical set up used to take observations) were used with GSWAVELENGTH. It examined the input spectrum and attempted to

identify the emission lines and thus wavelength calibrated the spectrum. This could be set to be automated, but it could also be done interactively to correct any misidentifications or double check identifications. With this performed on the arc fits file, the results were fed into GSTRANSFORM in order to wavelength calibrate all the science fits files.

With wavelength calibrated, sky subtraction was performed using GSSKYSUB. This was best performed interactively, whereby a wavelength slice (or range of slices, which were then averaged) was selected and a sample range chosen (in the slit height direction) from the displayed image. This sample was used to fit a sky background fit and this resulting fit then subtracted from the image. This sample can be a different selection size and position for each slit in the mask.

At this stage the science images were cut into separate slits as extensions of the fits file. They had been normalized using the flat field, bias subtracted, wavelength calibrated, sky background subtracted and cosmic ray rejection performed. This left the final reduction step of extracting the object spectrum (which was performed by GSEXTRACT). GSEXTRACT automatically attempted identification and extraction of object spectrum, but was best performed interactively. The spectrum was collapsed in a wavelength direction and, using the mdf information as a guide in object positioning, the task selected the feature it calculated to be the object, however, this could be overwritten. It was then possible for the user to specify whether the task should trace the spectrum or not. Tracing involved the task starting from the initial position selected, and then as it travelled along the wavelength axis direction, selected the strongest emission pixel nearby (i.e. giving the strongest signal to noise spectrum from the data). The results of this wander in slit height direction are output. For a strong spectrum the wander would be minimal (over about 5 to 10 pixels). For a very weak spectrum tracing was undesirable as large wander could occur due to the tracing algorithm selecting strong pixels of noise rather than the true object spectrum pixels. In this weak spectrum case, tracing was disabled and the spectrum was selected along the selected column height. The resulting extracted spectrum were output to a multi-extension fits file, with each extension having a labelled

and extracted one-dimensional object.

3.1.3 FORS2

The basic fits images fell into the following categories of bias, wave, science and flat frames for each set of observations. These data were reduced in a standard manner using Kelson's (2003) python scripts.

The first task was to normalise the frames by using the BIAS and FLAT fits frames. The BIAS frame gave the level of noise in the detector and, by applying this to the FLAT frame, returned a normalised flat frame and so on with the other frames. A blazing was also applied to the flat field. This was simply a laplacian curve fitted to the shape of the flat field. This fitting was used to define the slit edges (which could be altered by the user) and these were then copied to the fits file headers. Dividing the science frames against this flat field, which was simply a unity normalised quantum efficiency across all the CCD chips, removed the quantum efficiency variations across the CCDs.

A line list was constructed for each combination of grism and filter. This was required to enable the wavelength to be rectified for each slit. This consisted of inputting an arc lamp spectral image, identifying the lines and then constructing a list of the wavelengths and amplitudes corresponding to the spectroscopic equipment set up and arc lamp used. This line list determined the wavelength range of the slit and was used to calculate a wavelength solution derived for each slit. Using a similar method as set out in the previous GMOS reduction section, a fit was applied across the slits' height so that it could be subtracted across the slit to remove the sky background. In its standard form it would do this across the entire height (usually fine for faint object spectrum), but for bright spectral objects a range of sky across the slits height had to be selected so that the object was ignored and would not bias the sky fit.

With the science frames at this point, the script compared the chipped mean across the multiple exposures to remove any cosmic rays. The chipped mean was the average value of the same pixel on the CCD across all exposures. It then flux calibrated the

frames so that the spectrum was ready to be extracted. A list was prepared of the names for each of the slits and the position of the object's spectrum and then each spectrum for that object was sliced out into a separate image. These images contained several one-dimensional spectrum (sky spectrum, object spectrum etc.). The object spectrum can then be extracted out of these images and the redshift calculated for the object by identifying and comparing spectrum against a list of spectral lines.

3.1.4 Observations

Two sets of observations were taken for each cluster. One was observed with the FOcal Reducer and low dispersion Spectrograph 2 (FORS2; Appenzeller et al., 1998) on ESO's 8.2-m Very Large Telescope (VLT), through a single multi-slit mask. Slit placements were prioritised for each mask targeting potential multiple images and then on candidate cluster galaxies. The majority of observations were performed with the 300V grism (a grism determines the central wavelength of light let through and how much it spreads the light out) so as to maximize the observed wavelength range ($3500 \lessapprox \lambda_{\rm obs} \lessapprox 9000$ Å), but other grisms were also used. Table 3.1 lists the total integration time, along with the total number of exposures, filter, grism, airmass, seeing and the number of slits per mask (the same mask was used for the different combinations of grism and filter, so the number of slits were constant). These data were reduced in a standard manner using Kelson's (2003) python scripts, following the method described in §3.1.3.

The second set of observations were acquired (again using masks to pick out objects) with the Gemini Multi-Object Spectrograph (GMOS) mounted on the 8.1-m Gemini South telescope[1] using the $GG455\_G0329/R400+\_G5325$ filter/grating combination. These data were reduced using the GMOS package in IRAF following the standard reduction cookbook,

[1]Based on observations obtained at the Gemini Observatory, which was operated by the Association of Universities for Research in Astronomy, Inc., under a cooperative agreement with the NSF on behalf of the Gemini partnership: the National Science Foundation (United States), the Science and Technology Facilities Council (United Kingdom), the National Research Council (Canada), CONICYT (Chile), the Australian Research Council (Australia), Ministério da Ciência, Tecnologia e Inovação (Brazil) and Ministerio de Ciencia, Tecnología e Innovación Productiva (Argentina).

Table 3.1: FORS2 Observing Data

Name	Observing Date	Exposure Time	Grism	Filter	Airmass	Seeing	No. of Slits
Abell 2813	Aug 30, 2008	900x3	300V	GG375	1.015	0.91	43
	Aug 30, 2008	900x4	600I	OG590	1.012	0.73	
Abell 2895	no FORS2 observations were made						
Abell 368	Aug 31, 2008	900x2	300V	GG375	1.005	0.84	34
	Aug 31, 2008	900x3	600z	OG590	1.025	1.03	
Abell 3084	Aug 30, 2008	900x3	300V	GG375	1.032	1.04	41
Abell 3088	no FORS2 observations were made						
RXCJ 0528.2	no FORS2 observations were made						
Abell 3364	Aug 31, 2008	900x2	300V	GG375	1.284	1.13	36
Abell 2537	Aug 30, 2008	900x4	300V	GG375	1.163	0.86	43
	Aug 30, 2008	900x4	600I	OG590	1.087	0.73	

with additional steps for cosmic ray rejection and data combination, described in more detail in §3.1.2. The observational data are listed in Table 3.2.

The number of new object redshifts (z_{obs}) from these two data sources are shown in Table 3.3. On average $\sim 65\%$ of slits yielded a confident redshift for each mask, with six strong lensing image spectrums also found (this also enabled two sets of two images to be matched, yielding four separate source object redshifts).

3.2 Redshift Histograms

For each cluster any FORS2 or GMOS observations were reduced and redshifts measured. By examining the distribution in redshift space, constraints were put on the possible galaxy members of each cluster. Additional galaxy redshifts were provided from the REFLEX catalogue (see §3.2.1). With this combined redshift list, the resulting constrained cluster members were highlighted in each histogram plot (shown in the relevant cluster section in §3.7) in red (with the outliers shown in black) for each of the clusters. Dotted lines showed the selected constraint boundaries and a value for the mean redshift along with its one σ error in the top right of each histogram plot. Each plot was plotted in the range of $0.1 < z < 0.4$ for easy comparison, with the bin sizes fixed to 60 bins.

The redshift range within which a galaxy was included for cluster membership were based on an assumed velocity dispersion of each cluster being $\sigma \approx 1000$ km s^{-1}. This was near the top end value for 81 ACO clusters taken from Struble & Rood (1999) and near the largest velocity dispersion recorded for the 10 LoCuSS derived strong lensing clusters in Richard et al. (2010). This velocity dispersion value was used due to the low numbers of redshifts: low numbers might mean members appear unassociated due to gaps in redshift space. Then a 6σ ($z_{diff} \pm 0.02$) cut was applied to the centre of the clumping of redshifts in redshift space. Even if it is not assumed that the data has a normal distribution, Chebyshev's Inequality states that for a 6σ interval it should contain $\geq 97\%$ of the cluster galaxy sample, assuming the mean and σ were both true. The alternative

Table 3.2: GMOS Observing Data

Name	Mask[a]	Observing Date	Length of Time	Filter	GRISM	Mean Airmass	Seeing	No. Slits
Abell 2813	1	Aug 30, 2008	1200×6	GG455_G0329	R400+_G5325	1.157	≥ 0.75	23
	2	Aug 29, 2008	1200×6	GG455_G0329	R400+_G5325	1.125	≥ 0.75	20
Abell 2895	1	Oct 20, 2008	1200×6	GG455_G0329	R400+_G5325	1.002	≥ 1.05	27
	2	Sep 21, 2009	1200×6	GG455_G0329	R400+_G5325	1.036	≥ 0.45	25
Abell 368	3	Dec 31, 2008	1200×6	GG455_G0329	R400+_G5325	1.068	≥ 0.75	23
	4	Sep 14, 2009	1200×6	GG455_G0329	R400+_G5325	1.036	≥ 1.05	21
Abell 3084	5	Dec 25, 2008	1200×6	GG455_G0329	R400+_G5325	1.072	≥ 0.75	26
	6	Dec 28, 2008	1200×6	GG455_G0329	R400+_G5325	1.262	≥ 0.75	25
Abell 3088	7	Dec 21, 2008	1200×6	GG455_G0329	R400+_G5325	1.124	≥ 0.75	22
	8	Nov 27, 2008	1200×6	GG455_G0329	R400+_G5325	1.020	≥ 0.75	19
RXCJ 0528.2	5	Nov 25, 2008	1200×6	GG455_G0329	R400+_G5325	1.072	≥ 0.75	26
	6	Dec 2, 2008	1200×6	GG455_G0329	R400+_G5325	1.030	≥ 0.75	18
Abell 3364	9	Nov 24, 2008	1200×6	GG455_G0329	R400+_G5325	1.145	≥ 0.75	24
	10	Dec 20, 2008	1200×6	GG455_G0329	R400+_G5325	1.006	≥ 0.75	27
Abell 2537	11	Aug 8, 2008	1200×6	GG455_G0329	R400+_G5325	1.432	≥ 0.75	24
	12	Aug 31, 2008	1200×6	GG455_G0329	R400+_G5325	1.154	≥ 0.45	20
	13	Sep 24, 2008	1200×6	GG455_G0329	R400+_G5325	1.826	≥ 0.75	25
	14	Sep 25, 2008	1200×6	GG455_G0329	R400+_G5325	1.176	≥ 0.75	19

[a] This labelling system was simply listed for consistency with the mask labels the GMOS system assigned the masks.

Table 3.3: New object redshift yields

Name	GMOS slit No.	FORS2 slit No.	Total slit No.	z_{obs}[a]	Min Yield
Abell 2813	43	43	86	71	83%
Abell 2895	52	-	52	38	73%
Abell 368	44	34	78	47	60%
Abell 3084	51	41	92	65	71%
Abell 3088	41	-	41	29	71%
RXCJ 0528.2	44	-	44	25	57%
Abell 3364	51	36	87	43	49%
Abell 2537	88	43	131	69	53%

[a] Total number of redshifts measured across both GMOS and FORS masks.

method for selecting was by an iterative method. In this method a Gaussian was fitted to the data and then any galaxies outside the fitted 3σ range were discarded and the Gaussian refitted until the results were stable. The results were virtually identical, with only one or two galaxies different, to those of the prior method. This highlighted the sparsity of redshifts and the incompleteness in redshift space.

The results for the sample of the determined cluster redshift from this analysis are shown in Table 3.4, along with the cluster position, the literature BCG redshift, the redshift determined from the histograms and the number of members that were constrained by the histograms.

3.2.1 Reflex Additions

The priority given in the centre of each cluster for slits to be placed on potential multiple image systems left a deficit of possible galaxy members with measured redshifts near the cluster centre. This deficit could cause inaccuracies due to incompleteness but, for six of these clusters, their BCG redshift values and positions were listed in Guzzo et al. (2009), which used redshifts from the **ROSAT − ESO**[1] Flux-Limited X-ray (REFLEX) catalog. For several of these clusters more than one redshift has been measured near the BCG, which may fall within the redshift space constraints to be cluster members (see

[1]Based on data from the European Southern Observatory (ESO), La Silla, Chile

Table 3.4: Redshift Galaxy Cluster Sample

Name	α, δ [J2000]	z_{lit} [a]	z_{hist}	No. of Members + literature	σ_{gal} [$\mathrm{km\,s^{-1}}$]	σ_{X-ray} [e] [$\mathrm{km\,s^{-1}}$]
Abell 2813	10.854791 −20.616777	0.3004 ± 0.0003	0.2933 ± 0.0068	$38 + 8$	2010 ± 270	1044 ± 92
Abell 2895	19.546146 −26.96999	0.2310[b]	0.2244 ± 0.0044	$19 + 1$	1290 ± 150	1201 ± 147
Abell 368	39.3656048 −26.5079624	0.2219 ± 0.0003	0.2208 ± 0.0050	$25 + 11$	1440 ± 300	1053 ± 150
Abell 3084	46.016386 −36.940826	0.2177 ± 0.0002	0.2162 ± 0.0044	$37 + 2$	1290 ± 90	796 ± 95
Abell 3088	46.758535 −28.6658	0.2527 ± 0.0002	0.2515 ± 0.0058	$13 + 9$	1680 ± 270	1166 ± 177
RXCJ 0528.2	82.062728 −29.717505	0.1535 ± 0.0002	0.1548 ± 0.0035	$15 + 2$	990 ± 180	1160 ± 130
Abell 3364	86.907104 −31.873212	0.3693 ± 0.0002[c]	0.1484 ± 0.0038	$31 + 5$	1110 ± 120	1059 ± 101
Abell 2537	347.09294 −2.1925704	0.2950[d]	0.2960 ± 0.0056	$50 + 1$	1650 ± 180	787 ± 73

[a] Unless otherwise cited, all redshifts are from Guzzo et al. (2009).
[b] Redshift from Dahle et al. (2002).
[c] BCG redshift was exempted from analysis. See §3.7.7.
[d] Redshift from Cruddace et al. (2002).
[e] Calculated using eqn. 3.1 from Xue & Wu (2000).

§3.2). By adding in these extra redshifts, an improvement on the cluster redshift average error, lowering of basic cluster velocity dispersion and changes to the Dressler Shectman substructure plots might or might not be expected. A table of redshifts and positions for each galaxy member are listed in each cluster's subsection. Extra galaxies provided from the REFLEX catalog are also listed in these tables. Table 3.4 lists how many extra galaxy member redshifts were added onto the cluster galaxy member lists from these additions.

3.2.2 Velocity dispersion vs T_{X-ray}

The cluster galaxy redshifts should provide a basic estimation of velocity dispersion measurement for the cluster. By taking 1000 bootstrap samplings (with replacement) of the cluster samples a basic velocity dispersion and error was made for each cluster. This value was expected to have a bias from the selection method, because including more galaxies that were not part of the sample would affect any subsequent calculations. A sensible check to determine the accuracy of this velocity dispersion was to compare it to a relation linked with another observable. Any large aberrations would be indicative of the simple calculation or statistical errors being an issue. The advantage of this check was that it was simple to perform and it was essentially free data that had come as a result of reducing data in preparation for strong lensing analysis.

Work by Xue & Wu (2000) produced a velocity dispersion and X-ray temperature relation of the order $M \propto T^{\alpha}$. It showed a slightly different slope when fitted to groups or clusters, but for this comparison Xue & Wu's fit to clusters was used (see equation 3.1).

$$\sigma = 10^{2.49\pm0.02} \times T_X^{0.65\pm0.03} \tag{3.1}$$

With σ in $\mathrm{km\,s^{-1}}$ and T_X in keV. The resulting calculated velocity dispersions derived from the cluster X-ray temperatures against the estimated velocity dispersion are shown in Table 3.4, with a plot (see Fig. 3.1) of the velocity dispersion derived from the galaxy cluster members σ_{gal} plotted against the velocity dispersion σ_{X-ray}, calculated

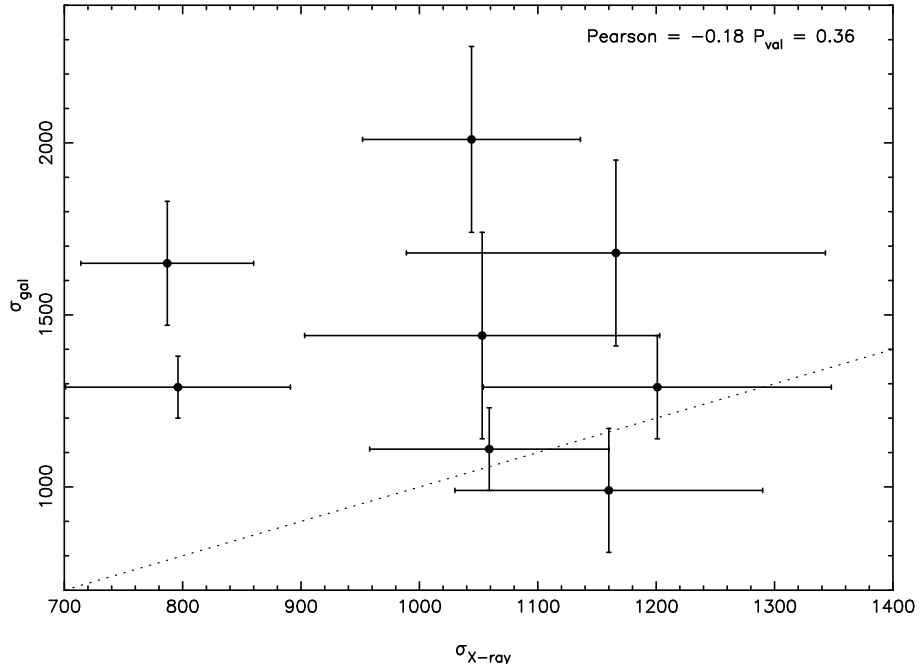


Figure 3.1: The galaxy cluster member derived velocity dispersion σ_{gal} plotted against the velocity σ_{X-ray} derived X-ray temperature relation for Xue & Wu (2000). The dotted line shows the one to one relation between the two axes, highlighting the systematically higher velocity dispersions on the y-axis.

using equation 3.1 and the cluster's X-ray temperature.

The results showed that the velocity dispersions derived from the galaxy cluster members were all much higher than the velocity dispersions derived from their X-ray temperatures, which suggested that the method of deriving σ_{gal} was not sufficient or that the lack of redshifts in the cores of the clusters caused a bias to occur. Fig. 3.1 shows a weak (see Table 5.4, Evans 1996) negative correlation with the returned Pearson value being 0.18 and a probability of 35% that this was due to sampling rather than the population. Since a positive correlation was expected, this again enforced the assumption that the calculation of σ_{gal} was not a true representation of the galaxies velocity dispersions or that the clusters themselves might not be relaxed, with the clusters velocities representing this relaxed mass distribution. The cluster galaxies had been disturbed and were no longer tracing the underlying mass distribution accurately and, therefore, a revised method for this calculation would be suggested as future work. The idea that not all the clusters were relaxed carried over over into utilising Dressler Shectman (see § 3.3) tests to search

for signs of substructure that could indicate the cluster was not relaxed (i.e. virialised).

3.3 Dressler Shectman

The fact that galaxy cluster members were in a potential well and moving in relation to that centre created a velocity offset (for each galaxy) in redshift space, which for a large enough sample should be evenly distributed around the cluster redshift. Any substructure in a galaxy cluster would cause this distribution to break down (for large enough substructure and galaxy velocity sampling, this would show up as superimposed peaks in a histogram of redshift space). The Dressler Shectman test uses velocity and position information for each of the cluster galaxy members. This information was used to look for variances with position of the velocity dispersion, which could indicate possible substructure.

The original testing method was first given in Dressler & Shectman (1988) and utilized a comparison of the global velocity mean and dispersion values against a local velocity mean and dispersion value for each individual galaxy. These local velocity values were calculated by incorporating the selected galaxy and the ten galaxies nearest (for a total of eleven).

This deviation from from the global values was defined as δ and was calculated from the equation given in Dressler & Shectman (1988):

$$\delta^2 = \left(\frac{11}{\sigma^2_{global}}\right)\left[\left(\bar{\nu}_{local} - \bar{\nu}_{global}\right)^2 + \left(\sigma_{local} - \sigma_{global}\right)^2\right] \tag{3.2}$$

The $\bar{\nu}$ and σ were the mean velocity and dispersion respectively for the local (ten nearest neighbours) and global (all cluster members). This equation used the ten nearest neighbours as its calculation value, which was acceptable for large numbers of galaxy cluster members but when the total sample got smaller the deviations from the global cluster kinematics was drowned out by adding in galaxies that were not deviated. To counteract this the equation can be altered to (Pinkney et al., 1996):

$$\delta_i^2 = \left(\frac{n_{nn}+1}{\sigma^2_{global}}\right)\left[\left(\bar{\nu}^i_{local} - \bar{\nu}_{global}\right)^2 + \left(\sigma^i_{local} - \sigma_{global}\right)^2\right] \tag{3.3}$$

Where $n_{nn} = \sqrt{N}$ (where N was the total sample size of galaxies used) following previous methodologies (Hou et al., 2012) and i ran from 1 to N. The resulting δ_i^2 for each galaxy cluster member was used to produce a Dressler Shectman plot with the positions of the galaxies plotted with circles of radii scaling as:

$$r \propto e^{\delta} \tag{3.4}$$

This produced plots with circles of varying radii often clustered together, giving indication of the presence and position of substructure.

3.3.1 DS Statistics

The produced plots had visual indicators of substructure, but a statistical value of substructure was obtained called the $\Delta - value$, which was defined as:

$$\Delta = \sum_{i=1}^{N} \delta_i \tag{3.5}$$

This value should yield $\Delta/N \approx 1$ for a cluster with a Gaussian velocity distribution (this come from N-body simulations for relaxed collisionless systems, Merrall & Henriksen 2003) that only contained random fluctuations. Any values above this indicated substructure. The standard for substructure was (see Dressler & Shectman 1988 and Hou et al., 2012):

$$\frac{\Delta}{N} > 1.0 \tag{3.6}$$

This was the critical values method for determining substructure. However, the sample of galaxies for a cluster might not accurately trace the underlying population and so would give values that were a product of the sample and not the population. This could yield

false values of Δ detecting substructure even if the cluster had no substructure. Therefore, in order to yield a useful statistic for comparison, Monte-Carlo shuffling for each sample was necessary. This was achieved by conserving galaxy positions, but shuffling (randomly assigning from the sample without replacement) the velocities and recalculating δ^2 and therefore the resulting Δ/N. This was repeated many times (in this thesis $10,000$ times in order to produce a well distributed and smooth histogram of Δ/N) and then compared to the original value over N shuffles to result in a probability that Δ/N was produced by chance. This latter method was the $P-value$ method and is described in greater detail below.

The $P-value$ method calculates the probability that the observed Δ value was a random occurrence of the sampling and was not an indicator of cluster substructure (see above). The $P-value$ is calculated using (see Hou et al., 2012):

$$P = \sum \frac{(\Delta_{shuffled} > \Delta_{observed})}{n_{shuffle}} \tag{3.7}$$

This uses Monte-Carlo shuffling (a Fisher-Yates algorithm was used to provide unbiased shuffling) as described in the critical value method (see above). The galaxy velocities were randomly assigned to each galaxy position and each shuffled sample had a calculated $\Delta_{shuffled}$ which was compared to the unaltered sample $\Delta_{observed}$. If it exceeded the observed value, the P value was incremented. At the end of the shuffling, the P value was normalised by dividing by the number of sample shuffles performed $n_{shuffle}$ ($n_{shuffle} = 10000$). As this sampled the distribution of Δ, the lower the P value, the more substructure was present. This was due to the small statistical chance that $\Delta_{observed}$ was due to a fortuitous sampling of the population. For an undisturbed cluster with a low substructure the reverse was true and so a large value of P resulted.

This statistic was useful as it required no prior information as to whether or not there was substructure or where it might be, but the two measurements were affected both by the sampling rate (nearest number of neighbours chosen) and also the number of galaxies associated with the cluster. As these changed the calculated statistics would also change.

Knebe & Müller (2000) found that for $n_{members} \geq 100$ the boundary between clusters with and without substructure was distinct even for very large values of N_{nn} (which desensitized the equation to small scale perturbations and lessened the effect of outliers).

Investigations of the stability of substructure measurements from either of these values (Hou et al., 2012) found that for low numbers of cluster members ($n_{nn} \approx 20$), a high level of false positives occured when measuring Δ/N (substructure where there was none). $P - Value$ was generally preferred as a measure of substructure, because it had a lower occurrence of false positives if selecting an appropriately low enough value ($P - value \leq 0.01$ minimised the false positive rate to 5%) to indicate substructure. To demonstrate this, for $\Delta/N > 1$ the false rate was found to be 81%, but for the $P - Value$, the rates were 5%, 10% and 15% for $P = 0.01$, $P = 0.05$ and $P = 0.10$.

3.4 Dressler Shectman Analysis

The main calculation performed with these redshifts was the Dressler Shectman (DS) test and the resulting statistics used for further analyses. A detailed description of the method for calculating the Dressler Shectman statistics was given in §3.3. Because of the low number of cluster member galaxies for some of the clusters, the number of nearest neighbours selected for calculating δ_i^2 were set by $n_{nn} = \sqrt{N}$ (see Table 3.5). This value was then used to calculate and produce a radius, which was normalized to be constrained within the field of view of the figure. This radius was used to plot a circle on a plot of each circles distance from the BCG centre. The larger the circle was, from the average group radius, signified a deviation from the global values and was a possible sign of substructure. Radii that were all approximately equal generally signified that the cluster had no significant amounts or substructure, no dynamical activity and were most likely to be well relaxed.

For each cluster the Dressler Shectman statistic Δ/N and P values and a Jackknife error for each were calculated and shown in Table 3.5, giving an estimation for the exis-

Table 3.5: Dressler Shectman Results

Name	n_{nn}[b]	Δ/N[c]	P[a]	Offset kpc	T_X	No. Halo[d]
Abell 2813	7	1.29 ± 0.41	$0.09^{+0.35}_{-0.09}$	116.45 ± 23.29	$6.51^{+0.89}_{-0.71}$	2
Abell 2895	4	0.89 ± 0.29	$0.76^{+0.24}_{-0.54}$	45.06 ± 9.01	$8.08^{+1.72}_{-1.18}$	1
Abell 368	6	1.58 ± 0.26	$0.04^{+0.08}_{-0.04}$	7.73 ± 1.55	$6.59^{+1.71}_{-1.09}$	1
Abell 3084	6	0.95 ± 0.21	$0.45^{+0.51}_{-0.45}$	46.09 ± 9.22	$4.29^{+0.91}_{-0.59}$	2
Abell 3088	5	1.03 ± 0.28	$0.61^{+0.39}_{-0.45}$	14.63 ± 2.92	$7.71^{+2.19}_{-1.31}$	-
Abell 3364	6	0.96 ± 0.32	$0.58^{+0.42}_{-0.58}$	40.94 ± 8.19	$7.65^{+1.45}_{-1.05}$	-
Abell 2537	7	1.44 ± 0.29	$0.01^{+0.05}_{-0.01}$	25.25 ± 5.05	$6.65^{+0.95}_{-0.85}$	2
RXCJ 0528.2	4	0.92 ± 0.47	$0.72^{+0.28}_{-0.72}$	31.32 ± 6.26	$4.21^{+0.59}_{-0.51}$	-

[a] Each test used 10,000 Monte-Carlo Shuffles to return a value plus jackknife error.
[b] n_{nn} states the number of nearest neighbours used in the DS calculations.
[c] Error was calculated using jackknife estimations.
[c] No. Halo = The number of DM halos used in strong lensing analyses. Clusters without SL constraints are marked with a "-".

tence of possible cluster substructure. The first observation was the high errors calculated for the P values, which were found in all but two cases of the clusters (A 368 and A 2537) with smaller errors calculated for the Δ/N values. The P values often spanned the whole probability range of 0 to 1. As a result of the high errors of the P values (see discussion of possible reasons for this in § 3.8), only Δ/N value was plotted against BCG - X-ray centroid offset, X-ray temperature and the number of large dark matter halos (produced and discussed in chapter 5) to show the values compared in §3.7. For each of these plots a Pearson correlation value and Null hypothesis probability (indicating the probability the correlation was due to the sample and not population) was indicated. The results of these plots are discussed in § 3.8.

3.5 Comparative Observations - *HST* & X-ray

To aid the cluster analysis, X-ray centroid offset was also compared to the other multi-wavelength data on each of these clusters (see Table 3.5). Additionally the X-ray data (from observations taken with *CHANDRA*) were used to construct surface brightness contours over-plotted onto the optical data from *HST* (Hubble Space Telescope). The

Table 3.6: HST Observing Data

Name	Observing Date	Obs Time (seconds)	Filter
Abell 2813	Sep 13, 2006	1200	F606W
Abell 2895	Aug 23, 2006	1200	F606W
Abell 368	Sep 14, 2006	1200	F606W
Abell 3084	Sep 12, 2006	1200	F606W
Abell 3088	Sep 12, 2006	1200	F606W
RXCJ 0528.2	Aug 15, 2006	1200	F606W
Abell 3364	Aug 15, 2006	1200	F606W
Abell 2537[a]	Oct 22, 2002	2080	F606W

[a] For this cluster the Observation ID is 9270, for all the rest it is 10881.

methods for their reduction are listed in chapter 4 as this chapter (3) is focused on redshifts of the clusters and their relation to observables. The X-ray surface brightness contours were overlayed onto the *HST* frames, with the X-ray centroid marked and its offset from the BCG centre listed in Table 3.5. Any confirmed multiple image systems were also marked on top of the combined *HST* and *CHANDRA*. Multiple image systems with measured redshifts were also listed for the relevant cluster in §3.7. In all produced plots showing *HST* data, North was up, East was left and each tick mark signified 10 arcseconds. The BCG position in each plot was indicated by the intersection of the dotted lines. Observing date, length and filter are listed in Table 3.6.

3.6 Data Summary

The work so far can be summarized as follows:

(i) 8 clusters were selected from the LoCuSS survey sample due to strong lensing features, along with unreduced spectroscopic data and as many redshifts measured as possible.

(ii) Additional redshifts from the REFLEX catalogue were added to help fill in the void in the centre of the clusters.

(iii) The redshifts were used to plot histograms and a 6σ (this assumed a velocity dispersion of $\sigma = 1000\,\mathrm{km\,s^{-1}}$) cut made to select galaxy cluster members.

(iv) A basic estimation of cluster velocity dispersion was made using bootstrapping of the cluster galaxies and compared to the values derived from a σ - T_X relation to see if they were in agreement.

(v) A Dressler Shectman (DS) test was carried out for each cluster using its member galaxies and a plot produced. A DS statistic was also calculated for each cluster.

(vi) A *HST* image, with X-ray contour overlays, was produced for each cluster (including identified arcs where applicable), to enable visual comparison of other multi-wavelength data to the DS plots.

(vii) A list of the number of dark matter cluster-scale halos (where applicable) was taken from chapter 5 to compare to the DS statistics to determine if there was agreement on substructure.

In §3.7 a combined plot containing the *HST*/X-ray contour overlay plot, with the DS plot and redshift histogram is shown for each cluster and the various values compared and discussed as the work progressed for ease of comprehension. Also each cluster has a list of the positions and redshifts of each of its cluster galaxies.

3.7 Analysis & Interpretation

3.7.1 Abell 2813

Fig. 3.2 shows the resultant figures (described above) for A 2813. The plot of X-ray contours overlaid onto the *HST* observations (top plot of Fig. 3.2) showed a large offset (116.45 ± 23.29 kpc) from the X-ray centroid and BCG centres, suggesting disturbance in the cluster. The redshift histogram showed some evidence for twin peaks, which also

agreed with the DS plot showing evidence for substructure in the South-East of the cluster. The DS statistics also agree there could be possible substructure $\Delta/N = 1.29 \pm 0.41$ (however, the error did cross the boundary signifying the presence of substructure). The strong lensing analysis (see chapter 5) showed that the cluster consisted of two DM halos, one centred on the X-ray centroid and the other on the BCG. Table 3.7 lists the cluster members used.

3.7.2 Abell 2895

The *HST* frame showed an offset between the X-ray centroid and the BCG centre (Fig. 3.3). This suggested a possible disturbance displacing the hot intra-cluster gas in the clusters past and and the DS plot showed roughly even sized radii circles on the outside of the cluster, but with much larger ones near the centre that suggested possible substructure. However, the redshift histogram appeared to show no bimodality. This could mean that any disturbance was either very localized to the cluster centre and had occurred in the clusters very recent or distant past. It could also simply be a product of the low cluster numbers and the lack of information in the North-West of the cluster (the coverage was patchy).

The histogram and DS plots (bottom of Fig. 3.3) conflicted on the substructures presence, however $\Delta/N = 0.89 \pm 0.29$ suggested no strong substructure. This lack of substructure was backed by the strong lensing (SL) model in chapter 5 which only had one DM halo. Table 3.8 lists the cluster members used. The spectrum for the multiple images A1 and A2 showed strong Lyman-Alpha emission (see Fig. 3.4) which put their respective redshifts at $z_{A1} = 3.39$ and $z_{A2} = 3.72$.

3.7.3 Abell 368

Fig. 3.5 shows in the *HST* frame that the X-ray centroid and BCG centre coincided with a very small offset (7.73 ± 1.55 kpc). The redshift histogram only showed one peak (possibly

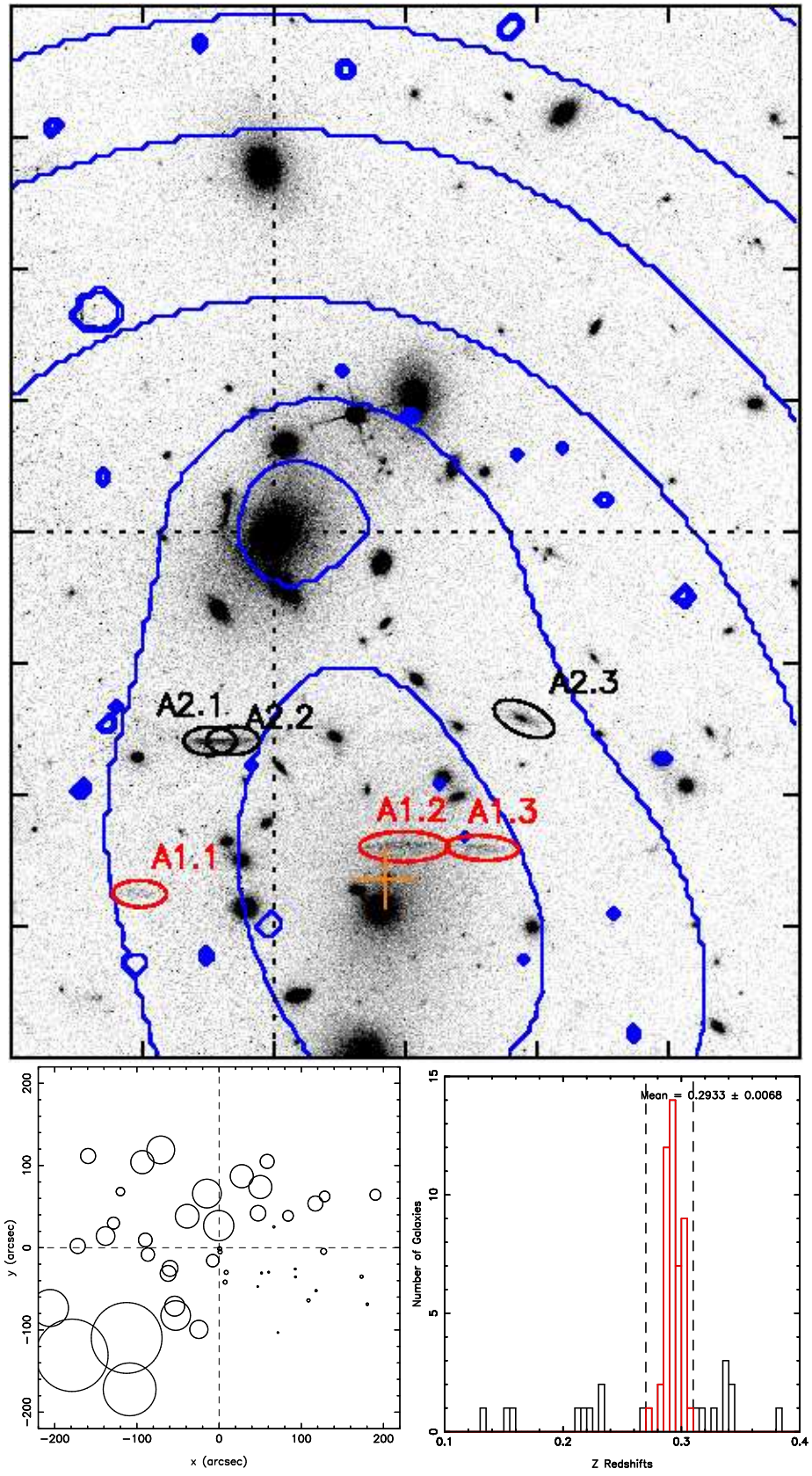


Figure 3.2: Top - Abell 2813 *HST* frame with X-rays overlaid in blue contours. Any arcs with measured redshifts are marked. Bottom Left - The resulting cluster member DS analysis. Bottom Right - Redshift histogram with the cluster members highlighted in red.

Table 3.7: Cluster Members for Abell 2813

Name	α, δ [J2000]	z	spectral features / Comments
37	10.854583, −20.617111	0.3004 ± 0.0003	BCG, Guzzo et al. (2009)
30	10.837417, −20.587611	0.2888 ± 0.0001	Guzzo et al. (2009)
32	10.846583, −20.592611	0.2930 ± 0.0007	Guzzo et al. (2009)
34	10.840708, −20.605111	0.2963 ± 0.0004	Guzzo et al. (2009)
35	10.854917, −20.609306	0.2932 ± 0.0004	Guzzo et al. (2009)
38	10.854417, −20.618194	0.2922 ± 0.0004	Guzzo et al. (2009)
39	10.852208, −20.625111	0.2858 ± 0.0003	Guzzo et al. (2009)
40	10.852583, −20.628389	0.2866 ± 0.0002	Guzzo et al. (2009)
F108	10.907917, −20.653111	0.3012 ± 0.0029	Hγ, Ca H & K, G, Hβ
F113	10.905833, −20.616167	0.3001 ± 0.0031	Ca H & K, G, Mβ, Hβ
F119	10.902083, −20.585778	0.2904 ± 0.0008	Hγ, Ca H & K, G, Mβ, Hβ
F133	10.859167, −20.598472	0.3074 ± 0.0013	Hγ, Ca H & K, G, Mβ, Hβ
F147	10.857083, −20.621139	0.2967 ± 0.0015	Hγ, Ca H & K, G, Mβ, Hβ
F277	10.862083, −20.644361	0.2988 ± 0.0017	Hγ, Ca H & K, G, Hβ
F283	10.892917, −20.608444	0.2922 ± 0.0017	Hγ, Ca H & K, G, Mβ, Hβ
F289	10.890417, −20.597806	0.3005 ± 0.0015	Hγ, Ca H & K, G, Mβ, Hβ
F301	10.887083, −20.664611	0.3005 ± 0.0013	Hγ, Ca H & K, G, Mβ, Hβ
F316	10.8825, −20.587861	0.2941 ± 0.0024	Hγ, Ca H & K, G
F336	10.875833, −20.583722	0.2932 ± 0.0011	Hγ, Ca H & K, G, Mβ, Hβ
F339	10.8725, −20.62375	0.3042 ± 0.0013	Hγ, Ca H & K, G, Mβ, Hβ
F94	10.915833, −20.637139	0.2987 ± 0.0011	Hγ, Ca H & K, G
F17	10.798333, −20.598889	0.2919 ± 0.0014	Ca H & K, G
F20	10.80125, −20.635861	0.2880 ± 0.0012	Hγ, Ca H & K, G, Hβ
F215	10.840833, −20.629889	0.3019 ± 0.0010	Hγ, Ca H & K, G, Mβ, Hβ
F244	10.835, −20.609722	0.2904 ± 0.0010	Hγ, Ca H & K, G, Mβ, Hβ
F27	10.803333, −20.626528	0.2942 ± 0.0003	Ca H & K, G
F64	10.816667, −20.599444	0.3014 ± 0.0010	Hγ, Ca H & K, G, Mβ, Hβ
F73	10.82, −20.601833	0.2899 ± 0.0013	Hγ, Ca H & K, G, Mβ, Hβ
F81	10.8225, −20.634556	0.2858 ± 0.0027	Hγ, Ca H & K, G, Hβ
1G48	10.87082, −20.636511	0.2999 ± 0.0010	Hγ, Ca H & K, G, Mβ, Hβ
1G27	10.88142, −20.614149	0.2889 ± 0.0014	Ca H & K, G, Mβ, Hβ
1G57	10.86631, −20.606131	0.2941 ± 0.0004	G, Mβ, Hβ
1G17	10.88816, −20.647249	0.2967 ± 0.0009	Ca H & K, Mβ, Hβ
1G215	10.839889, −20.59627	0.2889 ± 0.0009	Mβ, Hβ
1G208	10.83688, −20.62504	0.2835 ± 0.0010	Ca H & K, G, Mβ
1G204	10.83355, −20.645439	0.2864 ± 0.0013	Ca H & K, G, Mβ, Hβ
1G169	10.81977, −20.63129	0.2868 ± 0.0011	Ca H & K, G, Mβ, Hβ
1G189	10.8272, −20.62664	0.2978 ± 0.0012	Ca H & K, G, Hβ
1G195	10.829879, −20.606001	0.3022 ± 0.0016	Hβ, Mβ
2G6	10.895779, −20.61282	0.2939 ± 0.0027	G, Mβ, Hβ
2G50	10.87045, −20.639681	0.2949 ± 0.0008	Ca H & K, G, Mβ, Hβ
2G45	10.87319, −20.625441	0.2882 ± 0.0014	Ca H & K, G, Mβ
2G31	10.8805, −20.61907	0.2881 ± 0.0011	Ca H & K, G, Mβ, Hβ
2G216	10.83946, −20.625311	0.2708 ± 0.0007	Ca H & K, G, Mβ, Hβ
2G191	10.82728, −20.62396	0.2835 ± 0.0008	Ca H & K, G, Mβ
2G105	10.81701, −20.618019	0.2915 ± 0.0011	Ca H & K, G, Mβ, Hβ

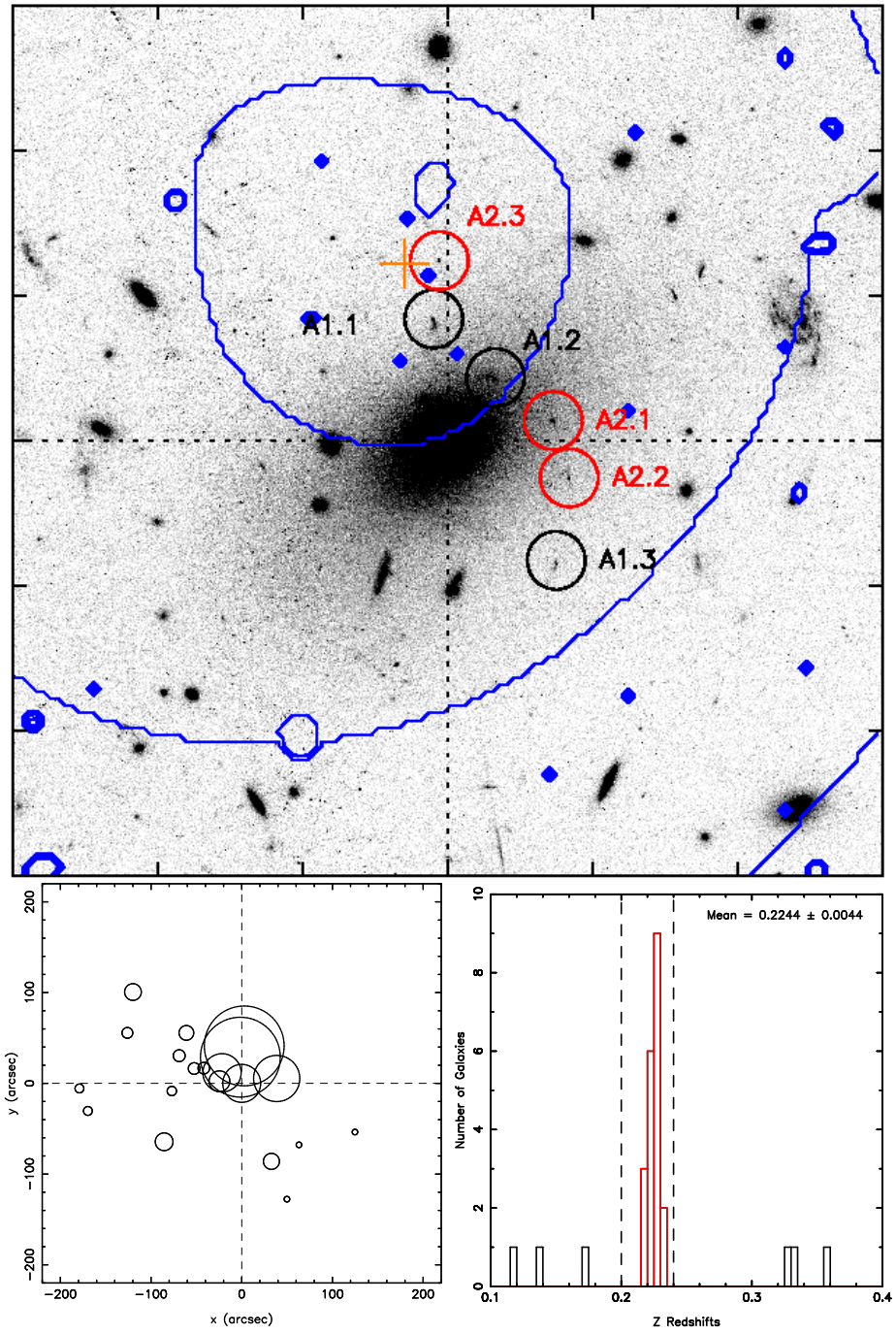


Figure 3.3: Top - Abell 2895 *HST* frame with X-rays overlaid in blue contours. Any arcs with measured redshifts are marked. Bottom Left - The resulting cluster member DS analysis. Bottom Right - Redshift histogram with the cluster members highlighted in red.

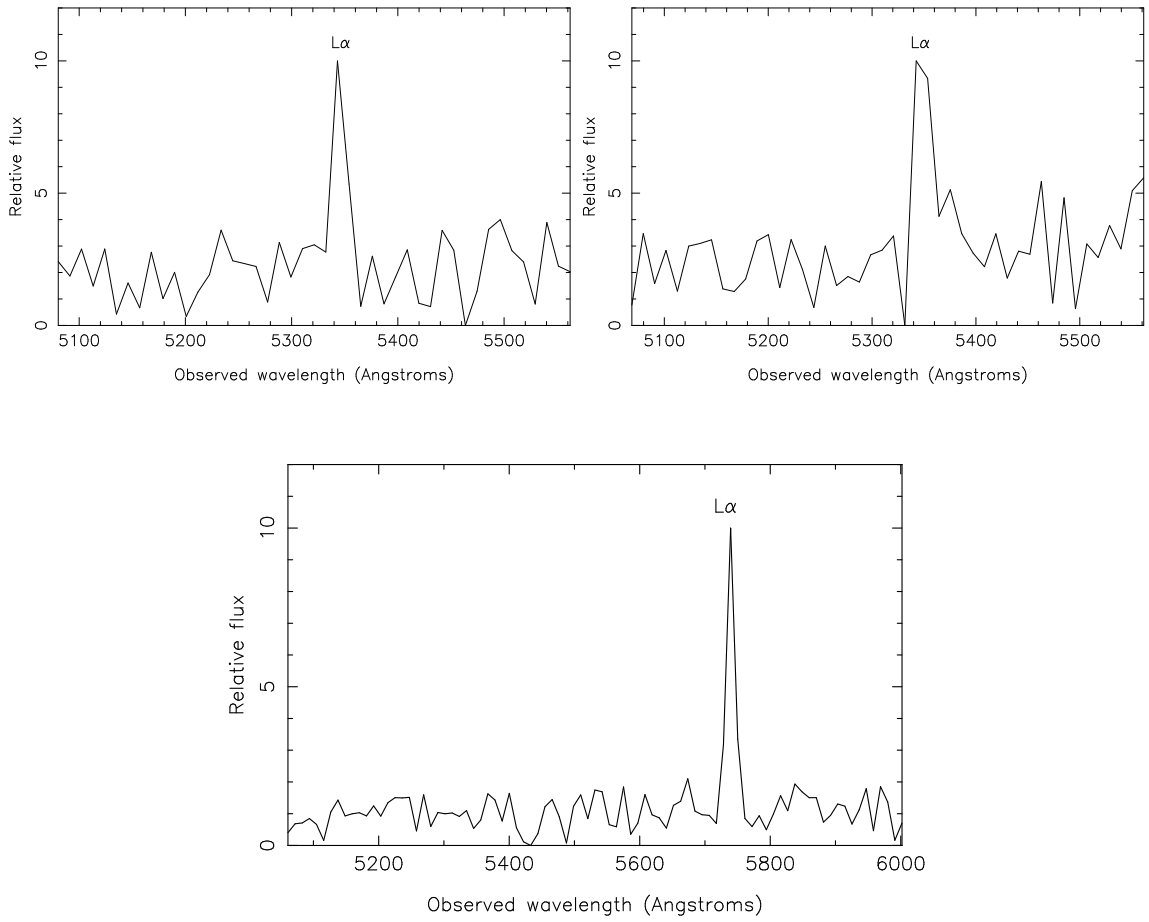


Figure 3.4: Abell 2895 multiple images. Top Left - The Lyman-Alpha emission line for the multiple image A1.1 (z=3.39). Top Right - The Lyman-Alpha emission line for the multiple image A1.2 (z=3.39). Bottom - The Lyman-Alpha emission line for the multiple image A2.1 (z=3.72).

Table 3.8: Cluster Members for Abell 2895

Name	α, δ [J2000]	z	spectral features / Comments
BCG	19.546146, −26.96999	0.2310	Cruddace et al. (2002)
1G40	19.599010, −26.978430	0.2325 ± 0.0006	Ca H & K, G, Mβ, Hβ
1G28	19.583530, −26.942030	0.2255 ± 0.0013	Hβ, Mβ
1G20	19.601920, −26.971550	0.2251 ± 0.0005	Hβ, [OIII]
1G200	19.572809, −26.987869	0.2205 ± 0.0009	Ca H & K, G, Mβ
1G395	19.552919, −26.966770	0.2209 ± 0.0010	G, Mβ, Hβ
1G654	19.546611, −26.961969	0.2288 ± 0.0012	G, Mβ, Hβ
1G119	19.565170, −26.954540	0.2290 ± 0.0011	G, Mβ, Hβ
1G518	19.535941, −26.993870	0.2262 ± 0.0011	Ca H & K, G, Mβ
1G471	19.526480, −26.988810	0.2254 ± 0.0006	Ca H & K, G, Mβ, Hβ
1G314	19.507231, −26.984890	0.2214 ± 0.0010	G, Mβ, Hβ
2G50	19.585440, −26.954540	0.2181 ± 0.0016	G, Mβ, NaD
2G156	19.570169, −26.972340	0.2187 ± 0.0008	Ca H & K, G, Mβ, NaD
2G195	19.562509, −26.965450	0.2231 ± 0.0012	G, Mβ, NaD
2G225	19.559171, −26.965340	0.2217 ± 0.0012	G, Mβ, Hβ
2G392	19.553801, −26.969339	0.2269 ± 0.0003	G, Mβ, NaD, Hβ
2G177	19.567630, −26.961519	0.2259 ± 0.0045	Hβ, NaD
2G659	19.545280, −26.958530	0.2287 ± 0.0154	G, Mβ
2G582	19.534130, −26.968500	0.2216 ± 0.0012	G, Mβ, Hβ
2G365	19.530600, −27.005449	0.2166 ± 0.0010	Mβ, SI, SII

exhibiting asymmetry) and the DS plot showed evidence for substructure in the centre of the cluster. The DS statistics gave very strong evidence for strong substructure with $\Delta/N = 1.58 \pm 0.26$ and $P = 0.04^{+0.08}_{-0.04}$. The lensing analysis in chapter 5 disagreed with the interpretation of large amounts of substructure, as the lensing model had only one DM halo that was coincident with the BCG. This discrepancy could have been due to the small scales the strong lensing model probed and so might not be sensitive to a large enough radii that encompassed the positions of the large clump of large radii circles in the DS plot that could be substructure. Table 3.9 lists the cluster members used.

For the multiple image system the redshift was determined by matching a P-cygni profile (both absorption and emission) around 4740 Å in the FORS2 data. This corresponded to CIV at around $z = 2.07$[1].

3.7.4 Abell 3084

This cluster was studied in detail (showing bimodality) in chapter 4, but for completeness a similar comparison to other clusters is included here.

The *HST* frame in Fig. 3.6 showed a large offset from the BCG and X-ray centroid 46.09 ± 9.22 kpc, suggesting the cluster was disturbed. The redshift histogram showed slight signs of asymmetry and the DS plot showed some evidence for substructure in the South-East of the cluster. The DS statistic cannot determine if there was substructure or not with $\Delta/N = 0.95 \pm 0.21$ being just below 1, so nothing conclusive could be drawn. The strong lensing analysis of this cluster in chapter 4 showed that two DM halos were required in order to reproduce the strong lensing constraints the cluster exhibited. This discrepancy could be due to the low cluster member numbers and the poor coverage in the centre of the cluster within which the SL constraints are situated. Cluster members were listed in Table 4.1 of chapter 4 for ease of the detailed analysis of that chapter.

[1]This redshift determination was performed by Johan Richard, CRAL, Observatoire de Lyon, Université de Lyon 1, 9 avenue Ch. André, F-69561 Saint-Genis Laval, France

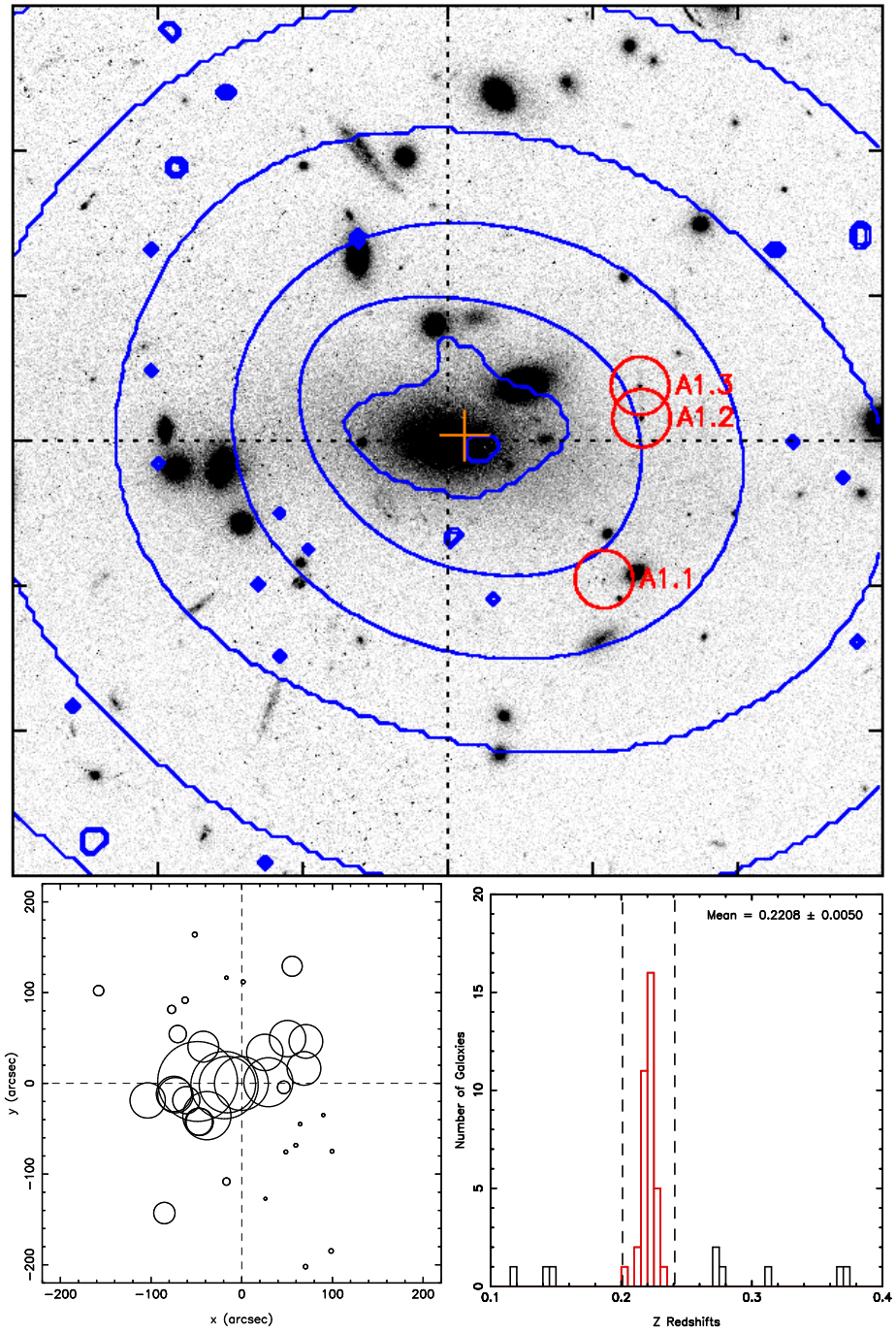


Figure 3.5: Top - Abell 368 *HST* frame with X-rays overlaid in blue contours. Any arcs with measured redshifts are marked. Bottom Left - The resulting cluster member DS analysis. Bottom Right - Redshift histogram with the cluster members highlighted in red.

Table 3.9: Cluster Members for Abell 368

Name	α, δ [J2000]	z	Spectrum features / Comments
150	39.365708, −26.508	0.2219 ± 0.0003	BCG, Guzzo et al. (2009)
145	39.344292, −26.503389	0.2196 ± 0.0002	Guzzo et al. (2009)
146	39.351208, −26.509194	0.2206 ± 0.0002	Guzzo et al. (2009)
147	39.3565, −26.507611	0.2208 ± 0.0002	Guzzo et al. (2009)
148	39.357708, −26.4985	0.2260 ± 0.0003	Guzzo et al. (2009)
151	39.3705, −26.508333	0.2300 ± 0.0001	Guzzo et al. (2009)
152	39.3715, −26.5085	0.2246 ± 0.0001	Guzzo et al. (2009)
153	39.380292, −26.519694	0.2209 ± 0.0002	Guzzo et al. (2009)
154	39.380708, −26.507389	0.2210 ± 0.0002	Guzzo et al. (2009)
156	39.388792, −26.511194	0.2254 ± 0.0003	Guzzo et al. (2009)
157	39.397792, −26.513194	0.2265 ± 0.0003	Guzzo et al. (2009)
F102	39.3875, −26.492861	0.2280 ± 0.0022	Hγ, Ca H & K, G, Mβ, Hβ
F164	39.384583, −26.513111	0.2243 ± 0.0027	Hγ, Ca H & K, G, Mβ, Hβ
F224	39.348333, −26.472111	0.2198 ± 0.0015	Hγ, Ca H & K, G, Mβ, Hβ
F231	39.370833, −26.475639	0.2191 ± 0.0024	Hγ, Ca H & K, G, Mβ, Hβ
F233	39.414583, −26.479583	0.2177 ± 0.0008	Ca H & K, G, Mβ, Hβ
F244	39.389583, −26.485333	0.2225 ± 0.0028	Hγ, Ca H & K, G, Mβ, Hβ
F261	39.37875, −26.496722	0.2233 ± 0.0020	Hγ, Ca H & K, G, Mβ, Hβ
F72	39.381667, −26.462417	0.2164 ± 0.0022	Hγ, Ca H & K, G, Mβ, Hβ
F90	39.385, −26.4825	0.2033 ± 0.0011	[OIII], Hα, NII, SI, SII
F14	39.335, −26.559278	0.2213 ± 0.0006	Hγ, Ca H & K, G, Mβ, Hβ
F177	39.38, −26.519778	0.2195 ± 0.0017	Hγ, Ca H & K, G
F192	39.347083, −26.526917	0.2186 ± 0.0014	Ca H & K, G, Mβ
F31	39.392083, −26.547667	0.2177 ± 0.0020	Hγ, Ca H & K, G, Mβ, Hβ
F36	39.3575, −26.543278	0.2219 ± 0.0011	Hγ, Ca H & K, G, Mβ, Hβ
F47	39.370833, −26.538028	0.2229 ± 0.0019	Hγ, Ca H & K, G, Mβ, Hβ
F7	39.34375, −26.564083	0.2145 ± 0.0013	Hγ, Ca H & K, G, Mβ
3G174	39.334692, −26.528749	0.2196 ± 0.0005	Ca H & K, G, Mβ
3G237	39.337682, −26.51767	0.2208 ± 0.0013	Ca H & K, G, Mβ
3G222	39.345621, −26.520439	0.2154 ± 0.0010	H, G, Mβ, Hβ
3G451	39.34364, −26.495119	0.2289 ± 0.0009	Ca H & K, G, Mβ, Hβ
4G164	39.350499, −26.528959	0.2106 ± 0.0009	Ca H & K, G, Mβ, Hβ
4G567	39.377528, −26.517851	0.2231 ± 0.0004	Mβ, NaD, Hβ
4G537	39.388311, −26.511589	0.2219 ± 0.0001	Mβ, NaD
4G457	39.34992, −26.49427	0.2218 ± 0.0010	Ca H & K, G, Mβ, NaD
4G297	39.36509, −26.4769	0.2188 ± 0.0010	G, Mβ, NaD, Hβ

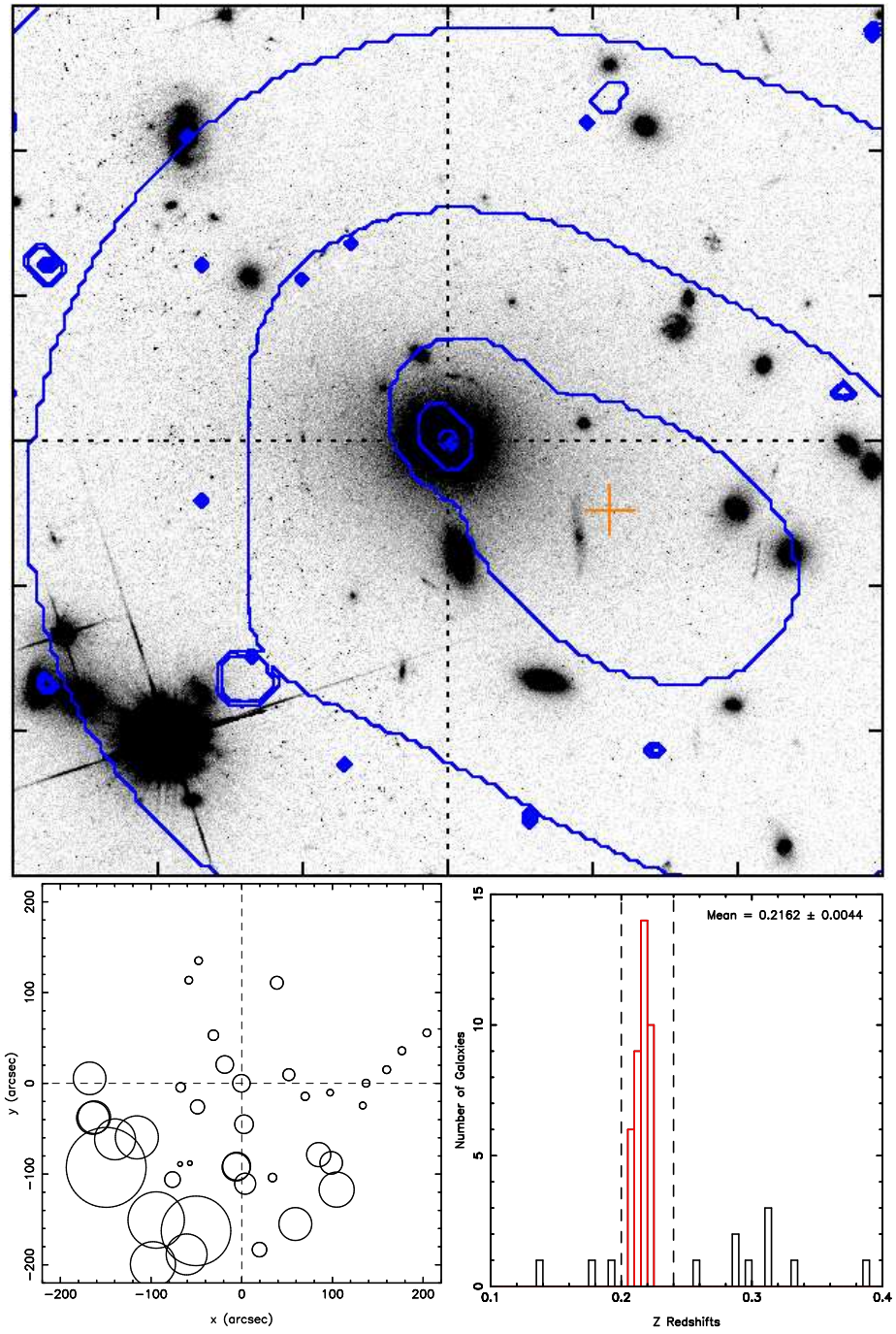


Figure 3.6: Top - Abell 3084 *HST* frame with X-rays overlaid in blue contours. Any arcs with measured redshifts are marked. Bottom Left - The resulting cluster member DS analysis. Bottom Right - Redshift histogram with the cluster members highlighted in red.

Table 3.10: Cluster Members for Abell 3088

Name	α, δ [J2000]	z	Spectrum features / Comments
220	46.758708, −28.665889	0.2527 ± 0.0002	BCG, Guzzo et al. (2009)
218	46.755583, −28.689611	0.2602 ± 0.0003	Guzzo et al. (2009)
219	46.752792, −28.664	0.2534 ± 0.0002	Guzzo et al. (2009)
221	46.761583, −28.668611	0.2512 ± 0.0004	Guzzo et al. (2009)
222	46.769292, −28.6755	0.2478 ± 0.0002	Guzzo et al. (2009)
223	46.7675, −28.662694	0.2538 ± 0.0002	Guzzo et al. (2009)
225	46.779708, −28.6695	0.2603 ± 0.0004	Guzzo et al. (2009)
226	46.7805, −28.657	0.2547 ± 0.0002	Guzzo et al. (2009)
228	46.792208, −28.664389	0.2537 ± 0.0001	Guzzo et al. (2009)
7G43	46.795231, −28.661209	0.2547 ± 0.0010	Ca H & K, G, Mβ, NaD, Hβ
7G37	46.786898, −28.670401	0.2526 ± 0.0009	Ca H, G, Mβ, NaD, Hβ
7G26	46.779900, −28.706770	0.2612 ± 0.0005	Ca H & K, G, Hβ
7G24	46.758578, −28.692310	0.2413 ± 0.0009	Ca H & K, G, Mβ, Hβ
7G27	46.746701, −28.671320	0.2407 ± 0.0009	G, Mβ, NaD, Hβ
7G22	46.731412, −28.669050	0.2483 ± 0.0010	G, Mβ, NaD, Hβ
7G13	46.736419, −28.703239	0.2513 ± 0.0008	Ca H & K, G, Mβ, NaD, Hβ
7G12	46.734731, −28.707050	0.2508 ± 0.0016	Ca H & K, G, Mβ, NaD, Hβ
8G47	46.798421, −28.650339	0.2465 ± 0.0008	Ca H & K, G, Mβ, NaD, Hβ
8G42	46.789799, −28.661551	0.2398 ± 0.0001	Hβ, Hα, NII, SI, SII
8G36	46.758410, −28.636530	0.2539 ± 0.0007	Mβ, NaD, Hβ
8G46	46.763241, −28.616739	0.2518 ± 0.0008	Mβ, NaD
8G28	46.738340, −28.659201	0.2534 ± 0.0005	Mβ, NaD

3.7.5 Abell 3088

The *HST* frame in Fig. 3.7 showed smooth contours with only a small offset (14.63 ± 2.92 kpc) from the X-ray centroid and BCG centre. The redshift histogram showed some evidence for asymmetry in its single peak, but this could be due to the low cluster member counts not sampling the distribution sufficiently. The DS plot appeared to show no real substructure (possibly again due to under sampling). The DS statistic also agreed with this interpretation with $\Delta/N = 1.04 \pm 0.28$ (it encompassed the boundary condition of $\Delta/N > 1$ for substructure being present). This could again be an anomaly due to the uneven distribution of the redshift cluster members across the sky and no strong conclusions could be drawn without further investigation. Table 3.10 lists the cluster members used.

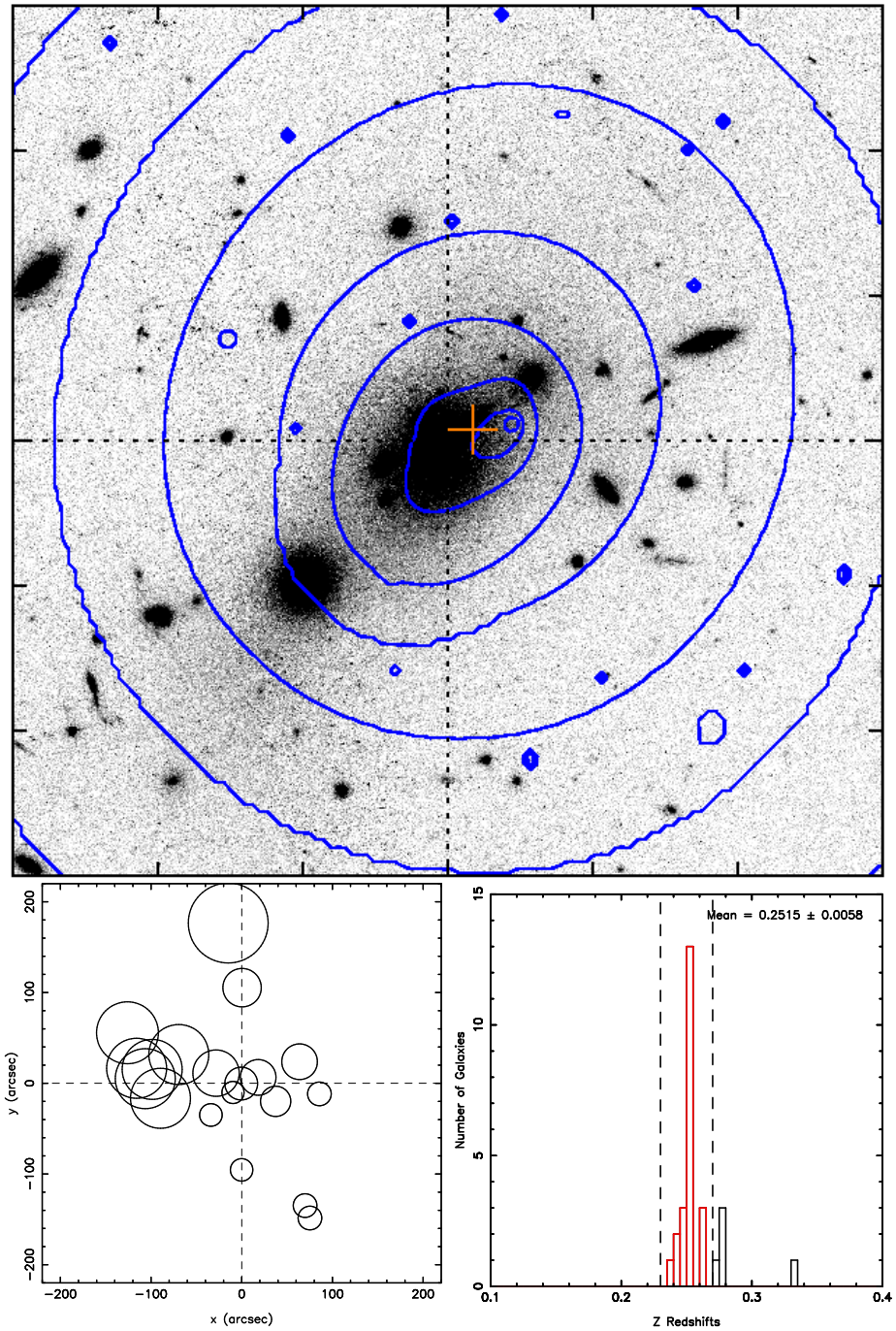


Figure 3.7: Top - Abell 3088 *HST* frame with X-rays overlaid in blue contours. Any arcs with measured redshifts are marked. Bottom Left - The resulting cluster member DS analysis. Bottom Right - Redshift histogram with the cluster members highlighted in red.

3.7.6 Abell 2537

The *HST* frame in Fig. 3.8 showed a slight offset from the X-ray centroid and BCG (25.25 ± 5.05 kpc) but, apart from the very centre, the rest of the X-ray contours appeared smooth. The redshift histogram showed evidence for asymmetry, but nothing conclusive. The DS plot showed a large amount of substructure in the cluster to both the East and West. The DS statistics backed this up, with very strong evidence for substructure with $\Delta/N = 1.44 \pm 0.29$ and $P = 0.01) - 0.01^{+0.05}$ ($P \approx 0.05$ is a low false positive rate accepted value for very strong substructure). The substructure was divided distinctly between East and West of the BCG centre and was again in agreement with the lensing analysis in chapter 5, which had two DM masses, one coincident on the BCG, the other to the west. Table 3.11 lists the cluster members used.

3.7.7 Abell 3364

In the *HST* frame in Fig. 3.9 the X-ray contours showed disturbance to the North-East of the BCG and the X-ray centroid was also offset from the BCG by 40.94 ± 8.19 kpc. The redshift histogram showed some slight asymmetry, but no evidence of twin peaks. The DS plot showed a large deviation from the global kinematics in the North of the cluster, again, agreeing with the *HST* frame. This suggested a large amount of substructure, possibly caused by a displaced DM halo from a merger, or more than one DM halo. The DS statistics however disagreed as both values showed no strong sign of substructure. Table 3.12 lists the cluster members used, although the measured redshift for the BCG was substantially different (possibly due to error or other unknown factor) from that of the cluster, which formed the majority of the histogram redshift space, so it was excluded from the analysis so as to not bias the results.

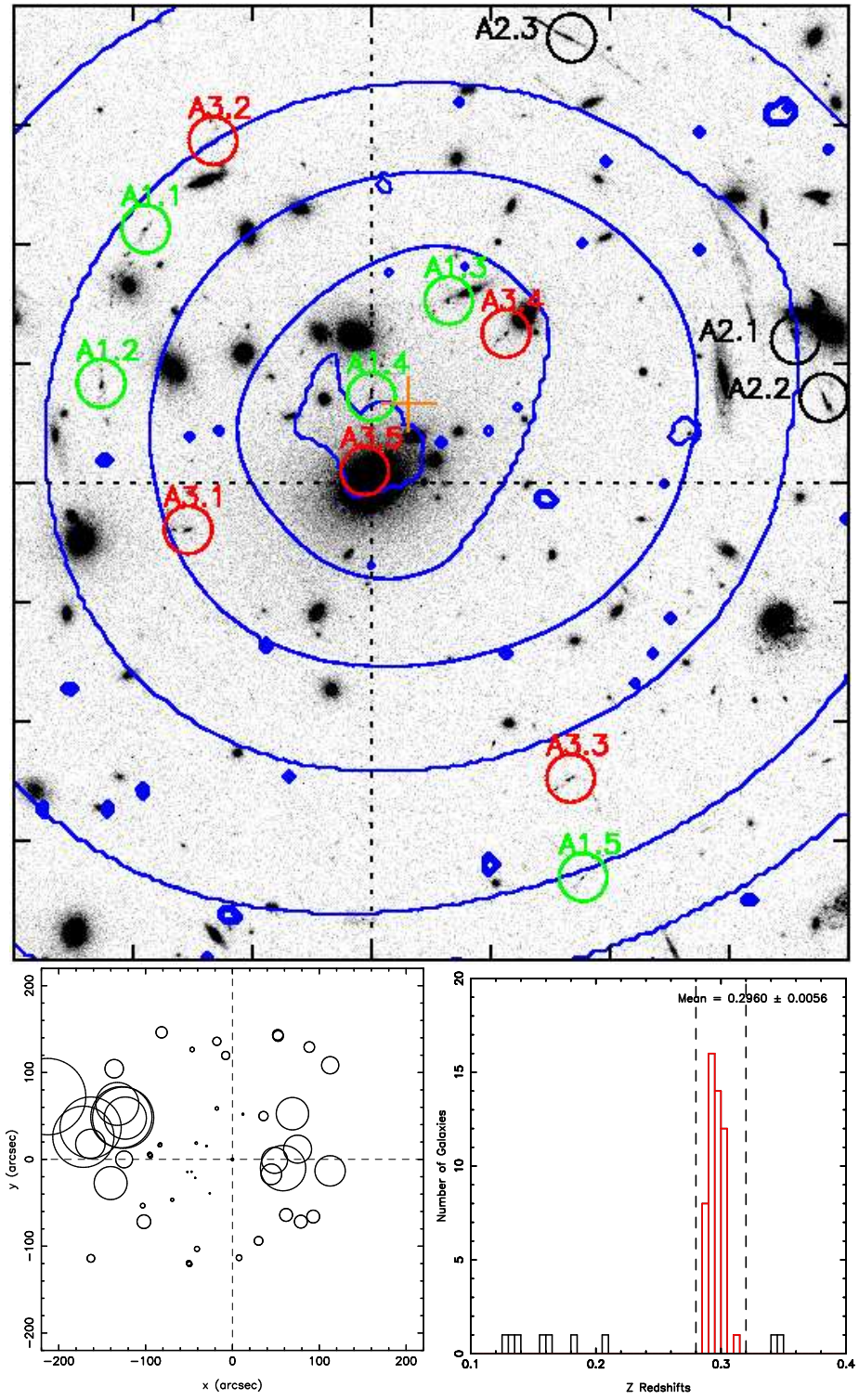


Figure 3.8: Top - Abell 2537 *HST* frame with X-rays overlaid in blue contours. Any arcs with measured redshifts are marked. Bottom Left - The resulting cluster member DS analysis. Bottom Right - Redshift histogram with the cluster members highlighted in red.

Table 3.11: Cluster Members for Abell 2537

Name	α, δ [J2000]	z	Spectrum features / Comments
BCG	347.09294, −2.1925704	0.2950	Dahle et al. (2002)
F88	347.08958, −2.1781111	0.2961 ± 0.0030	Hγ, Ca H & K, G, Mβ
F98	347.11917, −2.1913889	0.2948 ± 0.0023	Hγ, Ca H & K, G, Hβ
F77	347.09792, −2.1762778	0.2887 ± 0.0016	Hγ, Ca H & K, G
F60	347.06833, −2.1566389	0.2962 ± 0.0007	Hγ, Ca H & K, G
F47	347.07833, −2.1531111	0.2969 ± 0.0008	Hγ, Ca H & K, G, Mβ, 5268, Hβ
F64	347.12792, −2.17925	0.2982 ± 0.0017	Hγ, Ca H & K, G, Mβ
F62	347.13833, −2.1823889	0.2920 ± 0.0015	Ca H & K, G
F37	347.09792, −2.1548056	0.2897 ± 0.0022	Hγ, Ca H & K, G, Hβ
F36	347.10583, −2.1574167	0.2966 ± 0.0010	Hγ, Ca H & K, G, Mβ, Hβ
F34	347.15208, −2.1725	0.3040 ± 0.0012	Hγ, Ca H & K, G
F165	347.07208, −2.1893056	0.3021 ± 0.0013	[OII], Hβ, [OIII]
F204	347.06167, −2.19625	0.3012 ± 0.0008	Hγ, Ca H & K, G, Mβ, Hβ
F239	347.10667, −2.2260833	0.2913 ± 0.0066	Ca H & K, Mβ, 5268
F244	347.08458, −2.2186389	0.2863 ± 0.0012	Hγ, Ca K, H
F235	347.06708, −2.2108611	0.2917 ± 0.0010	Hγ, Ca H & K, G, Mβ, 5268, Hβ
11G24	347.12723, −2.17936	0.2942 ± 0.0005	Ca H & K, G, Hβ
11G33	347.12977, −2.17483	0.3044 ± 0.0002	Hα, NII
11G29	347.13078, −2.16357	0.2854 ± 0.0011	Hα, NII
11G43	347.13192, −2.20015	0.2918 ± 0.0010	Ca H & K, G, Mβ, Hβ
11G93	347.12128, −2.2125	0.2968 ± 0.0011	Ca H & K, Hβ
11G84	347.12165, −2.20738	0.2991 ± 0.0011	Ca H, G, Hβ
11G72	347.11621, −2.18805	0.2952 ± 0.0001	Ca H & K, Hβ
11G327	347.10428, −2.22121	0.3027 ± 0.0012	Ca H & K, Mβ, Hβ
11G108	347.10451, −2.18737	0.2897 ± 0.0010	Ca H & K, G, Mβ, Hβ
11G124	347.10603, −2.19652	0.2944 ± 0.0006	Ca H & K, G, Hβ
11G235	347.09081, −2.22403	0.3011 ± 0.0015	Ca H & K, G, Mβ
11G303	347.07945, −2.19288	0.2935 ± 0.0031	Hβ, Mβ
12G32	347.13841, −2.18764	0.3143 ± 0.0033	Ca H & K
12G28	347.14065, −2.18534	0.3009 ± 0.0010	Ca H & K, G, Mβ, Hβ
12G25	347.12849, −2.17908	0.2985 ± 0.0012	Ca H & K, G, Mβ, Hβ
12G116	347.11221, −2.20548	0.2921 ± 0.0014	Ca H & K, G, Mβ, Hβ
12G126	347.10125, −2.1883	0.2904 ± 0.0011	Ca H & K, G
12G324	347.083, −2.17871	0.2913 ± 0.0003	G, Mβ
12G280	347.07673, −2.19537	0.3041 ± 0.0012	Ca H, Hβ
12G176	347.08048, −2.19734	0.3020 ± 0.0019	Ca H & K, G
12G259	347.07576, −2.21034	0.3010 ± 0.0009	Ca H & K, G, Mβ, Hβ
12G237	347.07104, −2.21249	0.2894 ± 0.0009	Ca H & K, G
13G23	347.06165, −2.16246	0.3035 ± 0.0009	Mβ, 5268
13G21	347.07842, −2.15274	0.2964 ± 0.0016	Ca H, G, Mβ
13G26	347.11561, −2.15193	0.2992 ± 0.0066	[OII], Hβ, [OIII]
13G276	347.11942, −2.19099	0.2944 ± 0.0012	Ca H & K, G, Mβ
13G260	347.10485, −2.1985	0.2896 ± 0.0006	Ca H & K, G, Mβ
13G134	347.13824, −2.22424	0.3027 ± 0.0023	Ca H & K, Mβ
14G27	347.0951, −2.15933	0.2998 ± 0.0115	Hα, NII, SI, SII
14G55	347.07373, −2.17795	0.2939 ± 0.0115	Mβ, 5268, Hβ
14G15	347.10019, −2.20343	0.2878 ± 0.0019	[OIII]
14G255	347.1276, −2.19258	0.2929 ± 0.0015	Ca H & K, Mβ
14G278	347.11604, −2.18782	0.2958 ± 0.0012	Ca H & K, G, Mβ, Hβ
14G261	347.10735, −2.19658	0.2922 ± 0.0021	Ca H & K, G, Mβ, Hβ
14G147	347.10689, −2.22573	0.2956 ± 0.0008	Ca H & K, G, Mβ, Hβ

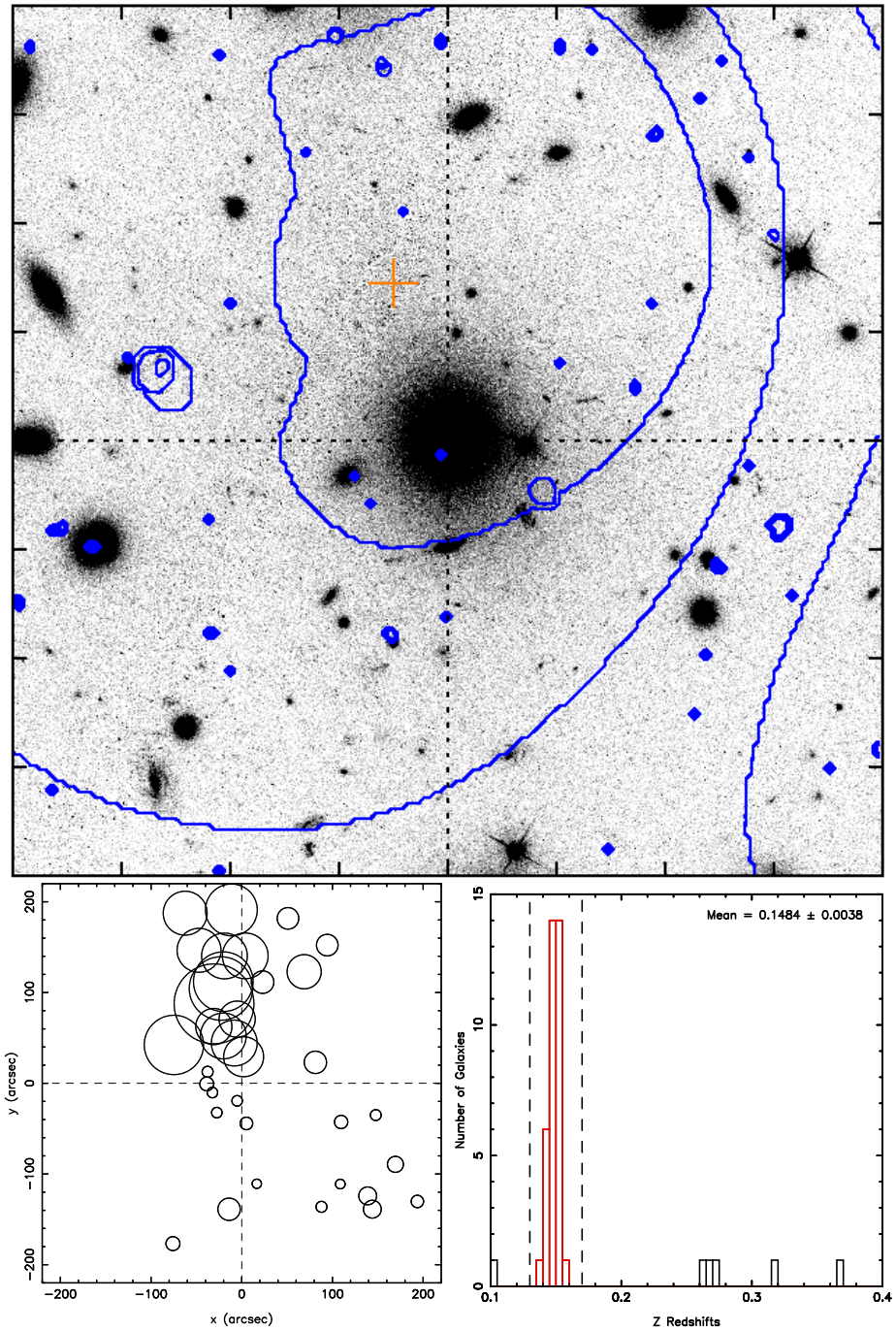


Figure 3.9: Top - Abell 3364 *HST* frame with X-rays overlaid in blue contours. Any arcs with measured redshifts are marked. Bottom Left - The resulting cluster member DS analysis. Bottom Right - Redshift histogram with the cluster members highlighted in red.

Table 3.12: Cluster Members for Abell 3364

Name	α, δ [J2000]	z	Spectrum features / Comments
516	86.907292, −31.873389	0.3693 ± 0.0002	BCG, Guzzo et al. (2009), removed
515	86.916083, −31.882194	0.1487 ± 0.0002	Guzzo et al. (2009)
518	86.919417, −31.869694	0.1472 ± 0.0003	Guzzo et al. (2009)
519	86.909792, −31.860889	0.1537 ± 0.0002	Guzzo et al. (2009)
522	86.917083, −31.848889	0.1506 ± 0.0002	Guzzo et al. (2009)
523	86.914583, −31.844194	0.1444 ± 0.0001	Guzzo et al. (2009)
F1007	86.91375, −31.842111	0.1524 ± 0.0016	Hγ, Ca H & K, G, Mβ, Hβ
F112	86.899583, −31.84225	0.1512 ± 0.0012	Hγ, Ca H & K, G, Mβ, NaD, Hβ
F113	86.910833, −31.820139	0.1478 ± 0.0014	Hγ, Ca H & K, G, Mβ, Hβ
F115	86.905833, −31.834194	0.1461 ± 0.0007	Hγ, Ca H & K, G, Mβ, NaD, Hβ
F146	86.90875, −31.853583	0.1514 ± 0.0009	Hγ, Ca H & K, G, Mβ, NaD, Hβ
F161	86.9275, −31.821139	0.1508 ± 0.0032	Hγ, Ca H & K, G, Mβ, NaD, Hβ
F166	86.9225, −31.8325	0.1522 ± 0.0007	Ca H & K, G
F177	86.91375, −31.858722	0.1462 ± 0.0010	Hγ, Ca H & K, G, Mβ, NaD, Hβ
F195	86.90875, −31.878583	0.1492 ± 0.0010	Ca H & K, Mβ, NaD
F51	86.87625, −31.830917	0.1537 ± 0.0014	Hγ, Ca H & K, G, Mβ, NaD, Hβ
F77	86.890417, −31.822722	0.1464 ± 0.0010	Ca H & K, G, Mβ, NaD, Hβ
F84	86.884583, −31.839056	0.1398 ± 0.0010	Hγ, Ca H & K, G, Mβ, NaD
F132	86.878333, −31.911028	0.1457 ± 0.0007	Hγ, Ca H & K, G, Mβ, NaD, Hβ
F190	86.905417, −31.8855	0.1512 ± 0.0011	Hγ, Ca H & K, G, Mβ, NaD, Hβ
F201	86.901667, −31.903972	0.1459 ± 0.0017	Hγ, Ca H & K, G, Mβ, Hβ
F230	86.911667, −31.911778	0.1468 ± 0.0081	Ca H & K, G
F52	86.84375, −31.909389	0.1510 ± 0.0015	Ca K, G
F54	86.851667, −31.898028	0.1565 ± 0.0014	Hγ, Ca H & K, G, Mβ, NaD, Hβ
F57	86.85875, −31.882944	0.1444 ± 0.0023	Ca H & K, G, Mβ
F93	86.87125, −31.885056	0.1485 ± 0.0018	Hγ, Ca H & K, G, Mβ, Hβ
F95	86.86, −31.91175	0.1432 ± 0.0007	Hγ, Ca H & K, G
F96	86.861667, −31.907667	0.1479 ± 0.0007	Hγ, Ca H & K, G, Mβ, NaD, Hβ
9G1	86.91333, −31.83424	0.1516 ± 0.0006	Mβ, 5268, Hβ
9G229	86.91767, −31.87605	0.1439 ± 0.0006	G, Mβ, NaD, Hβ
9G136	86.93157, −31.86147	0.1517 ± 0.0007	G, Mβ, NaD, Hβ
9G204	86.90642, −31.86509	0.1437 ± 0.0006	Mβ, NaD, Hβ
9G438	86.93191, −31.9223	0.1512 ± 0.0001	Hβ, Hα, NII, SI, SII
9G446	86.87154, −31.90406	0.1413 ± 0.0005	Mβ, NaD, Hβ
10G218	86.91972, −31.87341	0.1465 ± 0.0009	G, Mβ, NaD, Hβ
10G149	86.91716, −31.85585	0.1526 ± 0.0005	Mβ, Hβ
10G277	86.88053, −31.8668	0.1470 ± 0.0005	Mβ, NaD

Table 3.13: Cluster Members for RXCJ 0528.2 − 2942

Name	α, δ [J2000]	z	Spectrum features / Comments
460BCG	82.063, −29.717694	0.1535 ± 0.0002	Guzzo et al. (2009)
458	82.0765, −29.7095	0.1626 ± 0.0004	Guzzo et al. (2009)
5G84	82.054739, −29.747869	0.1529 ± 0.0006	Mβ, Hβ
5G183	82.059581, −29.727091	0.1532 ± 0.0009	Mβ, NaD, Hβ
5G124	82.067077, −29.726570	0.1569 ± 0.0002	Mβ, NaD
5G117	82.077792, −29.723730	0.1581 ± 0.0007	Mβ, NaD
5G202	82.028317, −29.730209	0.1550 ± 0.0005	Mβ, NaD
5G206	82.048724, −29.715710	0.1543 ± 0.0011	G, Mβ, NaD, Hβ
5G159	82.026994, −29.717480	0.1510 ± 0.0007	G, Mβ, NaD, Hβ
5G172	82.044761, −29.705151	0.1516 ± 0.0010	G, Mβ, NaD, Hβ
6G43	82.093828, −29.742630	0.1499 ± 0.0005	Mβ, NaD
6G81	82.051520, −29.747169	0.1538 ± 0.0007	G, Mβ, 5268, NaD, Hβ
6G237	82.062178, −29.723640	0.1601 ± 0.0009	G, Mβ, NaD, Hβ
6G68	82.094114, −29.732670	0.1514 ± 0.0007	Mβ, 5268, NaD, Hβ
6G210	82.077105, −29.703461	0.1555 ± 0.0006	Mβ, NaD, Hβ
6G199	82.038696, −29.718809	0.1589 ± 0.0004	G, Mβ, 5268, Hβ
6G167	82.044833, −29.703320	0.1533 ± 0.0008	G, Mβ, 5268, NaD, Hβ

3.7.8 RXCJ0528

The *HST* frame in Fig. 3.10 showed X-ray contours which were not smooth, but asymmetric, suggesting a disturbance of the intra-cluster gas, with an offset of 31.32 ± 6.26 kpc. However, the histogram showed a thin and slightly asymmetric distribution of member galaxies suggesting little disturbance. The DS plot showed slight variations in the cluster member radii (with two deviations) around the centre suggesting some substructure or disturbance. This would be in agreement with the X-ray overlay, but the DS statistics both give values ($\Delta/N = 0.92$ and $P = 0.72$) suggesting no significant substructure. The suggestion of substructure could be due to the low galaxy cluster members and the poor sky coverage of the members available. Table 3.13 lists the cluster members used.

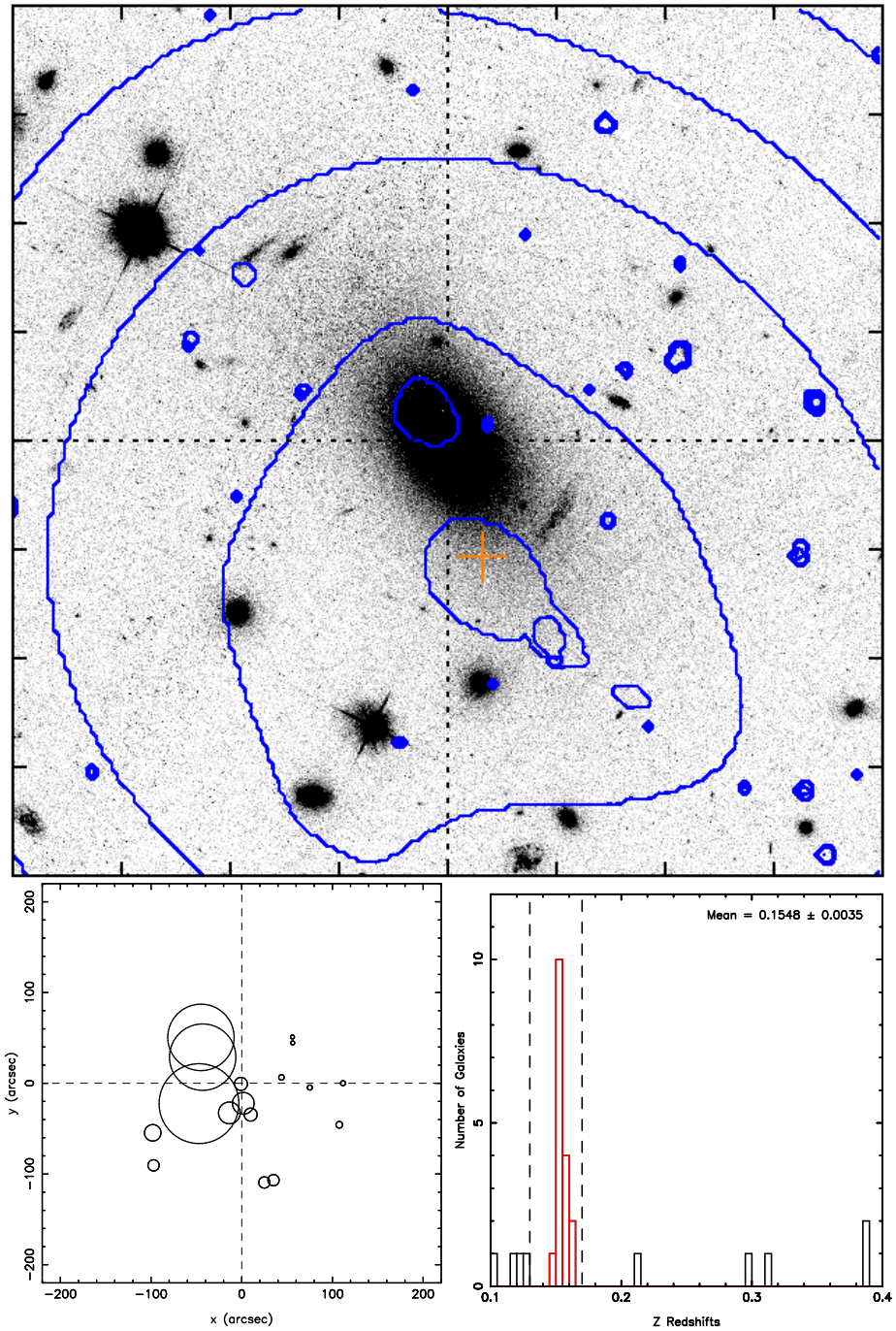


Figure 3.10: Top - RXCJ0528 *HST* frame with X-rays overlaid in blue contours. Any arcs with measured redshifts are marked. Bottom Left - The resulting cluster member DS analysis. Bottom Right - Redshift histogram with the cluster members highlighted in red.

Table 3.14: Strength of Pearson Correlation

Pearson Value	Strength
0.00 - 0.19	Very Weak
0.20 - 0.39	Weak
0.40 - 0.59	Moderate
0.60 - 0.79	Strong
0.80 - 1.00	Very Strong

3.8 Dynamical Comparisons

In § 3.4 Δ/N was found to have lower jackknifed derived errors than for errors on the P value. This was in contrast to Hou et al. (2012) who found that the P value method was more robust and accurate, especially down to low cluster members ($N \approx 20$) than the Δ/N method. This might be due to the distribution of galaxies in many of these clusters being asymmetric. It was often the case that more galaxies were measured in one area than others, or completely lacking close into the cores of the clusters) and so high errors resulted from shuffling the velocities. This was because poorer regions had to sample a larger area of the sky in order to include enough nearest neighbours, diluting or increasing the signal by including outliers as a result. The Δ/N calculation did not shuffle the velocities across the galaxy positions and so was perhaps insensitive to this clustering effect (it was not a factor of the calculation).

The correlation calculated was based on calculated Pearson correlations (this was described again later for clarity in chapter 5), which returned a value between 1 and -1. Where 1 indicated a "very strong" positive correlation and a negative value indicated a negative correlation, 0 showed no correlation at all and in between there were various grades of correlation as set out in Table 3.14 (Evans, 1996).

This large selection of galaxy cluster redshifts, coupled with alternate cluster information from X-ray and strong lensing sources, provided a good opportunity for a quantitative test (rather than qualitative as in §3.7) comparison of dynamical activity in cluster cores. Five of the clusters had strong lensing constraints and all eight in the sample had shallow

X-ray observations (approximately 20 kilo seconds).

The DS statistical analysis provided the critical values method (Δ/N) and the DS test that used probabilities ($P-Value$). Previous investigations of the stability of substructure measurements from either of these values (Hou et al., 2012) found that for low numbers of cluster members ($N_{nn} \approx 20$), a higher level of false positives would result (substructure where there was none) and so $P - Value$ was preferred as it was more resistant to these effects if a sufficiently low value was selected. For the distribution of redshifts of clusters in this work, Δ/N seemed more robust and was be used instead.

In the top plot of Fig. 3.11 two values are compared against one another and it was found that they agreed with each other, with three of the eight clusters showing strong evidence of substructure. The Pearson correlation returned a value of -0.91 which indicates a "very strong" negative correlation (which was in line with expectations), but the large errors in the P values made this correlation meaningless as values on the y axis could move up and down the whole range of 0 to 1 for all but the most strongly disturbed clusters (as indicated by their P and Δ/N values). Due to the large P errors, Δ/N was compared with other cluster core measures, such as offset and cluster-scale halos (derived from strong lensing). These values are given in Tables 3.4 and 3.5.

The X-ray centroid offset from BCG centre provided a useful indication of dynamical disturbance (since it traced the hot intra-cluster gas), which was independent of the cluster selection method. Bottom left in Fig. 3.11, the Δ/N value against X-ray centroid offset showed no strong agreement and the Pearson correlation value showed only a "very weak" negative correlation with a high probability this was due to the sampling. This lack of correlation could be due to lack of cluster members in the centre of clusters. Despite the addition of REFLEX galaxies, the number of galaxy cluster members in the centre of the clusters remained low.

Strong lensing (SL) should provide the strongest constraints of substructure, but only to small scales (strong lensing images tend to only extend a few 10s of arcseconds from the cluster centre). Only five of the eight clusters have observed SL constraints and, apart

from two clusters, had no correlation with Δ/N (see bottom right in Fig. 3.11). The Pearson values showed only a "very weak" negative correlation, with a high probability this was due to the sampling enforcing the lack of significant correlation between them.

The two clusters that had disagreements between SL detecting substructure and the derived $P - value$ were A 3084 and A 368. For A 3084, this could have been because very few cluster members close into the BCG were measured (see the DS plot in Fig. 3.6) and, with the two halos being close together, this meant that at larger radii the cluster members behaved as if there was only one halo of their combined mass. For the other outlier A 368, the DS plot (see Fig. 3.5) showed large circles near the BCG centre and small circles at radii larger than $r \approx 100$ kpc. This was outside the constraints of the multiple image system close to the BCG. The cluster members at large radii were relatively few when compared to the BCG centre and so evidence of substructure outside the BCG (where the SL constraints placed only one DM halo) had insufficient data to be detected by the DS tests. Only one set of multiple images exist, which only constrained the single halo and provided no additional evidence or constraint to add a second. It could be that the SL constraints were just not sensitive to the second halo that was present.

3.8.1 Effects of REFLEX additions

To examine the idea that removing the information provided at the centre of the clusters (where the SL signals dominated) affected the DS tests' ability to detect substructure, the REFLEX additions were removed and the DS statistical tests re-calculated. The number of nearest neighbours was re-adjusted, where necessary, in order to increase the sensitivity of the test for the numbers involved. If the value was left as it was, the test would read a lower substructure because of smoothing effects.

The results (see Table 3.15) showed no strong effects by the inclusion or exclusion of the REFLEX additions. In some cases the indicator of substructure was increased or decreased, which either agreed or disagreed with the strong lensing result (but was not consistent with each). The results suggested that, while logically adding more galaxies

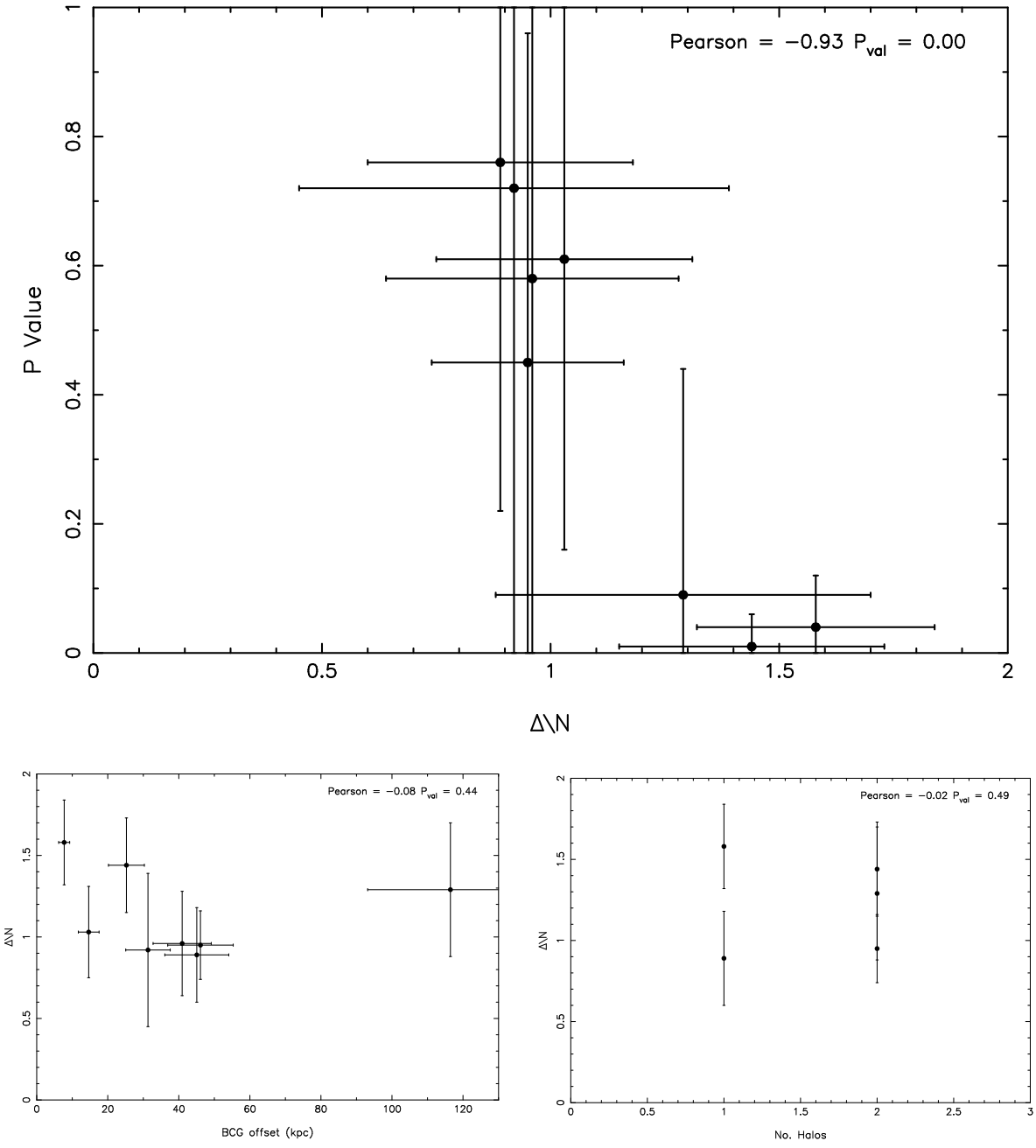


Figure 3.11: Top – Δ/N plotted against $P-Value$ to compare the two DS statistical tests of substructure evidence.Bottom Left – Δ/N plotted against X-ray centroid offset to compare the two values for dynamical evidence. Bottom Right – Δ/N plotted against number of DM halos from SL mass reconstructions. In each plot a Pearson correlation value and null hypothesis probability is given.

into the analysis would have improved results, the overall distribution of the galaxies played an effect too. This effect might also be what caused the resulting large errors in P values calculated (as discussed in § 3.4 & § 3.8). Another possible effect was that the removal of REFLEX additions accounted for a reduction in member numbers of $5\%-30\%$. This could push the number of cluster galaxy members to below the limit of $N \approx 20$: below this and the DS test would not pick up all substructure and would only be providing a lower limit (see Hou et al., 2012).

3.9 Chapter Discussion

In this section each of the analyses that were performed are discussed and then a comparison of DS statistics to other cluster values was made at the end..

The redshift histograms retrieved mean redshifts and calculated 1σ errors that were very close to the clusters' BCG literature values (see Table 3.4), except for the BCG excluded in A 3364. In this case the value was too close to the mean value for the other REFLEX galaxies of $z = 0.1489$, with $z_{hist} = 0.1484 \pm 0.0038$. A 2813 and A 2895 had differences of 0.003 and 0.022 respectively, but all values agreed well within 2σ errors. This would lend credence to the measured redshifts being members of the cluster.

The estimated velocity dispersion (see Fig. 3.1) showed no strong link to the $\sigma - T_X$ relation from Xue & Wu (2000). However due to the selection effects (a velocity dispersion and a cut with were defined), along with the simple method of calculating the velocity dispersion, did not prove that there was no link. In addition, the low numbers of galaxies measured might mean that the population simply was not sufficiently sampled.

This sampling effect, coupled with uneven coverage on the sky, meant that the Dressler Shectman test might suffer from biases during calculation. Any areas not covered by redshifts would render the DS tests blind to any substructure situated there, but also if the numbers were poorly distributed then the nearest neighbours might be non-ideally selected. Reducing the nearest neighbour count increased the sensitivity to substructure,

Table 3.15: DS results with/without REFLEX Additions

Name	With Reflex Additions				Without Reflex Additions				N_{halo}
	N	n_{nn} [b]	Δ/N	P[a]	N	n_{nn} [b]	Δ/N	P[a]	
Abell 2813	46	7	1.29 ± 0.41	$0.09^{+0.35}_{-0.09}$	38	6	1.37 ± 0.35	$0.04^{+0.20}_{-0.04}$	2
Abell 2895	20	4	0.89 ± 0.29	$0.76^{+0.24}_{-0.54}$	19	4	0.73 ± 0.24	$0.95^{+0.05}_{-0.23}$	1
Abell 368	36	6	1.58 ± 0.26	$0.04^{+0.08}_{-0.04}$	25	5	1.40 ± 0.26	$0.21^{+0.26}_{-0.20}$	1
Abell 3084	39	6	0.95 ± 0.21	$0.45^{+0.51}_{-0.45}$	37	6	1.01 ± 0.18	$0.31^{+0.39}_{-0.31}$	2
Abell 3088	22	5	1.04 ± 0.28	$0.61^{+0.39}_{-0.45}$	13	4	0.84 ± 0.46	$0.69^{+0.31}_{-0.69}$	-
Abell 3364	36	6	0.96 ± 0.32	$0.58^{+0.42}_{-0.58}$	31	6	1.04 ± 0.29	$0.35^{+0.58}_{-0.35}$	-
Abell 2537	51	7	1.44 ± 0.29	$0.01^{+0.05}_{-0.01}$	50	7	1.41 ± 0.27	$0.02^{+0.07}_{-0.02}$	2
RXCJ 0528.2	17	4	0.92 ± 0.47	$0.72^{+0.28}_{-0.72}$	15	4	1.12 ± 0.45	$0.24^{+0.46}_{-0.25}$	-

[a] Each test used 10,000 Monte-Carlo Shuffles to return a value.
[b] n_{nn} states the number of nearest neighbours used in the DS test calculations.
[c] N states the total number of galaxy cluster members used in the DS test calculations.

but below 20 cluster members the tests could only provide a lower limit on the clusters' substructure. Removing REFLEX galaxy members affected the DS statistics, but not in a consistent way, and so might be well within the errors of the two for the few numbers removed in many cases.

Other observable in the clusters revealed no strong correlation between any of them and the calculated DS statistics. Comparisons were dependent on direct linkages between observable values and merger activity. Substructure, that was tested for using DS statistics, would be present during a merger and an offset in the gas and BCG would be expected. The merging time for substructure to finish merging and hence decrease, and the BCG/X-ray centroid offset to fall to zero might be vastly dissimilar causing any correlation between the two to be held only for certain time periods. Additionally the DS test might not be sensitive to the cores of clusters due to the small number of redshifts in the region where the offset was.

The results of this test comparison of the dynamical status of cluster cores showed that the DS test could provide a good look at the centre of clusters if there were sufficient cluster members in the centre of the cluster. Without information in the centre of the clusters the DS tests (and statistics) were blind to substructure situated there. Also below a limit of 20 members, substructure measures were only a lower limit.

It was possible that this sample's galaxy density per cluster was too low to draw definite conclusions and required expansion and greater numbers of galaxies. The use of redshifts to look at substructure, as a free side effect from constraining cluster members and searching for strong lensing spectrum, showed some use providing qualitative information for priors when beginning mass models. In many cases it would be a very useful guide as to the number and position for halos. Deciding on priors was one of the big problems when building lensing models. If a model was too complex or physically unrealistic and poorly constrained, with poorly chosen parameters, then it would yield poorly fitting models. Understanding the physical interpretation of these models would become difficult.

3.10 Chapter Summary

In summary the analysis revealed the following:

(i) The masks on average produced a new object redshift for $\sim 65\%$ of the mask slits (see Table 3.3).

(ii) The literature values for the BCG (bar one exception) all laid within the calculated galaxy clusters 1σ errors (calculated from the sample, see Table 3.4) and within 2σ for A 2813 & A 2895.

(iii) Low cluster member statistics increased the chances of false positives for substructure. Larger errors appeared in P values than Δ/N, which was the opposite of that found by Hou et al. (2012). Future work should seek whether the cause was a low numbers and/or a variable sky galaxy density effect.

(iv) The need to prioritize mask slits over possible strong lensing images meant that the central area was devoid of redshifts, leaving the DS tests blind to these areas.

(v) A larger sampling for these clusters was needed over an even area of the sky, with emphasis on A 368 and A 3084 to see if their $P - Values$ changed significantly to match the SL data.

(vi) SL was only sensitive to substructure on small scales, while DS tests were sensitive to whatever scale the redshifts are distributed across sufficiently. These provide a useful indication of how far substructure extended.

(vii) DS tests could provide useful prior information when modelling clusters on the number and position of DM halos.

(viii) Comparing DS tests to other information showed quantitatively that there was no strong link to BCG offset or SL data correlating. However, this could have been due to small sample size or other effects. DS tests could be qualitatively used to assist in selecting strong lensing priors however.

CHAPTER 4

ABELL 3084 STRONG LENSING ANALYSIS

This chapter is based upon a paper in preparation. The senior author is Paul E. May with Graham P. Smith, Johan Richard, Alastair J. R. Sanderson, Arif Babul, Pasquale Mazzotta, Alastair C. Edge, Keelia R. Scott, Tim J. James, Victoria Hamilton-Morris, Tom A. Targett, and Eiichi Egami as co-authors.

Abell 3084 was picked from the sample analysed in this thesis for more in depth focus due to the serendipitous discovery of the location of the multiple images when constructing spectroscopic masks for redshift determination of any possible strong lensing arcs. The close arrangement of the multiple images to the BCG centre was unusual and a literature search yielded only one other example (Abell 1703) from Limousin et al. (2008). This image configuration has been classified as a "Hyperbolic Umbilic Catastrophe" (see § 4.1), which Orban de Xivry & Marshall (2009) determined to be a rare occurrence with only $\sim$ 1 expected to be observed per all-sky survey. The predicted rarity and the multi-wavelength data available made Abell 3084 a prime candidate for further investigation and analysis in this thesis.

4.1 Hyperbolic Umbilic Catastrophe

The Hyperbolic Umbilic Catastrophe is a unstable singularity in the caustic plane, with a typical shape of caustic that reproduces the quintet image system observed in A 1703 and

A 3084. Critical image lines located in the image plane are generally smooth, but tracing the critical lines using lensing equations onto the source plane yields the caustic lines (which are not necessarily smooth). The caustics can be calculated for any slice of redshift between the image and source plane. The caustic planes on these different redshift slices can evolve with redshift unless they are stable (if so, they are a constant shape throughout the whole redshift range between image and source redshifts). Typically caustics are composed of cusps joined by fold lines and are stable singularities, but other singularities exist where the Jacobian (see § 2.2.2, equation 2.19) becomes zero and the magnification becomes formally infinite. In reality it has a finite magnification due to finite object size and geometric optics breakdown. An unstable singularity, like the hyperbolic umbilic catastrophe, might only exist at certain points in redshift space or for a narrow range of it. This enables strong constraints to be put on the possible lens configurations capable of reproducing it and some singularities are capable of producing very strong magnifications. This can be useful observationally for spectroscopic or observational studies.

An example of the caustics produced by an elliptical mass distribution as a function of source redshift are shown in Fig. 4.1 taken from Orban de Xivry & Marshall (2009) and as can be seen the caustic shape persists over quite a wide range of redshifts. The caustics for the mass reconstruction in A 3084 matched this shape very well, as can be seen in the top plot of Fig. 4.2. The mass distribution used by Orban de Xivry & Marshall to recreate their multiple images was a non-singular isothermal ellipsoid density profile (NIE. as described by Kormann et al., 1994) that used a velocity dispersion of $\sim 1200\ \mathrm{km\,s^{-1}}$, core radius of ~ 30 kpc, and an ellipticity ~ 0.2. The NIE profile while similar to is not exactly comparable to a PIEMD and suffers from a drawback whereby isothermal lenses have infinite total mass (they extend forever), unless the mass density drops faster than r_{core}^{-2} (Kormann et al., 1994). This is the reason for the cut off radius in PIEMD models in order to keep the total mass from being infinite.

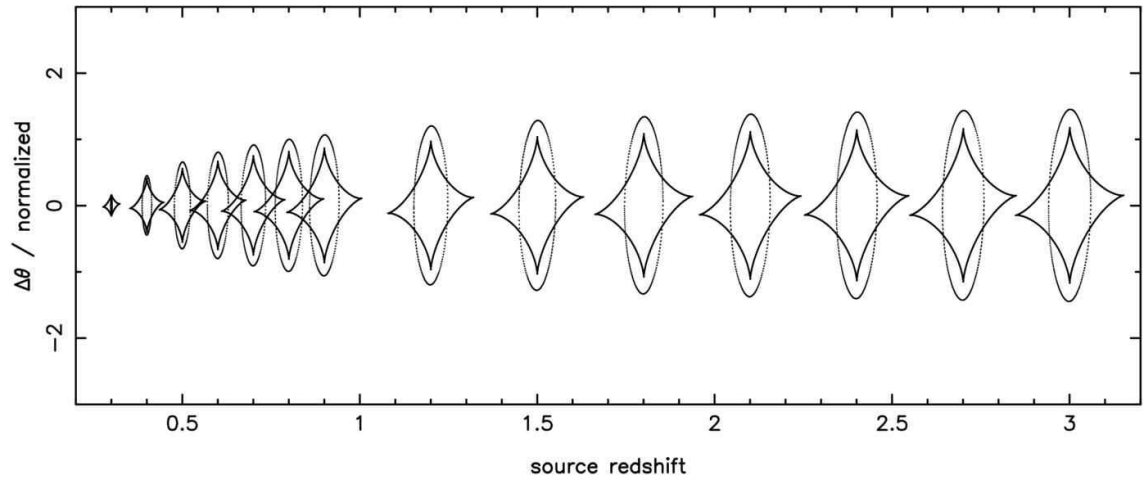


Figure 4.1: Figure 29 from Orban de Xivry & Marshall (2009) showing how the caustics for an elliptical cluster lens change as a function of source redshift.

4.2 Observations and Analysis

4.2.1 *Hubble Space Telescope* imaging

A 3084 was observed on February 2^{nd} 2007, with the *Hubble Space Telescope (HST)*[1] using the Advanced Camera for Survey (ACS). The observations totalled 1.2 ksec through the F606W filter (PID: 10881, PI: G. P. Smith). These data were reduced using standard MULTIDRIZZLE (Koekemoer et al., 2006) routines onto a final pixel scale of $0.03''$. The reduced frame revealed two candidate multiple image galaxies A1 and A2 (Fig. 4.2). Four images comprised of A1 were found by visual inspection of the reduced *HST* frame – A1.1, A1.2, A1.3, A1.5. All four images comprised two components, with the relative positions of the brighter ("a") and fainter ("b") blobs obeying the expected symmetry of a strongly-lensed galaxy. This image configuration matched that of a hyperbolic umbilic catastrophe (Orban de Xivry & Marshall, 2009), of which only one other was currently known in the literature (Limousin et al., 2008).

Limousin et al. (2008) studied Abell 1703 which had the same central circular grouping of four images (source image was at $z = 0.8885 \pm 0.0002$), with a fifth image located further

[1]Based in part on observations with the NASA/ESA *Hubble Space Telescope* obtained at the Space Telescope Science Institute, which was operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5-26555.

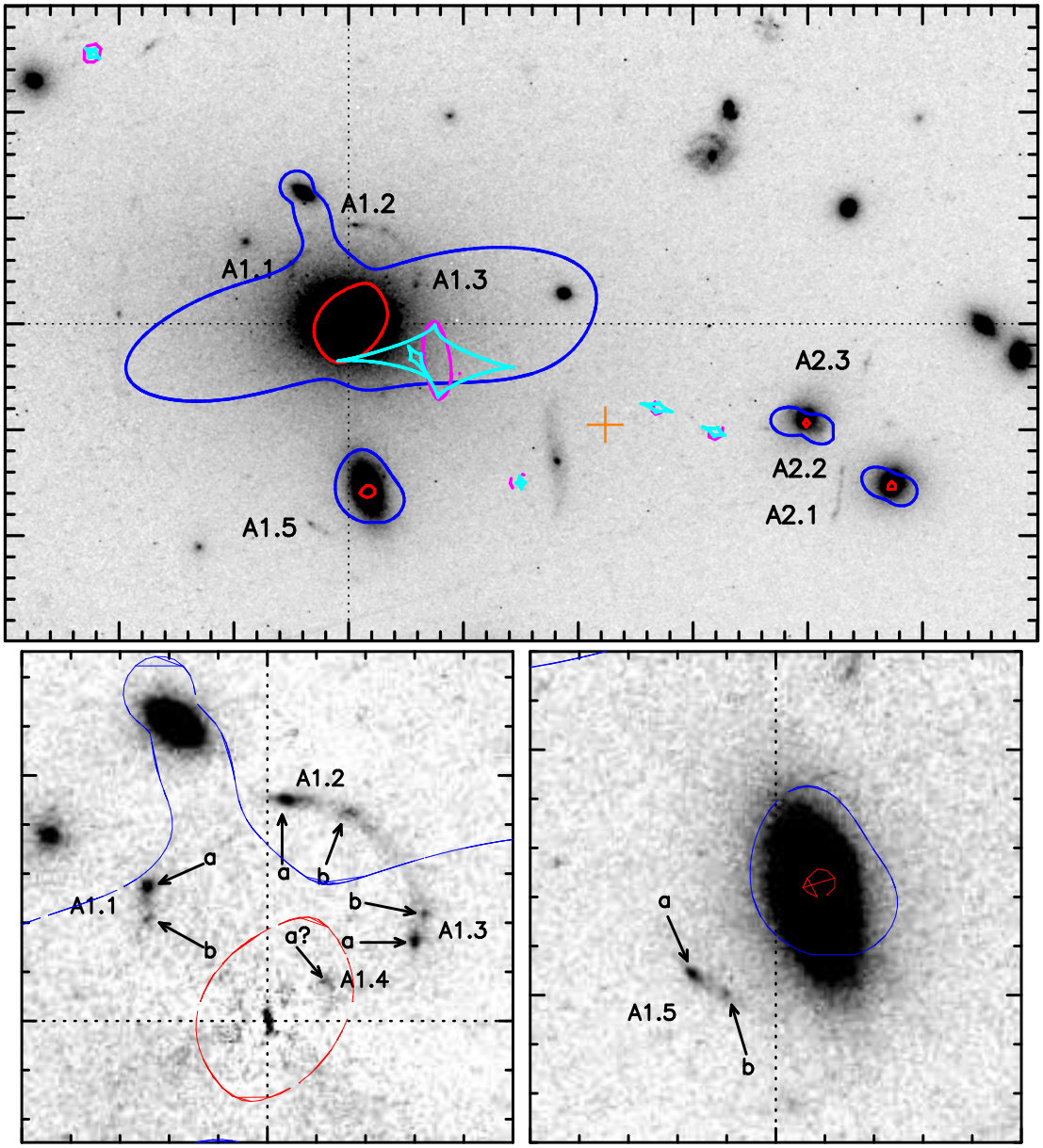


Figure 4.2: Top – The central 30×45 arcsec of A 3084, taken from the *HST*/ACS snapshot observations. Three of the five images comprising the hyperbolic umbilic catastrophe image configuration of the galaxy A1 were located adjacent to the BCG. The fourth image, A1.4, was detected after subtraction of the BCG (lower left panel), and the fifth image lay 10 arcsec South of the BCG. A second multiply-imaged galaxy laid $\sim$ 22 arcsec West of the BCG, marked A2.1, A2.2, and A2.3. Lower left – Zoom into the central 10×10 arcsec, after subtraction of the BCG. The brighter and fainter component (a, and b respectively) of A1.1, A1.2, and A1.3 were clearly visible. A candidate fourth image (A1.4) was detected following the subtraction of the BCG. Lower right – Zoom into the 10×10 arcsec region South of the BCG where the fifth image, A1.5, was found (two components visible). Note that all four of the securely detected images obeyed the mirror symmetry expected of strongly-lensed images. The dotted lines are centred on the BCG in all panels. Blue (red) lines are the tangential (radial) critical lines. The caustic lines for the mass distribution are shown in light blue (pink) for the tangential (radial) lines. North is up and East is left in all panels.

away to the East, similar to Abell 3084 (fifth image to the South), but the differences were that the multiple images were offset from the cluster BCG, and had two contributing galaxies within their scribed circle. Their analysis indicated a cluster potential that was described well by one DM halo modelled by a NFW potential and that the cluster was relaxed. In Orban de Xivry & Marshall (2009) they also were able to reproduce the image configuration with a single DM halo described by a NIE (Non-singular Isothermal Elliptical) potential Kormann et al. (1994). The existing analysis of Abell 1703 provided useful information on how to parameterise and model Abell 3084, suggesting a single halo with an ellipticity of around ~ 0.2 was capable of reproducing the multiple images.

A hyperbolic umbilic catastrophe actually comprises five images, suggesting that there may be a fifth, centrally located, image hidden by the BCG. Therefore the ELLIPSE task in IRAF was used to fit a model to the BCG light distribution, masking out objects such as likely cluster galaxies from the fit. ELLIPSE works by fitting elliptical isophotes to the galaxy image. It does this by tracing elliptical paths and analysing the one-dimensional harmonics of the surface brightness as a function of the ellipse tracing angle over several iterations of increasing fixed semi-major axis lengths. The model was then subtracted from the data. The resulting "BCG-subtracted" frame revealed a candidate fifth image, which was interpreted as the brighter component of A1.4, with the fainter component presumed lost in the fit residuals within ~ 1 arcsec of the centre of the BCG. It should be noted that the main strong-lensing results presented in §4.3 were insensitive to whether or not A1.4 was included in the constraints on strong lensing mass models.

The roughly linear arrangement of A2.1/A2.2/A2.3 (in an almost straight line) implied that the mass distribution in the cluster core was elongated in the direction orthogonal to the A2 system, and likely bi-modal. Both of these possibilities were studied (see §4.3.1).

4.2.2 Ground-based near-infrared imaging

A 3084 was observed through the J- and K-band filters (1 ksec each in FWHM = $0.9''$ seeing) with the ISPI near-infrared camera mounted on the Blanco 4-m telescope at the

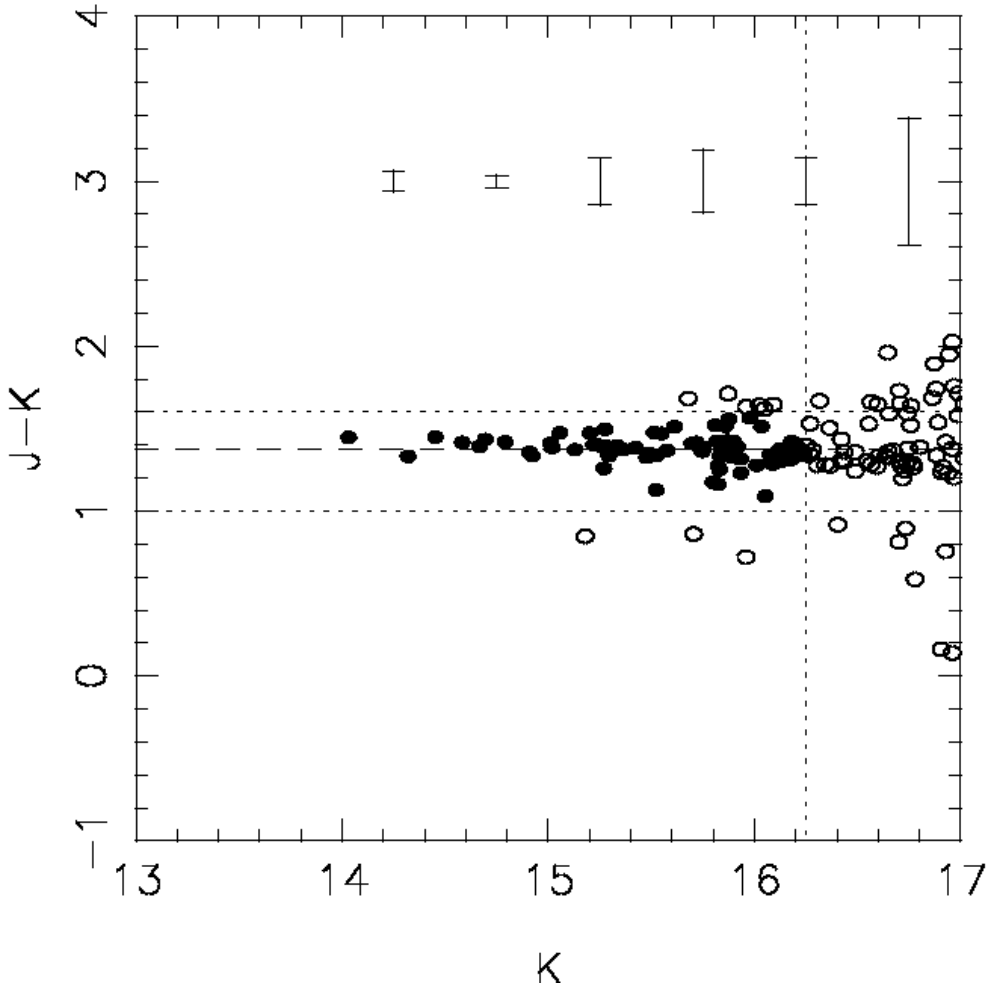


Figure 4.3: $(J-K)/K$ colour-magnitude diagram of galaxies within a 10×10arcmin field of view centred on the BCG in A 3084, revealing a tight red sequence ridge-line of likely cluster galaxies. The dotted lines mark the selection criteria applied to construct a catalog of cluster members for inclusion in the model of the cluster mass distribution in §4.3 and near-infrared luminosity density map, at right. The galaxies that satisfied this selection function, $1.0 < (J-K) < 1.6$ and $K < K^{\star} + 1 = 16.25$, were plotted as filled symbols. The error bars at the top of the figure show the mean error on $(J-K)$ as a function of K-band magnitude.

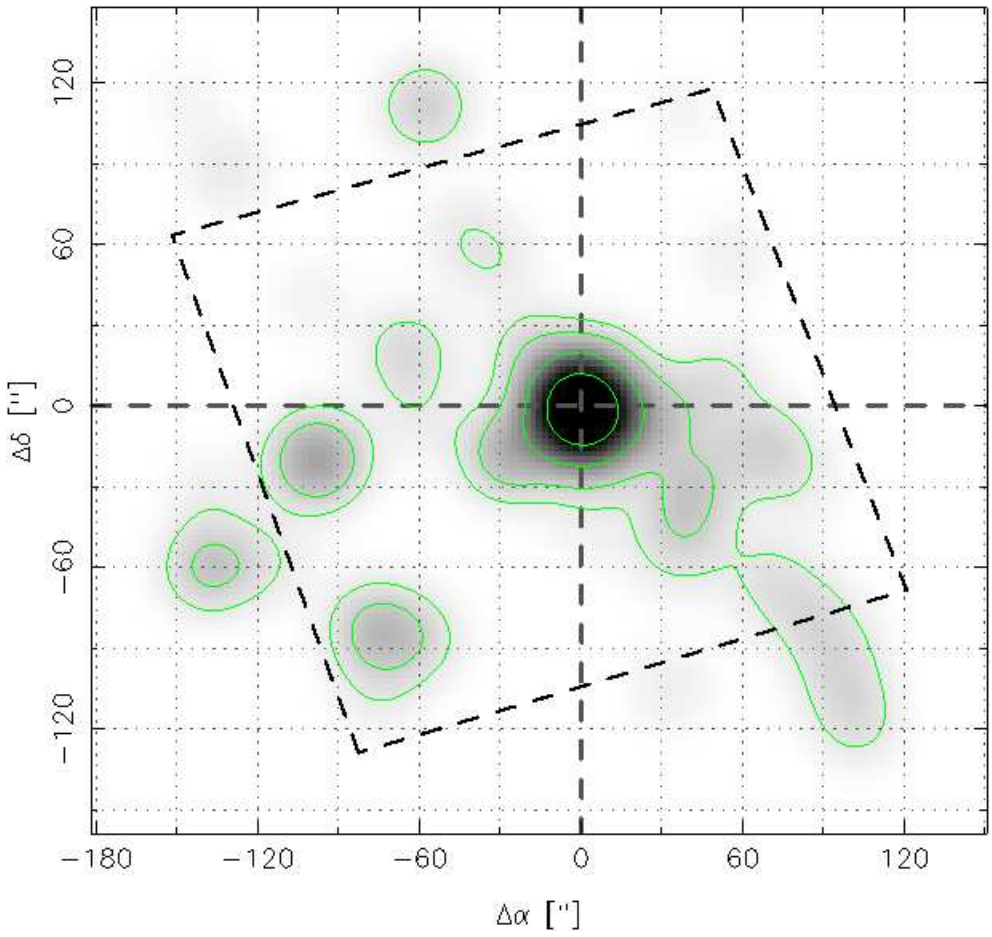


Figure 4.4: The rest-frame K-band luminosity density map of the core of A 3084 (grey-scale and green contours), based on the galaxies selected at left. The map is centred on the BCG, and the dashed box marks the ACS field of view. This panel shows the central $\sim$ 1Mpc $\times$ 1Mpc of the cluster in projection on the sky. North is up and East is left. Figure produced by Victoria Hamilton-Morris.

Cerro Tololo Inter-American Observatory[1] on February 11^{th} 2007. The data were reduced in a standard manner, using an automated pipeline of tasks in IRAF and calibrated photometrically and astrometrically to 2MASS[2] point source catalogues (Skrutskie et al., 2006) to root mean square (rms) precisions of 0.1mags and $0.1''$ respectively. The reduced frames were analysed using SEXTRACTOR, extracting sources that subtended > 4 contiguous pixels above 1.5σ/pixel, and the catalogues then merged. A red sequence of cluster galaxies was seen clearly in the $(J-K)/K$ colour-magnitude diagram (Fig. 4.3). Likely cluster members were therefore selected using the red sequence technique, as lying at $1.0 < (J-K) < 1.6$ and $K < K^{\star} + 1 = 16.25$, yielding a sample of 44 galaxies within the *HST*/ACS field of view. These limits were manually applied, based on the distribution in $J-K/K$ space, in order to select the overdensity which was the red-sequence within the cluster down to a K magnitude to remove poorly constrained and faint galaxies. K band was selected due to the better tracing of mass in this range than a bluer band and also due to the relative independence of redshift dimming with morphological type (k-correction) Loveday (2000). The K-band luminosity density map of the central $\sim 5 \times 5$arcmin of A 3084 (roughly the central 1 Mpc^2 in projection on the sky; Fig. 4.4) revealed that the luminosity density of the cluster core was centred on the BCG, with an extension to the South-West, and a number of prominent optical structures to the East.

4.2.3 Optical spectroscopy

A 3084 was observed with the FOcal Reducer and low dispersion Spectrograph 2 (FORS2; Appenzeller et al., 1998) on ESO's 8.2-m Very Large Telescope (VLT) on August 30^{th} 2008, through a single multi-slit mask. In order of priority this mask contained slits targeting (i) A1.2, A1.5, A2.3, (ii) candidate cluster galaxies from the sample defined in §4.2.2,

[1]Based in part on observations at the Cerro Tololo Inter-American Observatory, a National Optical Astronomy Observatory, which was operated by the Association of Universities for Research in Astronomy, under contract with the National Science Foundation.

[2]This publication makes use of data products from the Two Micron All Sky Survey, which was a joint project of the University of Massachusetts and the Infra-red Processing and Analysis Centre/California Institute of Technology, funded by the National Aeronautics and Space Administration and the National Science Foundation.

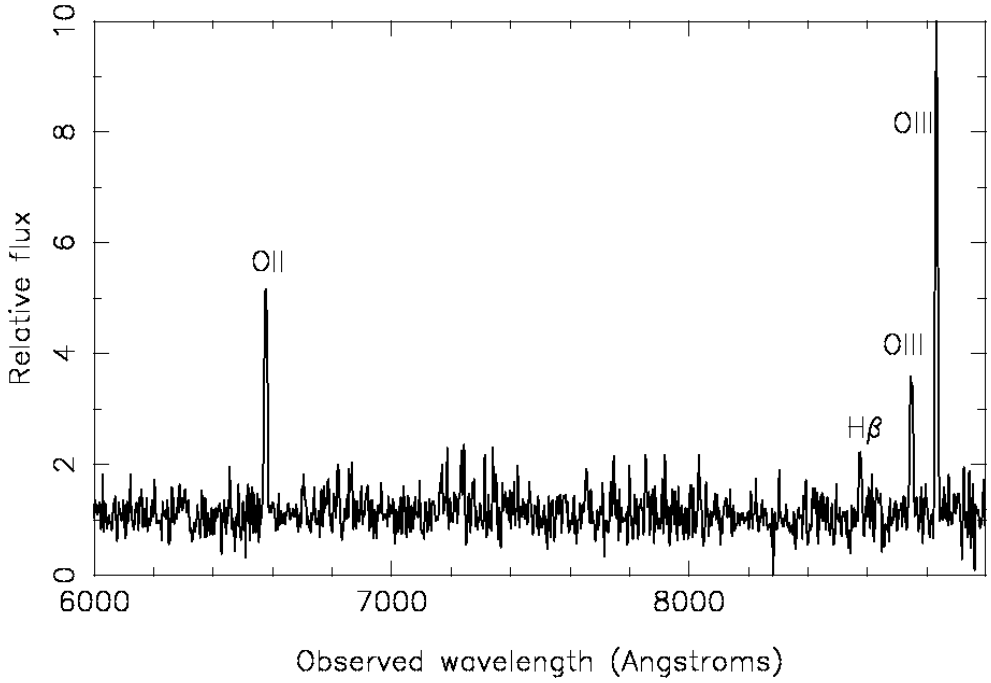


Figure 4.5: The reduced one-dimensional spectrum of A1.2, smoothed to the instrumental resolution of FORS2. Oxygen and hydrogen emission lines redshifted to $z = 0.764$ were detected at strong significance.

and (iii) other bright galaxies. The observations were performed with the 300V grism so as to maximize the observed wavelength range ($3500 \lesssim \lambda_{\rm obs} \lesssim 9000$ Å). These were integrated for a total of 3.6 ksec, split into four equal exposures, at a typical airmass of 1.04, through reasonably transparent sky, with seeing of FWHM $\simeq 1.5''$. These data were then reduced in a standard manner using Kelson's (2003) python scripts. Observations through two further masks were acquired with the Gemini Multi-Object Spectrograph (GMOS) mounted on the 8.1-m Gemini South telescope[1] on December 25^{th} 2008 using the $GG455\_G0329/R400+\_G5325$ filter/grating combination. These data were reduced using the GMOS package in IRAF following the standard reduction cookbook (see §3.1.2 for further details), with additional steps for cosmic ray rejection and frame combination.

The reduced 1-dimensional FORS2 spectrum of A1.2 and A1.5 contained prominent

[1]Based on observations obtained at the Gemini Observatory, operated by the Association of Universities for Research in Astronomy, Inc. under a cooperative agreement with the NSF on behalf of the Gemini partnership: the National Science Foundation (United States), the Science and Technology Facilities Council (United Kingdom), the National Research Council (Canada), CONICYT (Chile), the Australian Research Council (Australia), Ministério da Ciência, Tecnologia e Inovação (Brazil) and Ministerio de Ciencia, Tecnología e Innovación Productiva (Argentina).

Table 4.1: Spectroscopically Confirmed Galaxies at $0.2 < z < 0.23$ – "Members"

α, δ [J2000] (degrees)	K	Redshift	Spectrum features / Comment
45.94542, −36.92536	15.71 ± 0.07	0.2180 ± 0.0003	Hγ, Ca H & K, G
45.95500, −36.93086	16.17 ± 0.08	0.2120 ± 0.0005	Hγ, Ca H & K, G
45.96083, −36.93664	15.89 ± 0.07	0.2149 ± 0.0003	Hγ, Ca H & K, G, Hβ
45.96875, −36.94078	15.50 ± 0.06	0.2098 ± 0.0004	Hγ, Ca H & K, G, Hβ, Mgb, 5268, NaD
45.97000, −36.94761	15.59 ± 0.06	0.2180 ± 0.0003	Hγ, Ca H & K, G, Hβ, Mgb, 5268, NaD
45.98000, −36.97336	15.26 ± 0.05	0.2223 ± 0.0004	Hγ, Ca H & K, G, Hβ, Mgb, 5268, NaD
45.98208, −36.96517	16.40 ± 0.09	0.2186 ± 0.0004	Hγ, Ca H & K, G, Hβ, Mgb, 5268, NaD
45.98250, −36.94367	16.46 ± 0.10	0.2182 ± 0.0004	Ca H & K, G
45.98689, −36.96264	15.30 ± 0.06	0.2095 ± 0.0000	Hβ, H*alpha*, NII, SI, SII
45.99208, −36.94481	14.91 ± 0.05	0.2116 ± 0.0003	Hγ, Ca H & K, G, Hβ, Mgb, 5268, NaD
45.99583, −36.98389	18.05 ± 0.23	0.2106 ± 0.0001	[OII], [OIII]
45.99833, −36.93817	15.29 ± 0.05	0.2180 ± 0.0004	Hγ, Ca H & K, G
46.00292, −36.91003	16.27 ± 0.09	0.2085 ± 0.0010	Hγ, Ca H & K, Hβ
46.00458, −36.96967	15.75 ± 0.07	0.2179 ± 0.0003	Hγ, Ca H & K, G, Hβ, Mgb, 5268, NaD
46.00958, −36.99172	16.20 ± 0.09	0.2172 ± 0.0007	Hγ, Ca H & K, G
46.01507, −36.97154	16.64 ± 0.10	0.2232 ± 0.0002	Ca H & K, G
46.01550, −36.95335	16.12 ± 0.08	0.2213 ± 0.0003	Ca H & K, G, Hβ
46.01647, −36.94083	12.63 ± 0.02	0.2177 ± 0.0002	Brightest cluster galaxy; Guzzo et al. (2009)
46.01833, −36.96619	17.18 ± 0.15	0.2099 ± 0.0003	[OII], Hβ, [OIII], Hα, NII, SI
46.01855, −36.96644	17.18 ± 0.15	0.2092 ± 0.0000	Hα,NII,SI,SII
46.02289, −36.93506	15.11 ± 0.05	0.2106 ± 0.0001	Hβ, Hα, NII, SI, SII
46.02724, −36.92609	15.75 ± 0.07	0.2203 ± 0.0006	G,Hβ, Mgb
46.03292, −36.90328	15.88 ± 0.07	0.2195 ± 0.0009	Hγ, Ca H & K, G
46.03333, −36.94803	16.72 ± 0.10	0.2129 ± 0.0023	Hγ, Ca H & K, G
46.03390, −36.98598	16.64 ± 0.11	0.2157 ± 0.0003	Ca H & K, G
46.03625, −36.96525	15.47 ± 0.06	0.2210 ± 0.0003	Hγ, Ca H & K, G
46.03667, −36.90925	14.78 ± 0.04	0.2135 ± 0.0004	Hγ, Ca H & K, G
46.03750, −36.99319	15.32 ± 0.06	0.2158 ± 0.0006	Hγ, Ca H & K, G
46.03980, −36.94205	16.78 ± 0.12	0.2113 ± 0.0002	Ca H & K, G, Hβ
46.04000, −36.96553	17.47 ± 0.16	0.2145 ± 0.0002	[OII], Hβ, [OIII], Hα, SI
46.04292, −36.97028	15.79 ± 0.07	0.2180 ± 0.0003	Hγ, Ca H & K, G
46.04917, −36.98267	18.54 ± 0.26	0.2220 ± 0.0015	Hγ, Ca H & K, G, Hβ, Mgb
46.05011, −36.94623	14.03 ± 0.03	0.2203 ± 0.0003	Guzzo et al. (2009)
46.05657, −36.95737	15.69 ± 0.07	0.2099 ± 0.0001	Hβ, Hα, NII, SI
46.06494, −36.95797	14.58 ± 0.04	0.2205 ± 0.0002	Ca H & K, G
46.06833, −36.96656	16.61 ± 0.10	0.2184 ± 0.0006	Hγ, Ca H & K, G, Hβ
46.07292, −36.95119	16.93 ± 0.12	0.2210 ± 0.0003	Hγ, Ca H & K, G
46.07340, −36.95141	16.93 ± 0.12	0.2211 ± 0.0004	Ca H & K, G
46.07469, −36.93929	16.68 ± 0.11	0.2174 ± 0.0004	Ca H & K, G

Table 4.2: Other Spectroscopically Confirmed Galaxies – "Non-members"

α, δ [J2000] (degrees)	K	Redshift	Spectrum features / Comment
45.94083, −36.92447	15.98 ± 0.08	0.3136 ± 0.0002	[OII], Hβ, [OIII]
45.94542, −36.91825	17.09 ± 0.14	0.0966 ± 0.0010	[OII], Hα, SII
45.98354, −36.97320	15.73 ± 0.07	0.3131 ± 0.0015	Ca H & K
45.98542, −36.89261	17.62 ± 0.19	0.1904 ± 0.0002	[OIII], Hα, NII, SI, SII
45.99609, −36.91133	17.21 ± 0.14	0.4227 ± 0.0096	Ca H & K
45.99725, −36.95828	17.27 ± 0.15	0.1353 ± 0.0001	Hβ, [OIII], Hα, NII
46.00870, −36.94147	15.89 ± 0.08	0.4916 ± 0.0044	[OII], Hβ, [OIII]
46.00972, −36.94597	17.74 ± 0.18	0.1203 ± 0.0043	Ca H & K
46.01292, −36.99822	15.29 ± 0.06	0.2861 ± 0.0006	Hγ, Ca H & K, G
46.01341, −36.99874	15.29 ± 0.06	0.2853 ± 0.0003	Ca H & K
46.01516, −36.97007	17.76 ± 0.19	0.1755 ± 0.0001	[OIII]
46.01791, −36.96432	17.57 ± 0.17	0.6265 ± 0.0302	[OII], [OIII]
46.02500, −36.98658	15.78 ± 0.07	0.0011 ± 0.0002	Hγ, Ca H & K, G
46.02602, −36.99017	15.86 ± 0.07	0.4655 ± 0.0003	Ca H & K, G
46.02602, −36.99017	15.86 ± 0.07	0.4661 ± 0.0009	Ca H & K
46.02790, −36.91172	16.74 ± 0.12	0.4152 ± 0.0024	Hα, NII, SI
46.05076, −36.93889	17.12 ± 0.13	0.4932 ± 0.0009	Ca H & K, [OII]
46.05485, −36.94378	17.14 ± 0.14	0.5679 ± 0.0002	Hγ, Ca H & K, G
46.05737, −36.95364	16.56 ± 0.10	0.3889 ± 0.0001	Ca H & K, [OII]
46.05771, −36.92856	16.66 ± 0.11	0.2955 ± 0.0013	G, NaD
46.05771, −36.92856	16.66 ± 0.11	0.4942 ± 0.0002	[OII], Hβ
46.05840, −36.93579	17.20 ± 0.14	0.5675 ± 0.0011	Ca H & K
46.06366, −36.95443	17.87 ± 0.20	0.2564 ± 0.0024	Ca H & K, G
46.06667, −36.98258	17.30 ± 0.15	0.3117 ± 0.0010	Hα, NII, SI, SII
46.06833, −36.97925	17.43 ± 0.16	0.3333 ± 0.0001	Hγ, Ca H & K
46.06905, −36.92422	13.77 ± 0.03	0.7224 ± 0.0004	Ca H & K
46.07225, −36.95415	16.99 ± 0.13	0.4304 ± 0.0152	G, Hβ
46.08018, −36.94111	17.92 ± 0.21	0.4203 ± 0.0121	G, Mgb, Hβ

oxygen emission lines redshifted to $z = 0.764$, thus confirming that A1 was an imaged background galaxy (Fig. 4.5). A faint continuum was detected from A2.3, but no convincing spectral lines or breaks. In total from the combined FORS2/GMOS dataset 66 galaxy redshifts were obtained, of which 37 laid in the range $0.2 < z < 0.23$, and are hereafter referred to as "members" (Table 4.1; Fig. 4.6), and 29 laid at lower or higher redshifts (Table 4.2). Two further redshifts were added from the literature (Guzzo et al., 2009) to the newly confirmed members listed in Table 4.1, one of which was the BCG.

A mean cluster redshift of $z = 0.2161 \pm 0.0045$ was measured and a cluster velocity dispersion of $\sigma = 1358 \pm 118$ km s^{-1}. However, the redshift distribution of members appeared to be bi-modal or asymmetric (Fig 4.6), with peaks separated by ~ 2400km s^{-1}. This suggested that the measured velocity dispersion of this cluster might be inflated by dynamical activity along the line-of-sight.

To investigate the dynamical structure of the cluster further, each member was plotted as a circle, the radius of which scaled with the deviation of the local kinematics from the global kinematics following Dressler & Shectman (1988, Figure 4.7) and the method outlined in 3.3. The calculation was based on using the six nearest neighbour galaxies, using the procedure in Hou et al. (2012), rather than the original value of 10 outlined by Dressler & Shectman. This was carried out in order to sensitize the test to substructure when low member numbers are used. From this calculation various DS statistics (see §3.3 for the calculation method) yielded $\Delta/N = 0.95$ and a null hypothesis probability value of $P = 0.45$. A value above 1 for Δ/N normally indicates the presence of substructure in the cluster, with A 3084 having a calculated probability value of 45% for this value being a product of the sampling of the population. Neither of these values indicated the strong presence of substructure but, as was discussed further in Chapter 3, the data were missing redshifts from the central cluster area of interest and consequently were only providing statistical information from the edges of the cluster.

The strongest evidence for departures from the global cluster kinematics was found to the South East of the BCG in Figure 4.7. When examining the distribution of calculated

δ and plotting in red those that were below the median value of δ and others in black, the concentration of those above the median calculated values (and therefore some of the highest deviations from the global kinematics) were in the South East of the cluster. The peak (which is asymmetric, Fig. 4.6) for these lower deviated galaxies lay in the higher of the two redshift peaks; the redshift distribution of galaxies with larger deviations from global kinematics presented two peaks. However any statistical significance was moderate at best and, as stated in the previous paragraph, the lack of spectroscopic observations within 1arcmin of the BCG prevented firm conclusions on the dynamical structure along the line of sight through the cluster core. The best description was still one of the cluster being dynamically active along the line of sight, and that the BCG was associated with the higher redshift, dynamically less active of two peaks in the redshift distribution.

4.2.4 *Chandra* X-ray observations

A 3084 was observed with *Chandra*[1] on March 16^{th} 2008 (PID: 09800732; ObsID: 9413; PI: G. P. Smith) using ACIS-I in Very Faint mode. These data were reduced and analysed according to the procedure described in Sanderson et al. (2009). Briefly, the data were reprocessed using CIAO version 4.1 and incorporating CALDB version 4.1.2, to produce flare cleaned and point-source removed level 2 events files. No flares were found during the light curve cleaning, giving a total of 19.9 ksec of usable data. Corresponding blank sky background datasets were also produced and matched in normalization to the cluster events, to account for variations in the particle-dominated high energy background. No adjustment was made to the background to allow for any variation in soft Galactic foreground emission compared to each cluster observation. However, following Sanderson et al. (2006), the galactic absorbing column was fitted as a free parameter in the spectral modelling to allow for any differences in inferred low energy absorption associated with soft emission excesses or calibration uncertainties; the best-fit values were found to be

[1]This research made use of data obtained from the Chandra Data Archive and the Chandra Source Catalog, and software provided by the Chandra X-ray Center (CXC) in the application packages CIAO, ChIPS, and Sherpa.

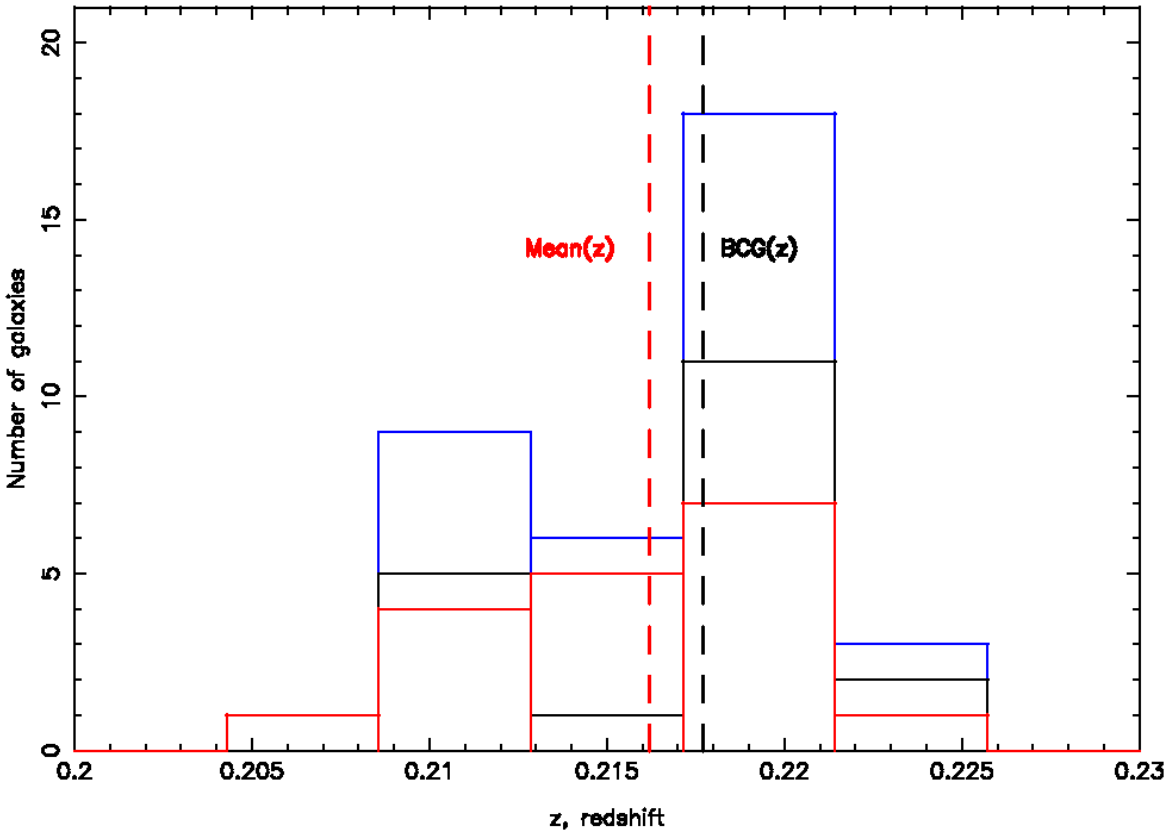


Figure 4.6: Redshift distribution of the 37 galaxies identified at $0.2 < z < 0.23$ from the multi-slit observations, plus the two galaxies from Guzzo et al. (2009). The red dashed line shows the mean redshift of $z = 0.2162 \pm 0.0044$; the black dashed line the measured BCG redshift of $z = 0.2177 \pm 0.0002$. The blue histogram shows the total number distribution, the red and black histograms the distribution of galaxies with low and high (respectively) local departure from the global cluster kinematics (see Fig. 4.7).

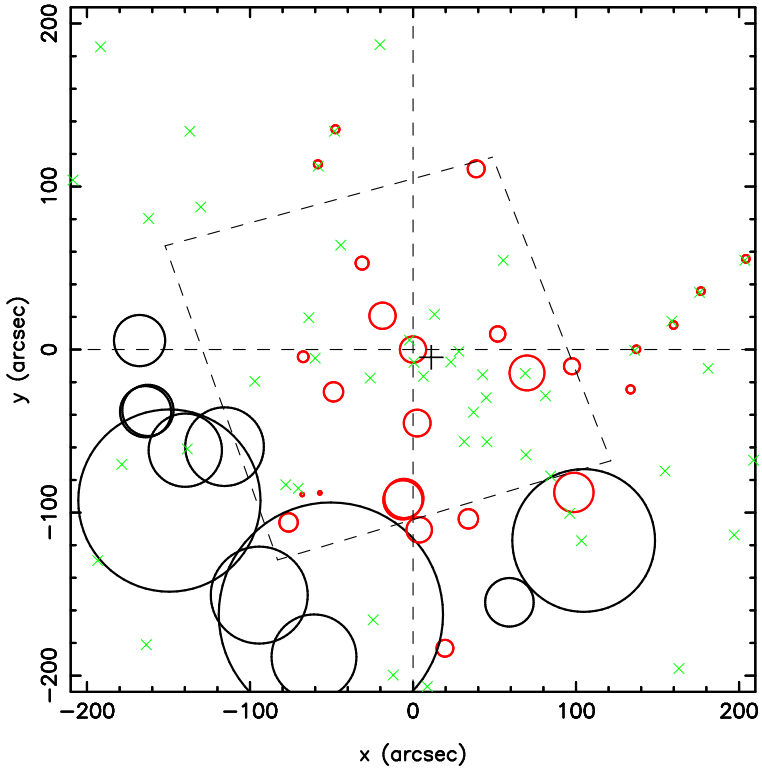


Figure 4.7: Spectroscopically confirmed cluster member galaxies plotted as circles of radius proportional to the departure of the local kinematics from global cluster kinematics, following Dressler & Shectman (1988). The figure is centred on the BCG and the plus symbol marks the X-ray centroid. Galaxies that satisfied the $(J-K)/K$ photometric selection described in §4.2.2 that were not spectroscopically confirmed as members are marked as small green crosses. Red and black circles are plotted separately in the redshift histogram shown in Fig. 4.6. The dashed box shows the ACS field of view. North is up and East is left.

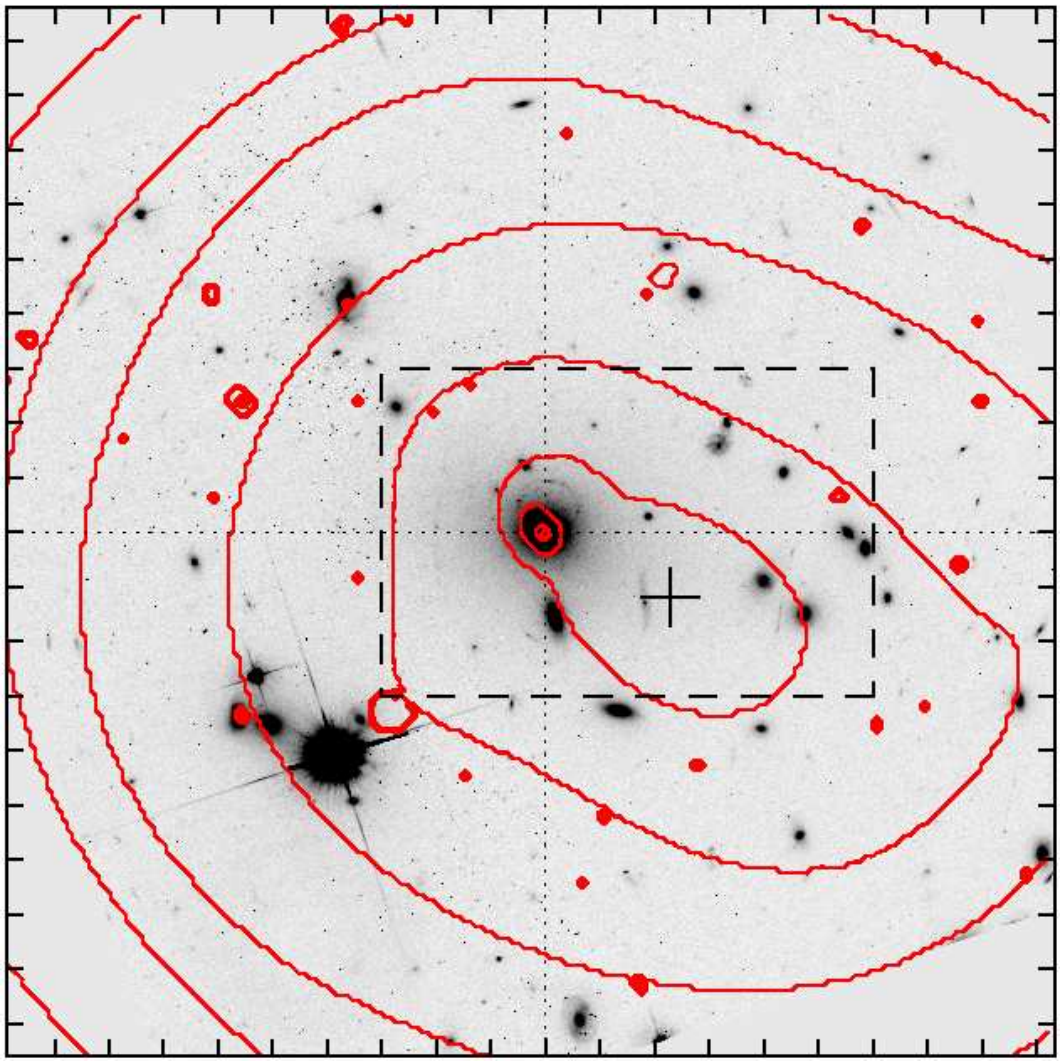


Figure 4.8: Smoothed X-ray surface brightness map based on the *Chandra*/ACIS-I observations (red contours) overlaid on the central 100×100 arcsec of the *HST*/ACS snapshot observation. The X-ray emission peak was within ~ 1 arcsec of the optical centroid of the BCG, whilst the X-ray centroid lay 11.1 arcsec West and 4.8 arcsec South of the BCG (black "plus" symbol). The dashed black box delineates the region shown in Fig. 4.2. Contours were spaced equally in the log, with each contour separated by 0.25dex. Each tick mark on the axes represents 5 arcsec. North is up and East is left.

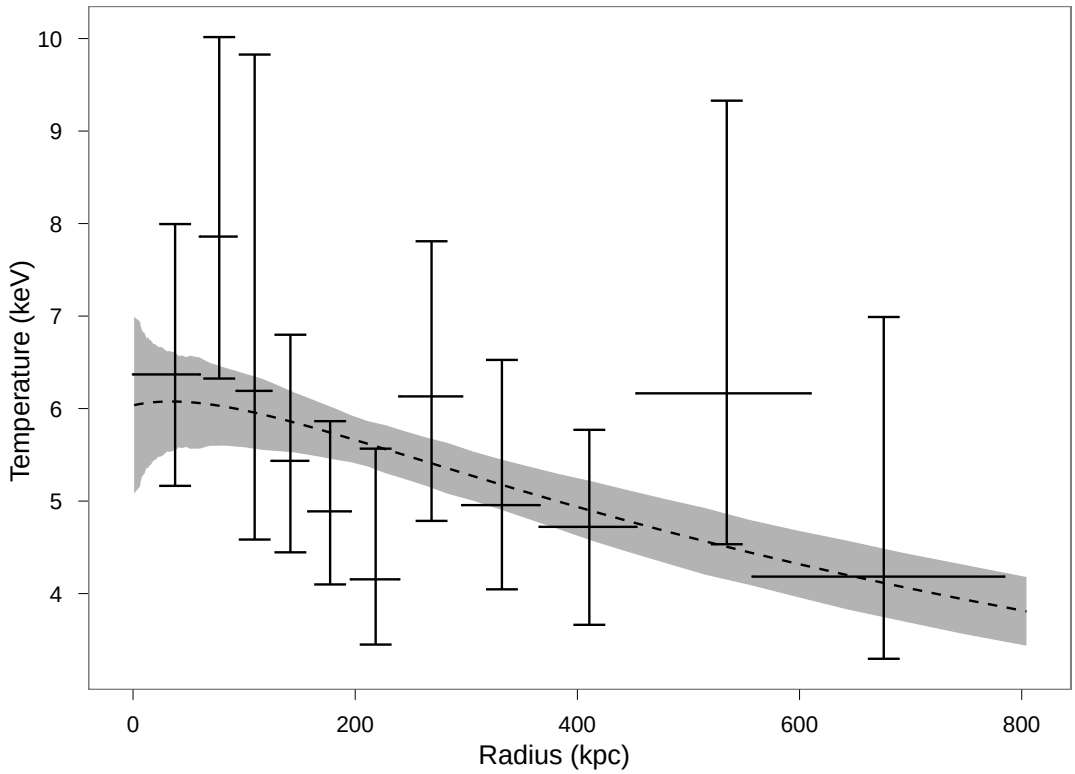


Figure 4.9: X-ray temperature profile of the intra-cluster medium, based on the *Chandra*/ACIS-I observations. The profile is centred on the X-ray centroid. The grey shaded area shows the 68% confidence interval around the best-fit model (dashed curve) described in §4.2.4. Figure produced by Alaistair Sanderson.

consistent with the HI inferred value of $2.32 \times 10^{20}\mathrm{cm}^{-2}$ from the FTOOLS task NH (based on the radio survey data of Kalberla et al., 2005). In the final reduced dataset, a total of ~ 5500 net cluster counts were detected in the energy range $0.5 - 7.0$keV.

To study the morphology of the hot gas, a smoothed X-ray surface brightness map was created using the wavelet reconstruction task WVDECOMP (Vikhlinin, 1998), employing the à trous wavelet kernel and the default parameter settings. The resulting wavelet smoothed image (Fig. 4.8) was used to identify the X-ray peak and centroid of the emission, as described in Sanderson et al. (2009). The X-ray peak laid within 1 arcsec of the optical centroid of the BCG, and the centroid 11.4 arcsec West and 4.8 arcsec South of the X-ray peak, i.e. offset 43 kpc from the BCG/peak in the same direction as the South-West extension of the K-band luminosity density map (Fig. 4.4).

Within a projected radius of 1 arcsec of the X-ray peak, just 13 photons were detected and it was not possible to perform a robust extension test on the X-ray peak. Indeed the data were formally consistent with the X-ray peak being an X-ray point source. Without deeper X-ray observations no conclusions could be made as to the origins of the X-ray peak other than as possibly due to nuclear activity inside the BCG. From here on the X-ray centroid was used as the centre of cluster emission and the X-ray peak no longer considered in discussions.

A mean temperature of $4.3^{+0.9}_{-0.6}$keV was measured in an annulus spanning $0.15-0.2r_{500}$, centred on the X-ray centroid, using the iterative method of Sanderson et al. (2006), and fitting an absorbed APEC[1] model in XSPEC version 11.3.2. APEC stands for "Atomic Plasma Emission Code", which calculates the emissivity for optically-thin and hot collisional plasmas (Foster et al., 2012). A spectral profile was extracted that comprised 11 annuli containing $500 - 1000$ net cluster counts in the energy range $0.5 - 7.0$keV, out to a maximum radius of $220''$ (corresponding to 770 kpc). These spectrum were fitted with an absorbed APEC model using the PROJCT scheme in XSPEC to yield de-projected gas temperature and density profiles. Weighted response matrix files were used for each

[1]http://www.atomdb.org/

spectrum, and element abundances measured relative to the abundance ratios of Grevesse & Sauval (1998). The temperature profile revealed no evidence for a cool core centred on the X-ray centroid (Fig. 4.9). The lack of evidence for a cool core centred on the centroid and the lack of photons on the X-ray peak to produce a similarly BCG centred temperature profile meant that no confirmation of A 3084 hosting a cool core could be made.

The phenomenological cluster model of Sanderson & Ponman (2010) was fitted to the data (see Sanderson & Ponman, 2010 and Ascasibar & Diego, 2008 for full details). In summary, the model described the total cluster mass distribution as a Hernquist (1990) density profile, and from this profile predicted the temperature and gas density profiles of the cluster, with the addition of four physical parameters: the central gas temperature, T_0, that was modified by a variable cool core component controlled by a parameter t ($0 < t < 1$), which becomes important inside a radius a fraction β ($0 < \beta < 1$) times the scale radius, and the gas density normalization, f, a fraction of the cosmic mean baryon fraction. The predicted temperature and gas density profiles were fitted to the observed profiles, and uncertainties on the fitted parameters obtained from Monte Carlo re-sampling of the error distributions of the underlying data to which the model is fitted. This model was used to measure the slope of the gas density profile at a projected radius of $0.04r_{500}$, following Vikhlinin et al. (2007), obtaining $\alpha = d\,ln\,\rho_{\rm gas}/d\,ln\,r = -0.36 \pm 0.05$ and the entropy of the intra-cluster gas within a projected radius of 20 kpc of the X-ray centroid of $S_0 = S(< 20\ {\rm kpc}) = (161 \pm 22)\,{\rm keV\,cm^2}$.

4.2.5 Summary

The main results of the analysis of the optical, near-infrared and X-ray observations of the core of A 3084 suggested the following:

(i) Two strongly-lensed galaxies were located within $\sim$ 20 arcsec of the BCG, one of which (A1; a hyperbolic umbilic) was spectroscopically confirmed at $z = 0.764$. The

morphology of other images (A2) indicated an elongated and possibly bi-modal mass distribution with a major axis aligned roughly North-East/South-West.

(ii) The rest-frame K-band luminosity density of the cluster core peaked on the BCG and presented an extension to the South-West that spanned $\sim$ 2arcmin (420 kpc at the cluster redshift) on the sky.

(iii) The redshift distribution of 39 spectroscopically confirmed cluster member galaxies was bi-modal (peak separation of $\sim$ 2400 $\mathrm{km\,s^{-1}}$), suggesting that the measured cluster velocity dispersion of $\sigma = 1360\ \mathrm{km\,s^{-1}}$ could have been inflated by dynamical activity along the line of sight. Taking the redshift catalog at face value suggested that the BCG belonged to a kinematically "cold" population, however, spectroscopic incompleteness within 1arcmin of the BCG precluded firm conclusions.

(iv) The X-ray centroid was offset 43 kpc to the South-West of the BCG. The X-ray emission presented no evidence of a cool core in this cluster.

4.3 Strong-lens Model and Results

The identified lensing constraint images were used (§4.2.1 & §4.2.3) to constrain a parametrized model of the projected mass distribution in the cluster core. The goal was to infer the structure of the total mass distribution, and thus of the dark matter distribution, and to compare that with the distribution of stellar mass (§4.2.2) and hot gas (§4.2.4). Examining the summary of results in §4.2.5, a bi-modal (two DM halo) model was expected to be required to fit the lensing constraints. This section (§4.3) examines this hypothesis and other possible fits in an attempt to determine the simplest model that describes the data best.

Table 4.3: Gravitational Lens Model Parameters[a]

Halo	Δx [b] (arcsec)	Δy [b] (arcsec)	ϵ	θ [c] (deg)	$r_{\rm core}$ (kpc)	$r_{\rm cut}$ (kpc)	σ_0 (km/s)	z_{A2}	$\ln(E)$ [d]	σ_i [e] (arcsec)
Model A								> 2.6	-67.0 ± 1.1	0.52
DM1	0.0	0.0	$0.50^{+0.11}_{-0.12}$	$5.1^{+3.0}_{-2.6}$	35^{+3}_{-2}	1000	712^{+35}_{-32}			
BCG	0.0	0.0	0.11	-12.9	0.5	99^{+34}_{-33}	202^{+27}_{-27}			
$L^\star$ galaxy					0.15	45	157			
Model B								$0.93^{+0.02}_{-0.03}$	-133.7 ± 0.4	0.84
DM1	11.2	-4.8	$0.86^{+0.03}_{-0.04}$	$-1.4^{+0.8}_{-1.1}$	97^{+2}_{-2}	1000	897^{+17}_{-17}			
BCG	0.0	0.0	0.11	-12.9	0.5	81^{+29}_{-25}	292^{+5}_{-6}			
$L^\star$ galaxy					0.15	45	157			
Model C								> 2.3	-98.1 ± 1.6	0.50
DM1	$3.0^{+1.1}_{-0.9}$	$0.4^{+0.1}_{-0.1}$	> 0.97	$-4.5^{+0.5}_{-0.5}$	$14.5^{+3.1}_{-3.0}$	1000	661^{+22}_{-27}			
BCG	0.0	0.0	0.11	-12.9	0.5	$73.9^{+42.3}_{-19.2}$	305^{+8}_{-10}			
$L^\star$ galaxy					0.15	45	157			
Model D								> 2.0	-40.3 ± 1.9	0.24
DM1	0.0	0.0	$0.59^{+0.07}_{-0.09}$	$+15.6^{+3.4}_{-4.2}$	$6.9^{+1.5}_{-0.8}$	1000	439^{+14}_{-22}			
DM2	11.2	-4.8	$0.79^{+0.13}_{-0.23}$	$-18.8^{+5.8}_{-4.4}$	$40.5^{+24.9}_{-3.6}$	1000	505^{+69}_{-72}			
BCG	0.0	0.0	0.11	-12.9	0.5	120^{+12}_{-55}	< 131			
$L^\star$ galaxy					0.15	45	157			
Model E								> 1.5	-55.2 ± 8.6	0.21
DM1	$0.3^{+0.3}_{-0.3}$	$0.2^{+0.3}_{-0.1}$	$0.76^{+0.05}_{-0.17}$	$+13.4^{+4.9}_{-5.3}$	$7.4^{+6.8}_{-3.1}$	1000	419^{+44}_{-32}			
DM2	> 6.1	< -2.6	> 0.13	$-20.7^{+10.5}_{-11.1}$	$49.4^{+30.8}_{-14.0}$	1000	562^{+115}_{-116}			
BCG	0.0	0.0	0.11	-12.9	0.5	102^{+27}_{-37}	< 270			
$L^\star$ galaxy					0.15	45	157			
Model F								> 1.5	-42.0 ± 2.5	0.24
DM1	0.0	0.0	$0.62^{+0.05}_{-0.08}$	$+12.6^{+4.5}_{-4.5}$	$7.6^{+1.0}_{-1.2}$	1000	447^{+23}_{-24}			
DM2	11.2	-4.8	> 0.65	$-6.3^{+5.7}_{-20.7}$	$49.7^{+23.3}_{-19.6}$	1000	382^{+236}_{-169}			
DM3 [f]	...	...	0.0	0.0	50.0	1000	301^{+206}_{-208}			
BCG	0.0	0.0	0.11	-12.9	0.5	99^{+33}_{-32}	< 134			
$L^\star$ galaxy					0.15	45	157			

[a] Parameters stated with error bars or as limits were free parameters, otherwise parameters were held constant.
[b] Offset from optical centroid of BCG, as measured using the *HST*/ACS data. Note, the X-ray centroid lies at $(11.2, -4.8)$.
[c] Orientation of major axis, positive clockwise from x-axis.
[d] The natural logarithm of the probability of the data, given the model (the Bayesian evidence). Lower is better.
[e] The rms deviation from the observations of the predicted image-plane positions of the multiple-images.
[f] The 95% confidence interval on the x and y coordinates of DM3 span the full range of the prior: $-30'' < x < 30''$, $-30'' < y < 30''$.

4.3.1 Model Fitting

The mass distribution was parameterised as a superposition of cluster- and galaxy-scale halos, informed by the analysis presented in §4.2, and summarised in §4.2.5. The models used have both single- and double-cluster scale halos; all models had 14 galaxy-scale halos, centred on those galaxies from the sample of 44 discussed in §4.2.2 that laid within 1arcmin of the BCG. All halos were parametrised as smoothly truncated pseudo elliptical isothermal mass distributions (PIEMD see §2.5.2) following Kneib et al. (1996), Smith et al. (2005), and Richard et al. (2010), and others. Each halo was therefore described by seven parameters: $\{x, y, \epsilon, \theta, \sigma_0, r_{\rm core}, r_{\rm cut}\}$, where the first four were the position, ellipticity and orientation on the sky respectively, and the remaining parameters were the central velocity dispersion, core radius and cut-off radius respectively. The position, ellipticity and orientation of the galaxy halos were matched to those of their light, based on analysis of the *HST*/ACS data. The velocity dispersion, core and cut-off radii were scaled with their luminosity, adopting scaling relations that use parameter values for an $L^\star$ galaxy as shown in Richard et al. (2010, see Table 4.3). The scaling relations work using the Faber-Jackson relation for an elliptical galaxy, with σ, r_{core} and r_{cut} determined by the following equations:

$$\sigma = \sigma^\star 10^{0.4\frac{m^\star - mag}{\sigma_{slope}}} \tag{4.1}$$

$$r_{core} = r^\star_{core} 10^{0.4\frac{m^\star - mag}{2}} \tag{4.2}$$

$$r_{cut} = r^\star_{cut} 10^{0.4\frac{m^\star - mag}{2}} \tag{4.3}$$

With $sigma^\star$, $m^\star$, σ_{slope}, $r^\star_{core}$ and $r^\star_{cut}$ being the velocity dispersion, $L^\star$ magnitude, velocity dispersion slope, core and cut radii (in arcseconds) respectively. The use of these relations, along with a magnitude for each cluster member galaxy, allowed a scaled PIEMD mass potential to be produced for each galaxy that contributed to the total cluster mass model being parametrised without running out of model constraints. Each model was

fitted to the data using the Markov Chain Monte Carlo sampler within LENSTOOL[1] (Jullo et al., 2007, see §2.4.3 for more details of its operation), using the positions and redshift of both components ("a" and "b" – see Fig. 4.2) of A1.1/1.2/1.3/1.5, component "a" of A1.4, and the positions of A2.1/2.2/2.3 as constraints. The redshift of A2 was a free parameter in all models. As described in §2.4.3, the likelihood of each step in the Markov Chains was set to be calculated in the image plane (instead of source plane), adopting a positional uncertainty of 0.2 arcsec.

Models containing one cluster-scale halo

All previous galaxy cluster strong-lens models contained a massive cluster-scale halo centred within a few arcseconds of the BCG. For every model tested a separate PIEMD was used to model the BCG component and the cluster member galaxies were modelled as PIEMD's with magnitude scaled values (see previous section), with all cluster-scale potentials being modelled using PIEMD's. The analysis was started with the simplest model produced by fitting a model containing one cluster-scale halo, centred on the BCG, to the data – Model A. This model reproduced the positions of the observed multiple images to a rms precision of $\sigma_i = 0.52''$, with σ_i defined as (Richard et al., 2010):

$$\sigma_i = \sqrt{\sum_{j,k} (xobs_{j,k} - xpred_{j,k})^2 + (yobs_{j,k} - ypred_{j,k})^2} \tag{4.4}$$

With $xobs_{j,k}$ being the observed position and $xpred_{j,k}$ being the predicted position of an image j from source k in the x direction against the observed values (and the same for y), producing a root mean square deviation in the image plane of the lensing system.

The ellipticity of the cluster-scale halo was extreme ($\epsilon > 0.95$), indicating that a more complex model might be required. The cluster core radius was also unusually small at $r_{\rm core} = 35^{+3}_{-2}$ kpc, with the typical radius from Richard et al. (2010) being around $r_{\rm core} \sim 50 - 100$ kpc.

[1] http://www.oamp.fr/cosmology/lenstool

As the mismatch may have been due to the halo not being fixed on the BCG, the next model had a single cluster-scale halo fixed on the X-ray centroid – Model B. This choice of centre was motivated by the extension of the K-band luminosity density map to the South-West of the BCG (see Fig. 4.4). However in the absence of an obvious second peak in the luminosity map, the X-ray centroid was adopted as the centre of the halo. Model B had a poorer fit to the data than Model A (Table 4.3), which was not completely unexpected because of the location and positioning of the lensing system A1. Orban de Xivry & Marshall (2009) reproduced the same configuration with the halo being within the area described by the images.

As a final single cluster-scale halo model, the next model allowed the halo to move within a box of full-width 1arcmin centred on the BCG – Model C. This box included the X-ray centroid, which lay 12arcsec from the BCG. In terms of fit, Model C, did not improve dramatically on previous models, however the best-fit position of cluster-scale halo deviated from the BCG at $\sim 3\sigma$ significance in a direction roughly coincident with that of the X-ray centroid with respect to the BCG. It also had an extremely high ellipticity in the direction coincident with BCG and X-ray centroids. Whilst this was an unfeasibly high ellipticity it could possibly indicate the model trying to spread mass over both areas, indicating the need for another DM halo.

Models containing two cluster-scale halos

The failure of the single halo models to reproduce the lensed image positions accurately suggested that bi-modal models needed to be used. A PIEMD was still used to model the BCG and cluster member galaxies and, based on the previous model cluster-scale positions, one cluster-scale halo was placed on the BCG and one on the brightest galaxy within the group of galaxies that, in projection, inhabited the cluster core. However, as discussed in the context of Model B, there was no obvious group of galaxies and thus no obvious galaxy on which to centre the second cluster-scale halo. The next suitable candidate was the X-ray centroid – Model D. This model fitted the data very well, with

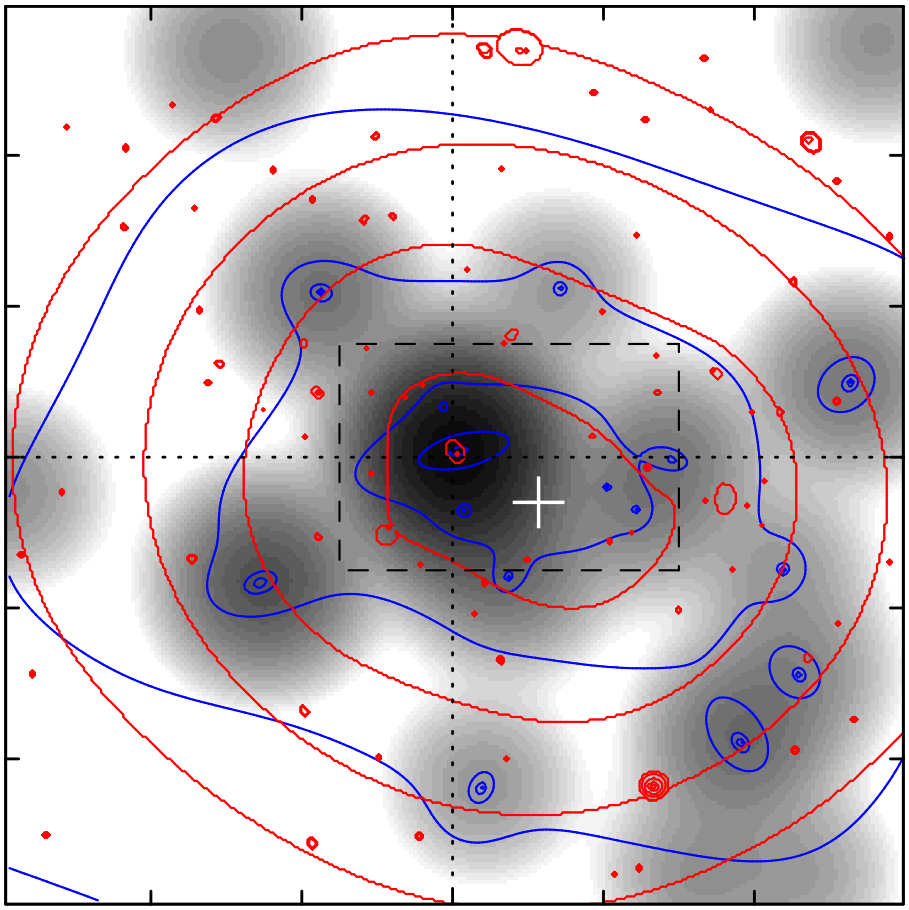


Figure 4.10: The central 2×2arcmin of A 3084, centred on the BCG, and shows the rest-frame K-band luminosity density of cluster members (as defined in §4.2.2; grey-scale), X-ray flux from *Chandra* (§4.2.4; red contours), and the projected mass density map (§4.3.1; blue contours). The large white plus marks the position of the X-ray centroid. The mass and X-ray contours are centred on the X-ray centroid, reflecting the large-scale structure of the gravitational potential. In contrast, the BCG is off-set from the centre of the potential. All contours and the grey-scale were plotted in the log. North is up and East is left.

$\sigma_i = 0.24''$, compared with an error circle of $0.2''$. The extra complexity of this model over Models A, B, and C was supported by the data, with the Bayesian evidence of Model D exceeding that of models discussed above by large factors.

To determine whether the assumption of cluster-scale halo positions was constraining the fit by having fixed positions, they were allowed to move by $\pm 5.5''$ in x, y around the BCG and X-ray centroid – Model E. This box-size was chosen as being the largest size possible without the boxes overlapping and so simplifying analysis of the chains generated by the LENSTOOL MCMC sampler. The extra flexibility of Model E fitted the data slightly better ($\sigma_i = 0.21$) than Model D ($\sigma_i = 0.24$), but the Bayesian evidence of the former was ~ 10 lower than that of the latter, indicating the additional model flexibility was not justified by the data. From this Model D was (from the five discussed thus far) the preferred model, because it maximized the probability of the data, and could reproduce the observed image positions to a rms precision comparable with the adopted error circle.

Three cluster-scale halos and model tests

So far the best description of the data was offered by Model D but to properly explore possible cluster descriptions, a model that contained three cluster-scale halos examining the residuals was used. This three-halo model – Model F – was identical to Model D, except for the addition of a third halo that was free to move within a box of full-width 1arcmin, centred on the BCG. The aim was to probe the possible location of additional substructure in the cluster core.

Model F fitted the data just as well as Model D, and the Bayesian evidence was consistent within the numerical precision with the evidence for Model D. The location of DM3 was not constrained, but the posterior probability density of the x, y coordinates of DM3 suggested that any additional substructures in the core of A 3084 might lie to the South-West of the BCG in the vicinity of the X-ray centroid (Fig. 4.11).

The multiple image A1.4 that was included as part of A1 was also investigated for model sensitivity. A1.4 was identified after subtracting light from the BCG from the

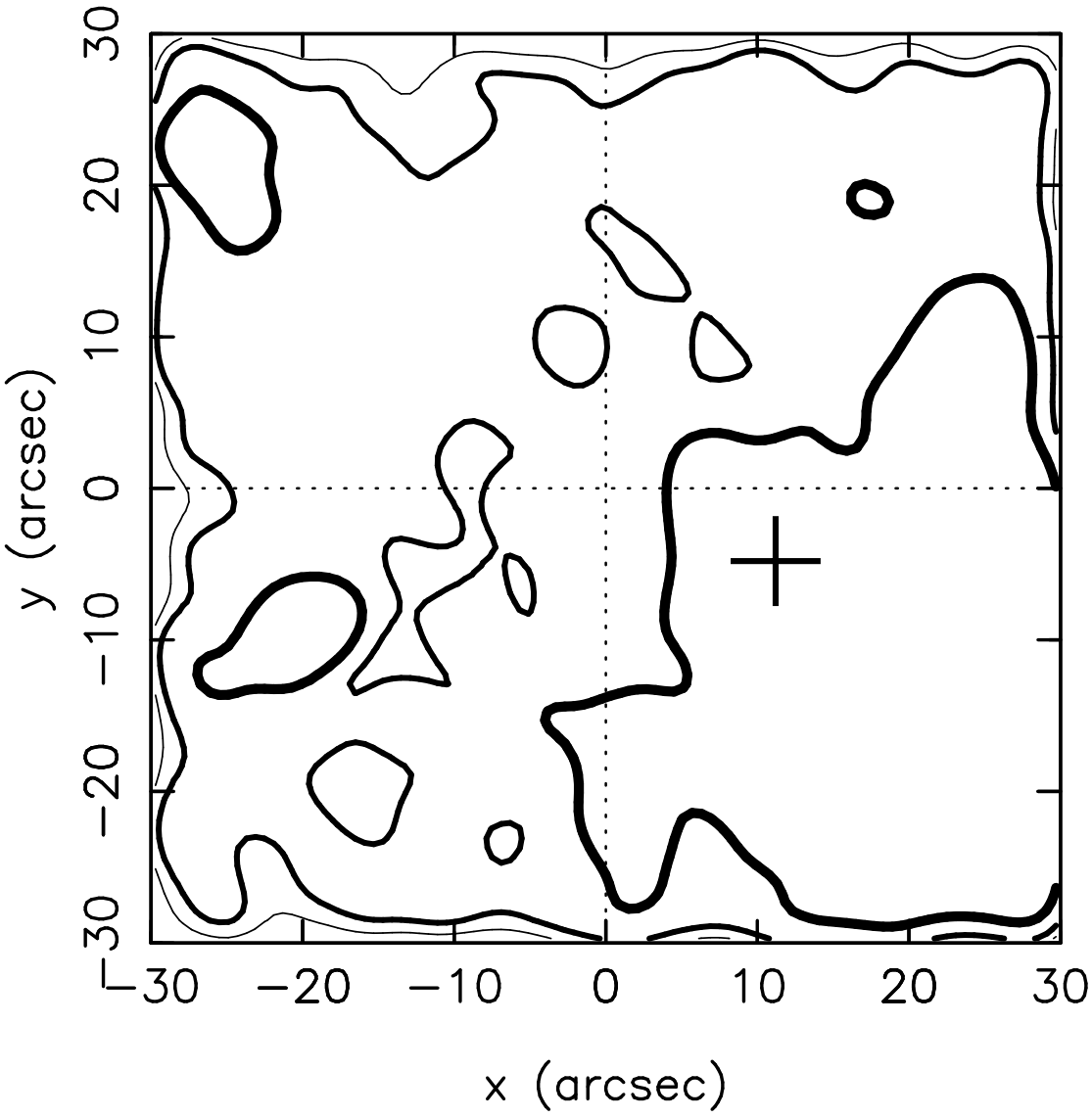


Figure 4.11: Posterior probability density of the 3-halo model discussed in §4.3.1. All model parameters were marginalised over except the x, y co-ordinates of the third cluster-scale halo. Thick, medium, and thin curves showed 68%, 95%, and 99% confidence levels respectively. The figure was centred on the BCG, and the black "plus" shows the location of the X-ray centroid. Despite the very low probability of the data, it revealed that the 3rd halo tended to lay to the South-West, following the K band luminosity map (Fig. 4.4) and the X-ray contours (Fig. 4.8).

HST/ACS frame (§4.2.1). It was therefore possible that the feature identified as A1.4 was a residual from the subtraction of the BCG. Instead of returning to details of the BCG subtraction, the issue was addressed by re-running all of the models listed in Table 4.3 excluding A1.4 from the constraints. These re-run models were fully consistent with those listed in Table 4.3. In particular, Model D reproduced the observed image positions to an rms accuracy of $\sigma_i = 0.2$, and maximized the probability of the data relative to the other four models.

Another source of model sensitivity was the masses assigned to each of the bright galaxies that were immediately adjacent to A1.5, and either side of the A2 triple-image system (Fig. 4.2). All the models listed in Table 4.3 were run again, but this time the velocity dispersion of the three galaxies in question were allowed to be free parameters (i.e. not scaled on their luminosity as described at the beginning of §4.3.1). The preference for, and properties of, Model D were again unchanged within the numerical precision.

In conclusion, Model D was the simplest model of those that were capable of reproducing the observational constraints ($\sigma_i \sim 0.2$) and maximized the probability of the data (the Bayesian evidence) compared to the other models. Model D was selected as the cluster model and was used to measure the mass and structure of the cluster core in §4.3.2.

4.3.2 The Mass and Structure of the Cluster Core

Model D was used to measure the mass of A 3084 within a projected radius of $R < 250$ kpc, and the fraction of that mass that resided in substructures, $f_{\rm sub}$. Richard et al. (2010) centred the aperture on the BCG in their study of 20 strong lensing cluster cores. This was a straightforward choice because the most massive cluster-scale halo in all of their models was either fixed on the BCG or the best-fit position was within a few arcseconds of the BCG. The situation was less clear cut in A 3084 because the cluster-scale halo centred on the X-ray centroid had a velocity dispersion comparable with that of the halo centred on the BCG. The physical interpretation of Model D is discussed in §4.4, however

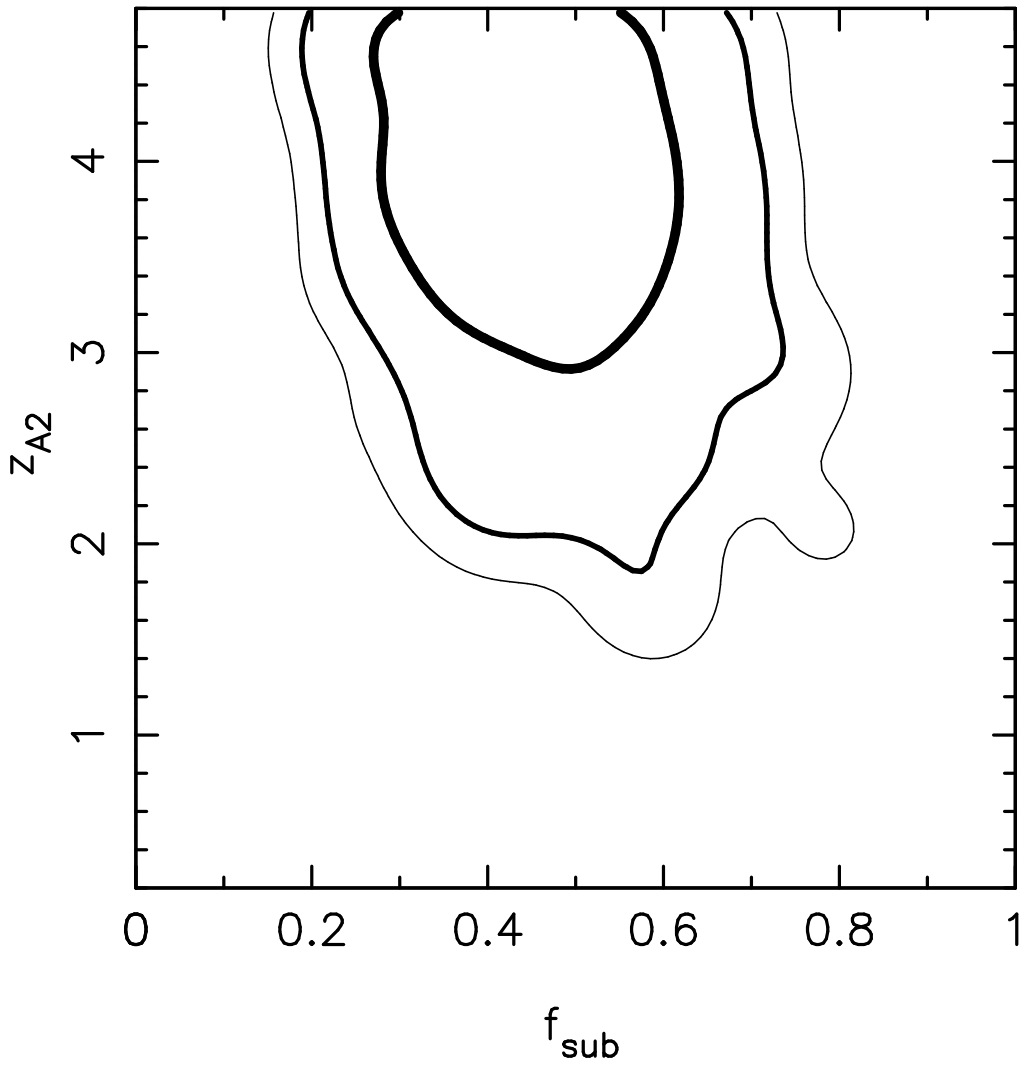


Figure 4.12: Posterior probability density in the $f_{\rm sub} - z_{\rm A2}$ plane for Model D, marginalizing over all other free parameters in the model. The full range of the prior on $z_{\rm A2}$ was plotted on the y-axis – $z_{\rm A2} < 4.8$ because it was detected through the F606W filter with ACS. The degeneracy between $f_{\rm sub}$ and $z_{\rm A2}$ was weak. Thick, medium, and thin curves show 68%, 95%, and 99% confidence levels respectively.

the obvious question raised by the model, that had practical implications for basic global measurements of mass and substructure fraction, is *"where was the centre of the cluster mass distribution?"*.

By adopting the BCG as the centre of the mass measurement aperture the following was obtained: $M_{SL}(R < 250 \text{ kpc}) = (9.3 \pm 0.5) \times 10^{13}$ $\text{M}_{\odot}$ and $f_{\text{sub}} = 0.47 \pm 0.17$. Shifting the centre of the aperture to the X-ray centroid changed the result to: $M_{SL}(R < 250 \text{ kpc}) = (8.3 \pm 0.5) \times 10^{13}$ $\text{M}_{\odot}$ and $f_{\text{sub}} = 0.73 \pm 0.13$. In both cases f_{sub} was mildly degenerate with the unknown redshift of the A2 multiple-image system (Fig. 4.12), but was not strong enough to alter any conclusions. The substructure fractions appeared to show a larger substructure fraction when associated with the larger dark matter halo (centred on the X-ray centroid) because when taking mass measurements it was done within a fixed aperture size of radius, $R = 250$ kpc. As a result the BCG was very compact (see blue contours on Fig. 4.10) and so more of its mass was contained within this aperture, while the other cluster-scale component had much more of its mass contained outside the aperture radius. The total mass measured was relatively insensitive to the choice of aperture centring because of the large aperture used for the mass measurement when compared to the small offset (42 kpc) between X-ray centroid and BCG (not much mass crossed the aperture boundary by the move of centres). In contrast, f_{sub} probed the distribution of mass within the aperture, and was therefore more sensitive to choice of centre. f_{sub} was defined as:

$$f_{\text{sub}} \equiv M_{\text{sub}}/M_{\text{total}} \tag{4.5}$$

Where M_{sub} was the mass residing in sub-halos (defined as any halo not centred on the centre of the mass-measurement aperture), and $M_{\text{total}} = M_{2D}^{SL}$ included all halos within the lens model. In the case of centring on the BCG, M_{sub} therefore included DM2 and all galaxies other than the BCG; in the case of centring on the X-ray centroid, M_{sub} included all halos except DM2.

The main reason to measure f_{sub} was to characterise the overall structure of the cluster core, and to compare it with other strong-lensing clusters (specifically to combine with

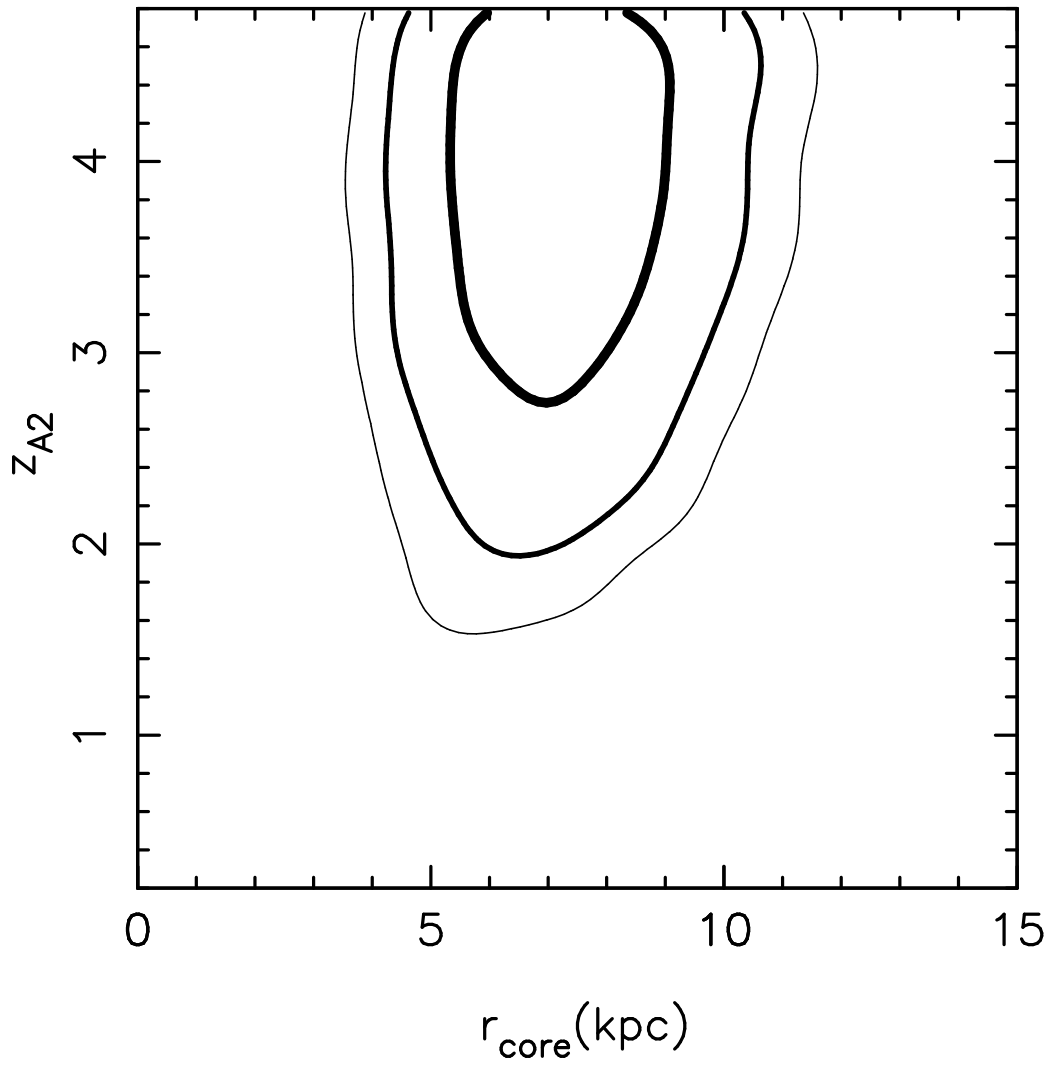


Figure 4.13: Joint confidence intervals on the core radius of DM1 and $z_{\rm A2}$ from Model D (Table 4.3). Thick, medium, and thin curves show 68%, 95%, and 99% confidence levels respectively.

values given in Smith et al., 2005 & Richard et al., 2010). Regardless of the ambiguity over aperture centre, it was clear from the relatively large values of $f_{\rm sub}$ obtained, and the shallow gas density profile (§4.2.4) that the overall structure of the mass and gas distribution in the core of A 3084 was typical of a strong-lensing cluster that required two or more cluster-scale halos to fit the strong-lensing constraints (Richard et al., 2010).

The most stringent constraints on the mass distribution in the cluster core came from the hyperbolic umbilic catastrophe, A1, four images of which laid within 5arcsec (17.5 kpc) of the centre of the BCG, and the cluster-scale halo centred on it, DM1. The hyperbolic umbilic catastrophe was therefore chiefly responsible for the tight constraints obtained on

the core radius of the DM halo within which the BCG was embedded, $r_{\rm core} = 6.9^{+1.5}_{-0.8}$ kpc (see Model D in Table 4.3), as a certain surface density was required to reproduce them and so dictated the total mass that could be within the rough circle that A1 outlines on the sky. The core radius of DM1 was not degenerate with any other parameter in the models, for example, the joint confidence intervals on $r_{\rm core}$ and $z_{\rm A2}$ in Fig. 4.13.

A typical BCG-centred cluster-scale halo in the core of a strong-lensing cluster has a core radius of $r_{core} \sim 20 - 100$ kpc (Richard et al., 2010). The core of A 3084 was therefore unusual in that DM1 had a very small core radius, and thus the BCG resided in a very compact DM halo. Moreover, among previously studied multi-modal cluster cores at similar redshifts, it was unusual to have such tight constraints on the radial profile of DM1.

4.4 Summary and Discussion

4.4.1 Main Results

In this chapter the distribution of DM, X-ray emitting gas, and stellar mass in the core of Abell 3084 using data from *HST*, VLT, Gemini-S, *Chandra*, and the Blanco telescope was studied. The *HST*/ACS snapshot observation revealed a rare strong-lensing image configuration (a hyperbolic umbilic catastrophe) surrounding the BCG in this cluster. The spectroscopically strong-lensing interpretation was confirmed and the lensed galaxy placed at $z = 0.764$. This system provided very tight constraints on the distribution of matter on $\lesssim 20$ kpc scales around the BCG. A wide range of mass models were fitted to the strong-lensing constraints to explore the detailed structure of the cluster core. These models were motivated by the distribution of hot X-ray emitting gas and stellar mass in the cluster core.

The centroid of the X-ray emission was offset 42 kpc to the South-West of the BCG, suggesting that the core might be disturbed. Despite the offset between X-ray centroid

and BCG, the overall X-ray morphology was quite "relaxed", being elliptical in projection and showed no obvious signs of shock or cold fronts. However, the ability to detect cold fronts was severely limited by the shallowness of the *Chandra* observations ($\sim$ 20 ksec).

Consideration of the stellar content of the core showed that the K-band luminosity density of the cluster, and thus its stellar mass density, peaked strongly on the BCG and was extended to the South-West, i.e. in the same direction as the axis connecting the BCG and X-ray centroid. However, there was no sub-peak in the luminosity density map – i.e. no obvious group of galaxies within the cluster core that might be interpreted as having penetrated the cluster core and thus influenced the recent thermodynamic history of the cluster gas.

The simplest mass model that could reproduce the current strong-lensing constraints to a precision comparable with the observational uncertainties contained two cluster-scale halos, Model D. In common with previous studies of bi-modal clusters (Richard et al., 2010), one of the cluster halos was centred on the BCG. However, this halo was very compact, with a core radius of just 7 kpc, i.e. an order of magnitude smaller than BCG-centric halos in other clusters. The second cluster halo in this model resided on the X-ray centroid and had a core radius of $r_{\rm core} \sim 40-70$ kpc, i.e. typical of cluster halos in strong-lensing cluster cores. Consequently, the "mass morphology" of the cluster core was well matched to the X-ray morphology (Figure 4.10), with the BCG embedded in a compact halo offset from the main centre of mass and X-ray emission.

To investigate the dynamical structure of the cluster further a Dressler Shectman plot was produced. Each cluster galaxy member was plotted as a circle, the radius of which scaled with the deviation of the local kinematics (based on the six nearest neighbour galaxies) from the global kinematics following Dressler & Shectman (1988, Figure 4.7) and Pinkney et al. (1996). The strongest evidence for departures from the global cluster kinematics was found to the South-East of the BCG. Galaxies with the largest deviations from global kinematics dominated the lower of the two redshift peaks, with the dynamically "colder" population presenting just one peak in the redshift distribution. Detailed

interpretation was complicated by the incompleteness of the spectroscopic observations within 1arcmin of the BCG. Nevertheless it can be concluded from the current data that the cluster was dynamically active along the line of sight, and that the BCG lay closer in redshift to the more prominent, higher redshift of two peaks (offset by ~ 300 km s^{-1}) than to the lower redshift peak (offset by ~ 1000 km s^{-1}).

4.4.2 Physical Interpretation

The physical interpretation of the unique properties of the core of A 3084 concentrated on the spatial and dynamical offset of the BCG from the underlying mass distribution of the cluster core. Three possibilities were considered:

(i) That the "BCG" was actually not physically associated with the cluster, and was seen in projection along the line of sight through the cluster.

(ii) That the "BCG" had fallen into the cluster pre-formed, and was seen in the data "en-route" to the bottom of the cluster potential well.

(iii) That the "BCG" was the *bona fide* brightest galaxy in A 3084, and it had been displaced from the bottom of the potential well by a cluster-cluster merger.

Section 4.2.3 showed that the velocity offset between the BCG and the mean cluster redshift was $\Delta v \simeq 450$ km s^{-1}. But when taking into account errors of ± 1320 km s^{-1}, this value did not favour interpreting the BCG as not physically associated with the cluster, regardless of the ambiguity in interpreting this as a velocity and/or physical separation. When included in the Dressler Shectman plot (Fig. 4.7), the small radius of its circle also pointed to it being part of the cluster as it did not deviate greatly from the global kinematics (see §3.3, equations 3.3 & 3.4). Another pointer to it being associated with the cluster and with the biggest of the redshift histogram peaks (Fig. 4.6) was that if the cluster disturbance was down the line of sight and at the maximum separation of the BCG from the largest cluster halo (the X-ray Centroid) and the velocity offset the

BCG had from the redshift histogram peaks was its true velocity, the BCG was only gravitationally held to the largest of the peaks (the velocity was too high for the BCG to be held in a stable orbit). Interpreting the BCG as falling into the cluster pre-formed also appeared implausible, because the galaxies that laid within $\sim 30''$ of the X-ray centroid (i.e. the bottom of the cluster potential well, as determined from the strong-lens model in §4.3), and satisfied the red sequence selection in §4.2.2 had absolute rest-frame K-band magnitudes of $M_K \simeq -25$ – i.e. typical of 2nd ranked cluster galaxies and not of 1st ranked cluster galaxies at $z \simeq 0.2$ (Smith et al., 2010, see their Fig. 4). The magnitude gap for the BCG to the rest of the cluster galaxies was $m_{12} = 1.4$. Values greater than $m_{12} > 1$ generally indicated homogeneous clusters with low substructure fractions and high concentrations. It was also an indicator of formation epoch and recent infall history of the cluster (Smith et al., 2010). The BCG had a K-magnitude of $m_K = 12.63$ and the next brightest galaxy existed in the the South-East of the cluster with $m_K = 14.03$, which was in the wrong area to coincide with either of the two large DM components SL restraints required. This also enforced the possibility that a second BCG candidate that stands out from the cluster galaxies was not visible.

The absence of strong morphological disturbance in the X-ray flux map of A 3084, indicated that the X-ray gas was not being strongly disturbed at the epoch of observation. The absence of obvious shocks or cold fronts was to be expected; due to the shallow nature of the X-ray observations, it was not possible to detect them even if they were there. This suggested that if the cluster-cluster merger interpretation was correct, then the event must have been viewed either early or late – i.e. before such disturbances developed, or if they had developed after they had dissipated. The optical morphology of A 3084 argued against interpreting this as an "early-stage" merger because there was no obvious concentration of galaxies that could be identified with an infalling galaxy system and associated DM halo. In fact the large luminosity gap would indicate the cluster was later in its formation history so that the BCG had time to become more dominant. Another fact is the intra-cluster gas was not displaced from the main cluster potential, but was

instead coincident with it; this is expected to take between $2-4$ Gyr depending on the ratio of merging halo masses and their impact parameter (Ascasibar & Markevitch, 2006). To estimate the BCG sloshing time required estimation of the BCG-BCG merging and the significance of BCG-X-ray peak offsets. The estimated time for galaxies in a cluster merger to nucleate can range from $2-7$ Gyr for $\delta < 300$ km s^{-1} (Kitzbichler & White, 2008), with the possible link between BCG offset from X-ray peaks being an area of current study, suggesting a typical non-zero offset between them and the possibility of offset being able to characterise merger and relaxation histories of clusters (Zitrin et al., 2012).

The most plausible interpretation therefore appeared to be that the core of A 3084 was in the process of settling down after a cluster-cluster merger, and that the BCG was currently displaced from the bottom of the potential well in the aftermath of the merger.

4.5 Summary

In summary the analysis revealed the following:

(i) A 3084 required two DM halos in order to reproduce the strong lensing image constraints. Single halo models were strongly rejected.

(ii) A 3084 was coincident on the smaller of the two DM halos and was separated from the X-ray centroid lying in the South-East by 43 kpc (the largest halo laid on this point).

(iii) The X-ray contours were smooth and undisturbed, but the redshift distributions (and Dressler Shectman test) of member galaxies was bimodal. This and the above information suggested the cluster could be in the "early" or "late" stages of a merger.

(iv) A 3084 exhibited rare lensing images in the form of a hyperbolic umbilic catastrophe. This configuration provided tight constraints to within ~ 18 kpc of the centre of the BCG.

(v) The DS plot showed disturbance to the South-East of the BCG, not coincident with the X-ray centroid (South-West) or BCG.

(vi) The most plausible interpretation of the data was that the BCG had been displaced via a merger from the cluster potential well and was situated in a compact DM halo.

CHAPTER 5

STRONG LENSING MASS VS X-RAY TEMPERATURES

Chapter 4 described the methods and procedures of lensing image constraints to reproduce a mass model for A3084. However, A 3084 was not the only strong lensing cluster that had multiple image lensing constraints. In chapter 3 multiple images in five clusters A 2813, A 2895, A 2537, A 368 and A 3084 were identified by reducing redshifts and constraining galaxy members, but not all had confirmed redshifts. This chapter uses these redshifts to produce more mass models that can be used on their own or added to existing samples for analysis.

In this chapter the new strong lensing derived mass models for clusters that had multiple images available are described. These were then added to the existing LoCuSS sample. The sample was then compared to T_X measurements for the same clusters andacr examined for any possible relation. Any scatter in the fitted relation was then studied by comparing the residuals (actual fit against relation derived) with other multi-wavelength derived measurements for each cluster. Finally the relation between the residuals is considered and refined by using other measurements.

For the models of A 368 and A 2537 the redshifts were provided by Johan Richard (in A 2537, the multiple image groupings as well). To increase signal to noise and make confident detections on the faint lensed image spectrum, Johan's expertise was used to help further constrain existing models. These would have had larger parametisation errors

without these redshifts. The resulting improved models are outlined here.

5.1 Strong Lensing Models

The models followed a similar iterative method as described in chapter 4 for A 3084. Methods set out in Richard et al. (2010) were also followed to allow the results to be added to the existing sample (described later). Richard et al. modelled the cluster galaxies (unless otherwise stated) as PIEMD's, with their parameters scaled with their K magnitudes, using the $L^{\star}$ star galaxy values in Table 5.2.

All mass distributions were constructed using LENSTOOL[1] (Jullo et al., 2007), with a combination of mass distributions to count for baryonic and dark matter components. The projected 2D total mass M_{SL} was measured within a radius of $R < 250$ kpc and the substructure mass was defined as the mass within the same radius, minus any dark matter or baryonic halos that were centred on the BCG. This was then converted into a fraction f_{sub}.

The additional strong lensing clusters and their models are outlined in the following sections. The best fitting models are shown in Table 5.2 and the mass measured within a 250 kpc radius and the substructure fraction measured within the same radius are listed in Table 5.3.

HST and X-ray contour overlays, along with the spectral features of multiple images, were taken from chapter 3 and repeated here for the ease of the reader.

5.1.1 Abell 2813

A 2813 (Fig. 5.1) had a very bright companion to its BCG to the South. Between the two brightest galaxies were two multiple image systems A1 & A2 that all laid in a West to East pattern suggesting they were caused by a saddle point between two gravitational potentials.

[1] http://www.oamp.fr/cosmology/lenstool

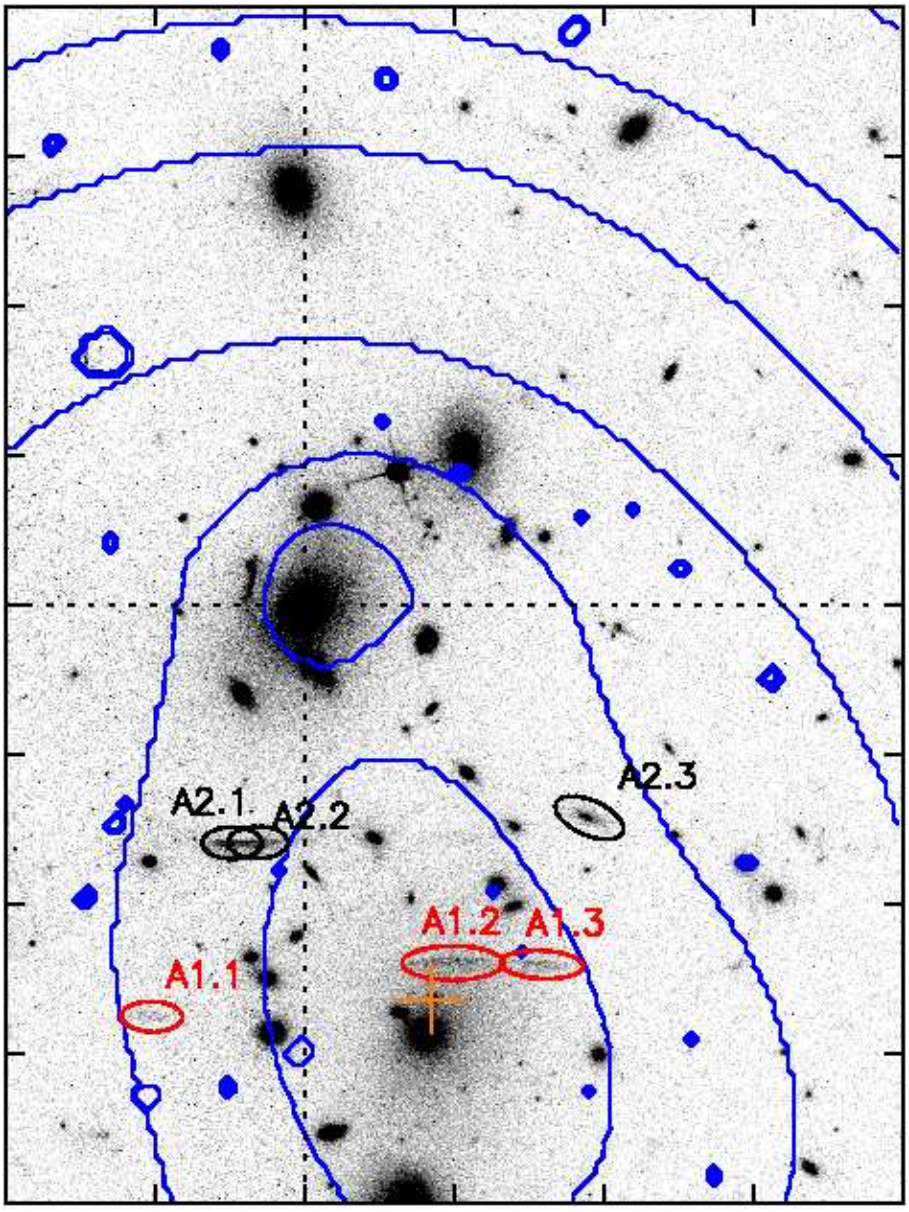


Figure 5.1: Abell 2813 with its BCG to the North and the second brightest cluster member situated to its South. The two multiple image systems between the two are labelled A1 and A2. The cross denotes the X-ray centroid position. North is up and East is to the left.

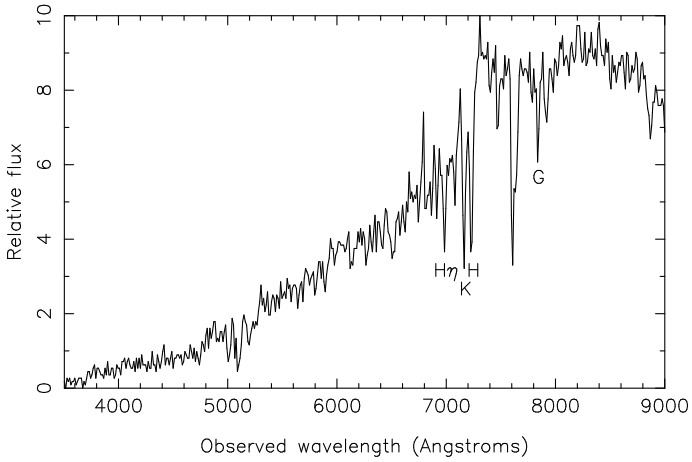


Figure 5.2: A2.3 spectrum shows the 4000 Å break and several other lines. These gave a redshift of $z = 0.8209 \pm 0.0006$.

Only one of the image systems (A2.3, see Fig. 5.2) had a confirmed redshift at $z = 0.8209 \pm 0.0006$, whilst the other was left as a free parameter in the model constraints. The positions and redshifts or selected range of these images are given in Table 5.1.

The positioning of both these image systems in parallel lines to one another suggested the presence of two large cluster sized halos. Placement of a DM halo on the next brightest cluster galaxy, as well as the BCG, provided a strong fit to the image constraints. A single halo model was tried, but fitted the constraints poorly, even if the position was not fixed for the halo. The mass was measured to be $M_{SL} = (2.44 \pm 0.11) \times 10^{14}$ $M_\odot$ and $f_{sub} = 0.71 \pm 0.05$ and the model results are listed in Table 5.2.

5.1.2 Abell 368

A 368 (Fig. 5.3) had one set of triple image systems labelled A1, with the redshift set at $z = 2.07$ (see Table 5.1) around its BCG. The image system suggested the presence of only a single DM halo. A single Pseudo-Isothermal Mass Distribution (PIEMD) Dark Matter halo was placed on the BCG and provided a good constraint to the image system.

The mass was measured to be $M_{SL} = (0.71 \pm 0.12) \times 10^{14}$ $M_\odot$ and $f_{sub} = 0.21 \pm 0.12$ and the model results are listed in Table 5.2.

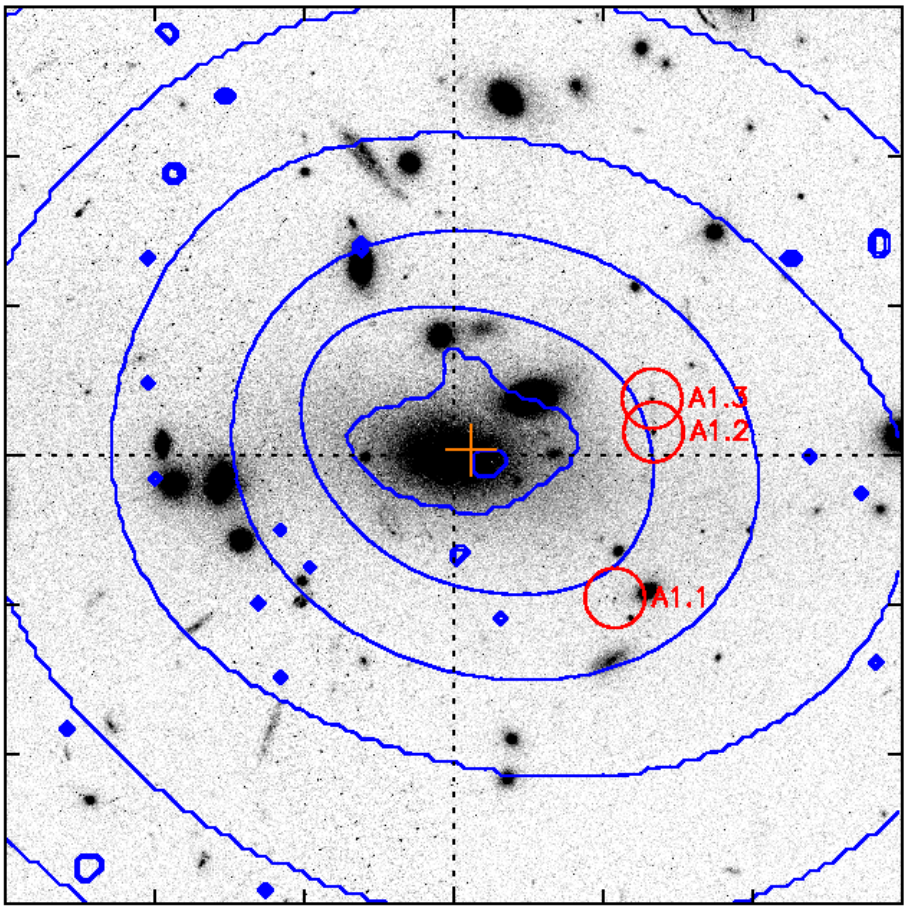


Figure 5.3: Abell 368 with its BCG and one set of multiple images labelled A1 to the West. The cross denotes the X-ray centroid position. North is up and East is to the left.

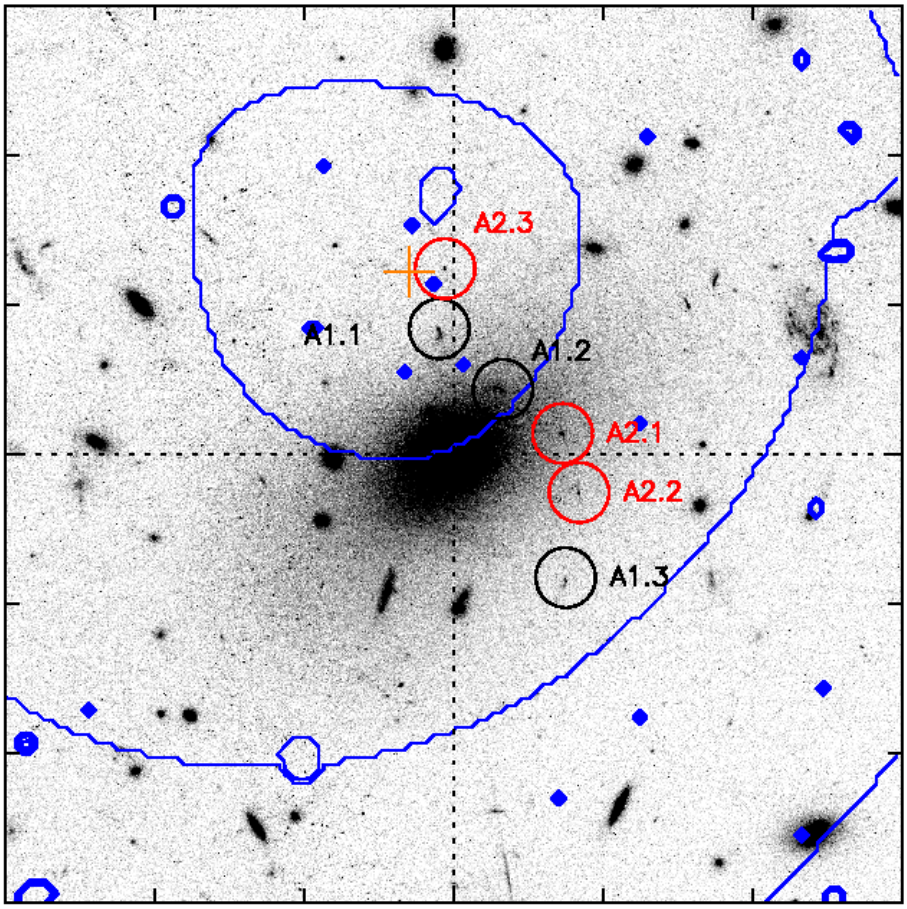


Figure 5.4: Abell 2895 with its BCG and the two multiple image systems A1 and A2 arrayed around it. The cross denotes the X-ray centroid position. North is up and East is to the left.

5.1.3 Abell 2895

A 2895 (Fig. 5.4) had two sets of triple image systems labelled A1 and A2 around its BCG. These were measured to be $z = 3.39$ for system A1 and $z = 3.72$ for system A1 (see Fig. 5.5) and the positions are listed in Table 5.1. This arrangement of increasing redshifts suggested one (PIEMD) halo centred on the BCG.

The mass was measured to be $M_{SL} = (0.92 \pm 0.31) \times 10^{14}$ $M_\odot$ and $f_{sub} = 0.08 \pm 0.04$ and the model results are listed in Table 5.2.

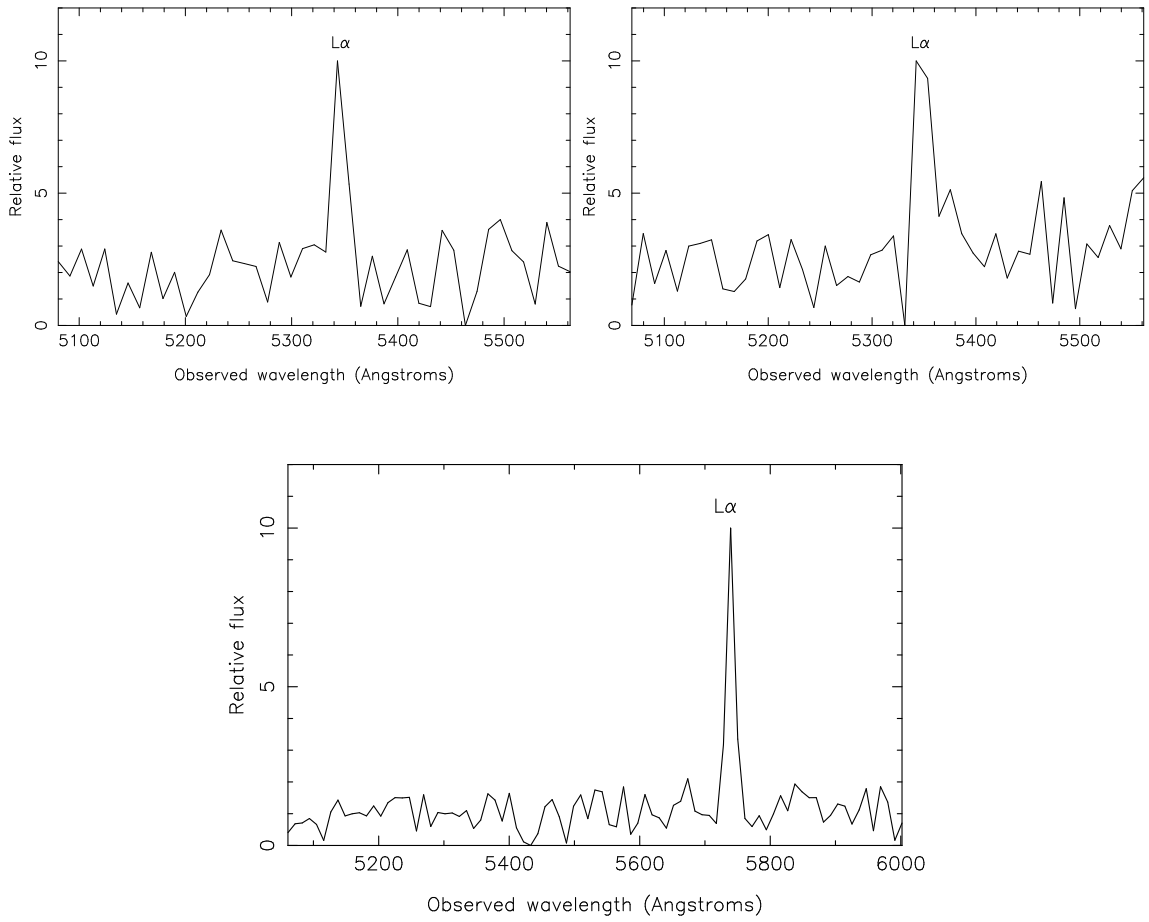


Figure 5.5: Abell 2895 multiple images. Top Left - The Lyman-Alpha emission line for the multiple image A1.1 (z=3.39). Top Right - The Lyman-Alpha emission line for the multiple image A1.2 (z=3.39). Bottom - The Lyman-Alpha emission line for the multiple image A2.1 (z=3.72).

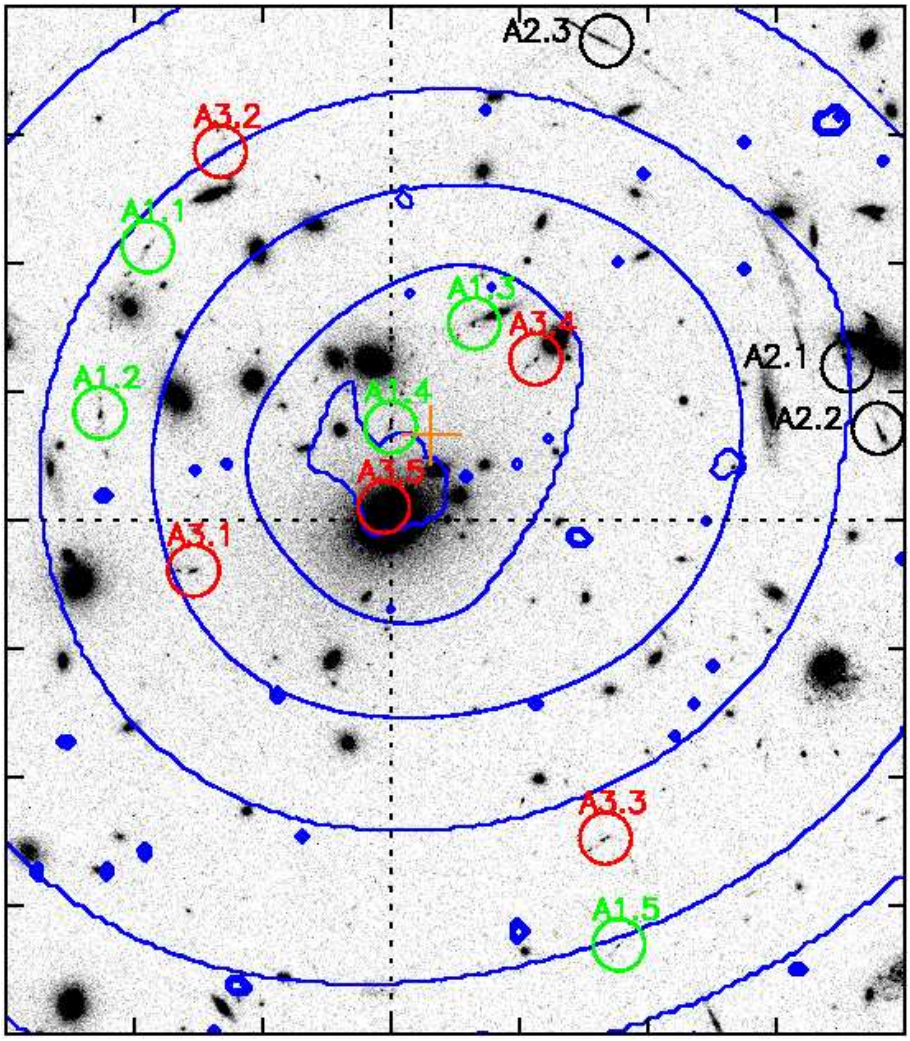


Figure 5.6: Abell 2537 with its BCG to the North and the second brightest cluster member situated to its South. The three multiple image systems are labelled A1, A2 and A3. The cross denotes the X-ray centroid position. North is up and East is to the left.

5.1.4 Abell 2537

Abell 2537 (Fig. 5.6) had three sets of multiple image systems; one quintet labelled A1 ($z = 3.607$), a triple system A2 ($z = 2.786$) and another quintet system A3 ($z = 1.970$). Their positions are listed in Table 5.1. These redshifts were reduced and measured by the author (confirming redshifts determined prior by Johan Richard).

Two PIEMD halos were used with constraints around the BCG and found to fit the image constraints reasonably well with an image rms of $\sigma_i = 1.09$.

The mass was measured to be $M_{SL} = (2.64 \pm 0.04) \times 10^{14}$ $M_\odot$ and $f_{sub} = 0.47 \pm 0.01$ and the model results are listed in Table 5.2.

5.1.5 Results

From the strong lensing analysis of the four above clusters, the model results for the DM and BCG halos are shown in Table 5.2. The table also shows any redshifts that were left as free parameters as well as the root mean square of the difference between the predicted and observed position of multiple images in the image plane.

5.2 Expanded Sample

From the results of this analysis four more cluster masses and substructure fractions could be combined with the table of cluster values from Richard et al. (2010) to increase the sample size. By also adding A 3084 from chapter 4 and the recently published values for the "Bullet Cluster" (1E 0657-55) from Paraficz et al. (2012) the total sample size was increased to 23, up from the original 17. The complete list is shown in Table 5.3. Only those clusters possessing all the values listed in Table 5.3 were selected from Richard et al. (2010), due to the requirement that these values can be compared with the mass measurements. In order that values were measured consistently, only clusters where the actual data was available, rather than quoted literature values, were included. From the

Table 5.1: Strong Lensing Image Constraints

ID	α	δ [J2000]	z
A 2813	.....	.....	.....
A1.1	10.857934	−20.624421	$[0.74 - 0.86]$
A1.2	10.851853	−20.623411	.
A1.3	10.850072	−20.623511	.
A2.1	10.856247	−20.621191	0.8209
A2.2	10.855731	−20.621191	.
A2.3	10.849196	−20.620758	.
A 368	.....	.....	.....
A1.1	39.3622666	−26.5106151	2.07
A1.2	39.3614628	−26.5075268	.
A1.3	39.3614929	−26.5069096	.
A 2895	.....	.....	.....
A1.1	19.546454	−26.967657	3.45
A1.2	19.545129	−26.968782	.
A1.3	19.543820	−26.972282	.
A2.1	19.543883	−26.969602	3.72
A2.2	19.543540	−26.970699	.
A2.3	19.546333	−26.966539	.
A 2537	.....	.....	.....
A1.1	347.09822	−2.1866524	3.607
A1.2	347.09926	−2.1902530	.
A1.3	347.09115	−2.1883190	.
A1.4	347.09296	−2.1905619	.
A1.5	347.08802	−2.2017542	.
A2.1	347.08307	−2.1892378	2.786
A2.2	347.08240	−2.1905849	.
A2.3	347.08829	−2.1822048	.
A3.1	347.09723	−2.1936616	1.970
A3.2	347.09665	−2.1845970	.
A3.3	347.08831	−2.1994475	.
A3.4	347.08982	−2.1890881	.
A3.5	347.09311	−2.1922910	.

Table 5.2: Strong Lensing Modelling Results, with 1σ[a] Errors

Halo	Δx (arcsec)	Δy (arcsec)	ϵ	θ[b] (deg)	$r_{\rm core}$ (kpc)	$r_{\rm cut}$ (kpc)	σ_0 (km/s)	z_{A2}	σ_i [c] (arcsec)
A 2813								$0.7580^{+0.0175}_{-0.0123}$	0.20
DM1	0.0	0.0	$0.15^{+0.16}_{-0.11}$	$+23.2^{+46.5}_{-77.5}$	40	1000	577^{+38}_{-47}		
DM2	6.9	−41.0	$0.36^{+0.08}_{-0.09}$	$+7.6^{+5.4}_{-11.4}$	40	1000	910^{+24}_{-28}		
BCG	0.4	−0.8	0.35	−35.6	0.3	98	231		
$L^\star$ galaxy					0.15	45	157		
A 368									0.03
DM1	0.0	0.0	$0.95^{+0.03}_{-0.06}$	$-8.2^{+2.8}_{-2.7}$	40^{+31}_{-24}	1000	718^{+129}_{-94}		
BCG	−1.5	−0.3	0.60	+15.8	0.3	98	232		
$L^\star$ galaxy					0.15	45	157		
A 2895									0.37
DM1	0.0	0.0	$0.78^{+0.01}_{-0.04}$	$+30.1^{+0.6}_{-0.6}$	96^{+64}_{-43}	1000	908^{+234}_{-182}		
BCG	0.0	0.0	0.37	+28.8	0.3	102	238		
$L^\star$ galaxy					0.15	45	157		
A 2537									1.09
DM1	$0.5^{+0.2}_{-0.2}$	$-0.8^{+0.3}_{-0.3}$	$0.40^{+0.04}_{-0.04}$	$+28.6^{+2.2}_{-2.0}$	38^{+3}_{-2}	1000	817^{+29}_{-27}		
DM2	$-1.1^{+0.3}_{-0.4}$	$12.4^{+0.4}_{-0.4}$	$0.60^{+0.07}_{-0.07}$	$+34.8^{+1.9}_{-1.9}$	161^{+14}_{-13}	1000	$904^{+49.8}_{-51}$		
BCG	0.0	0.0	0.32	+38.0	0.3	84	215		
$L^\star$ galaxy					0.15	45	157		

[a] All values were held as fixed constants unless quoted with error bars.
[b] Orientation of major axis, positive clockwise from x-axis.
[c] The rms deviation between the predicted and observed positions of the images in the image plane.

sample the mean mass was $M_{SL} \approx 1.97 \times 10^{14}$ $M_{\odot}$ and mean temperature $T_X \approx 7.48$ KeV.

5.3 Data Summary

The work so far can be summarised as follows:

(i) Using multiple image constraints four more clusters (A 368 , A 2537 , A 2813 and A 2895) were selected from the redshift sample in chapter 3.

(ii) A mass model was produced for each cluster and their best fit values presented in Table 5.2.

(iii) The cluster masses (and other multi-wavelength measurements) added to the existing sample from Richard et al. (2010), brought the total to 23 strong lensing clusters with complete multi-wavelength values to enable a full comparison to be made.

In section §5.4, the extended sample was compared to X-ray temperature and a fit attempted on a scaling relation. This was followed by a comparison of the residuals to other cluster measurements.

5.4 Analysis

The extended sample contained 23 clusters with a large selection of cluster observables, including X-ray temperatures. This enabled comparisons to be made between values and resulting correlations given a statistical value utilising the Pearson correlation, with a corresponding probability P_{value} for their null hypothesis also calculated (giving a probability the correlation was a sampling effect).

The Pearson correlation returns a value between -1 and 1, with 1 being a positive correlation. An increase in y has a corresponding increase in x and *vice versa*, with -1 being the reverse correlation. A value of zero corresponds to no correlation. The absolute

Table 5.3: Galaxy Clusters

Name	α, δ [J2000]	z	ϵ_{BCG}[b]	M_{SL} (10^{14} $M_\odot$)	T_X (KeV)	Δm_{12}[a]	f_{sub}	$\Delta M/M_{fit}$
Abell68	00:37:06.81 +09:09:24.00	0.255	0.26	2.16 ± 0.23	$8.89^{+3.41}_{-1.79}$	1.40	0.33 ± 0.04	-0.23 ± 0.08
Abell2813	00:43:25.13 -20:37:01.41	0.292	0.35	2.44 ± 0.11	$6.51^{+0.89}_{-0.71}$	0.16	0.71 ± 0.05	0.19 ± 0.05
Abell2895	01:18:11.05 -26:58:11.65	0.228	0.37	0.92 ± 0.31	$8.08^{+1.72}_{-1.18}$	0.47	0.08 ± 0.04	-0.64 ± 0.12
Abell368	02:37:27.80 -26:30:29.13	0.222	0.60	0.71 ± 0.12	$6.59^{+1.71}_{-1.09}$	1.26	0.21 ± 0.12	-0.66 ± 0.06
Abell383	02:48:03.38 -03:31:45.70	0.187	0.07	1.87 ± 0.26	$5.01^{+0.49}_{-0.41}$	1.90	0.02 ± 0.01	0.19 ± 0.17
Abell3084	03:04:04.01 -36:56:26.97	0.219	0.11	0.93 ± 0.05	$4.29^{+0.91}_{-0.59}$	1.40	0.47 ± 0.17	-0.38 ± 0.04
Abell521	04:54:06.86 -10:13:23.00	0.247	0.22	0.61 ± 0.33	$7.08^{+1.52}_{-1.18}$	0.05	0.13 ± 0.04	-0.73 ± 0.15
Abell611	08:00:56.80 +36:03:23.41	0.288	0.33	1.76 ± 0.33	$7.94^{+1.26}_{-1.04}$	2.16	0.10 ± 0.01	-0.29 ± 0.13
Abell773	09:17:53.37 +51:43:37.20	0.217	0.29	3.01 ± 0.58	$7.50^{+1.00}_{-0.90}$	0.13	0.78 ± 0.03	0.28 ± 0.25
ZwCl0949.6+5207	09:52:49.19 +51:53:05.26	0.214	0.17	1.74 ± 0.14	$5.08^{+0.42}_{-0.38}$	2.33	0.04 ± 0.02	0.09 ± 0.09
Abell1835	14:01:02.05 +02:52:42.30	0.253	0.13	2.83 ± 0.41	$9.82^{+0.48}_{-0.42}$	1.57	0.13 ± 0.01	-0.08 ± 0.13
Abell963	10:17:03.57 +39:02:49.20	0.206	0.19	1.74 ± 0.44	$6.73^{+0.77}_{-0.53}$	1.26	0.13 ± 0.07	-0.18 ± 0.21
Abell1201	11:12:54.50 +13:26:09.24	0.169	0.58	0.8 ± 0.33	$5.56^{+0.84}_{-0.56}$	2.54	0.02 ± 0.01	-0.54 ± 0.19
Abell1689	13:11:29.49 -01:20:27.18	0.183	0.24	4.53 ± 0.13	$8.86^{+0.84}_{-0.66}$	0.68	0.22 ± 0.03	0.63 ± 0.05
Abell2204	16:32:46.71 +05:34:30.90	0.152	0.22	2.29 ± 0.5	$9.64^{+0.56}_{-0.54}$	0.14	0.25 ± 0.10	-0.24 ± 0.17
RXJ1720.1+2638	17:20:10.06 +26:37:32.13	0.164	0.33	1.18 ± 0.59	$7.96^{+0.84}_{-0.96}$	1.60	0.10 ± 0.05	-0.53 ± 0.24
RXCJ2129.6+0005	21:29:39.96 +00:05:21.74	0.235	0.42	1.37 ± 0.37	$8.27^{+1.83}_{-1.47}$	1.26	0.15 ± 0.06	-0.47 ± 0.14
Abell2218	16:35:49.22 +66:12:44.80	0.171	0.46	3.00 ± 0.24	$7.17^{+0.63}_{-0.47}$	0.46	0.54 ± 0.01	0.33 ± 0.11
Abell2219	16:40:19.82 +46:42:41.50	0.226	0.29	2.33 ± 0.23	$11.52^{+0.88}_{-0.82}$	0.75	0.57 ± 0.04	-0.36 ± 0.06
Abell2390	21:53:36.84 +17:41:43.67	0.233	0.03	1.99 ± 0.07	$9.78^{+0.52}_{-0.58}$	1.53	0.03 ± 0.03	-0.35 ± 0.02
Abell2537	23:08:22.22 -02:11:32.06	0.297	0.32	2.64 ± 0.04	$6.65^{+0.95}_{-0.85}$	0.53	0.47 ± 0.01	0.26 ± 0.02
Abell2667	23:51:39.44 -26:05:02.21	0.226	0.40	2.41 ± 0.07	$5.66^{+0.84}_{-0.66}$	0.87	0.14 ± 0.03	0.36 ± 0.04
1E 0657-55	06:58:38.126 -55:57:25.87	0.296	0.26	1.96 ± 0.02	$12.01^{+1.25}_{-1.22}$	0.03	0.48 ± 0.01	-0.48 ± 0.01

[a] The error in Δm_{12} was dominated by uncertainties on the photometric calibration of 0.1 magnitudes (Smith et al., 2010).
[b] The ellipticity error was dominated by isophote fitting uncertainties of order 0.07.

Table 5.4: Strength of Pearson Correlation

Pearson Value	Strength
0.00 - 0.19	Very Weak
0.20 - 0.39	Weak
0.40 - 0.59	Moderate
0.60 - 0.79	Strong
0.80 - 1.00	Very Strong

value of any correlation has an associated correlation strength such as weak, strong etc., which is shown in Table 5.4 (Evans, 1996).

It was possible to produce a hypothesis test and return a P_{value} for each correlation, which indicated the probability that the null hypothesis was true and any correlation was as a result of the points sampled from the population and not intrinsic of the population itself. A value of $P_{value} \leqslant 0.05$ is usually used to discard the null hypothesis, strengthening the Pearson correlation statistic as not being due to a fortuitous arrangement of sample points, but as a real correlation to the population the sample is taken from. However, the Pearson correlation cannot provide any evidence of a causal link between the two correlated values.

The obvious comparison, that of lensing mass versus X-ray temperature, has been found to show a link (see, Richard et al., 2010). Though to thoroughly check all links other cluster measurements were also compared to the lensing mass in the following sections.

5.4.1 Mass Observables

Fig. 5.7 showed three of the cluster measurements compared with the strong lensing derived two dimensional mass within 250 kpc. Each of the three plots showed only a weak correlation (see Table 5.4) to lensing mass, but the null hypothesis probability values were slightly higher than 0.05 and therefore the results could be due to the sample and not the population, a selection of which formed the sample. For substructure fraction, luminosity gap and BCG ellipticity the weak correlation suggested there was no direct link between

these observables and lensing mass, except in the case of X-ray temperature. It was this latter measurement that was investigated in the subsequent sections of this chapter.

5.4.2 Mass vs Temperature

A correlation should exist between the strong lensing derived mass, M_{SL} (a 2 dimensional quantity, integrating the mass along the line of sight) for a cluster and its X-ray gas temperature, T_X. It was predicted by Kaiser (1986) for a virialised cluster for there to be a relation of order $M \propto T^{3/2}$. This was due to the gas temperature being affected by the cluster's potential well. As long as the cluster was relaxed and the gas was only being compressed by the cluster's gravity, then a hydrostatic equilibrium should exist and make both M_{SL} and T_X degenerate with one another. Cluster disturbance could be difficult to check if there were only a few different wavelength observations.

Following previous studies between these two variables, the relation between the two values should be in the form:

$$M_{SL} = \beta T_{X-Ray}^{\alpha} \tag{5.1}$$

Using the sample derived so far in this chapter, these two values were plotted against one another (see Fig. 5.8), resulting in a Pearson value of 0.34, showing a "weak" correlation to one another.

To apply a best fit line to these two variables, both sides of the equation were logged and a similar method for a χ^2 fit used by Mahdavi et al. (2008) performed. This fit was based on a straight line fit in log space and was based upon Press et al. (1992), which gave the χ^2 merit function to fit $y(x) = a + bx$ (with errors in y and x) as:

$$\chi^2(a,b) = \sum_{i=0}^{N-1} \frac{(y_i - a - bx_i))^2}{\sigma_{yi}^2 + b^2\sigma_{xi}^2} \tag{5.2}$$

The resulting χ^2 values vary as α is altered, with the lowest χ^2 value giving the best fit value and, by looking at $\chi^2_{lowest} + 1$, the 68% error values (Press et al., 1992). The best

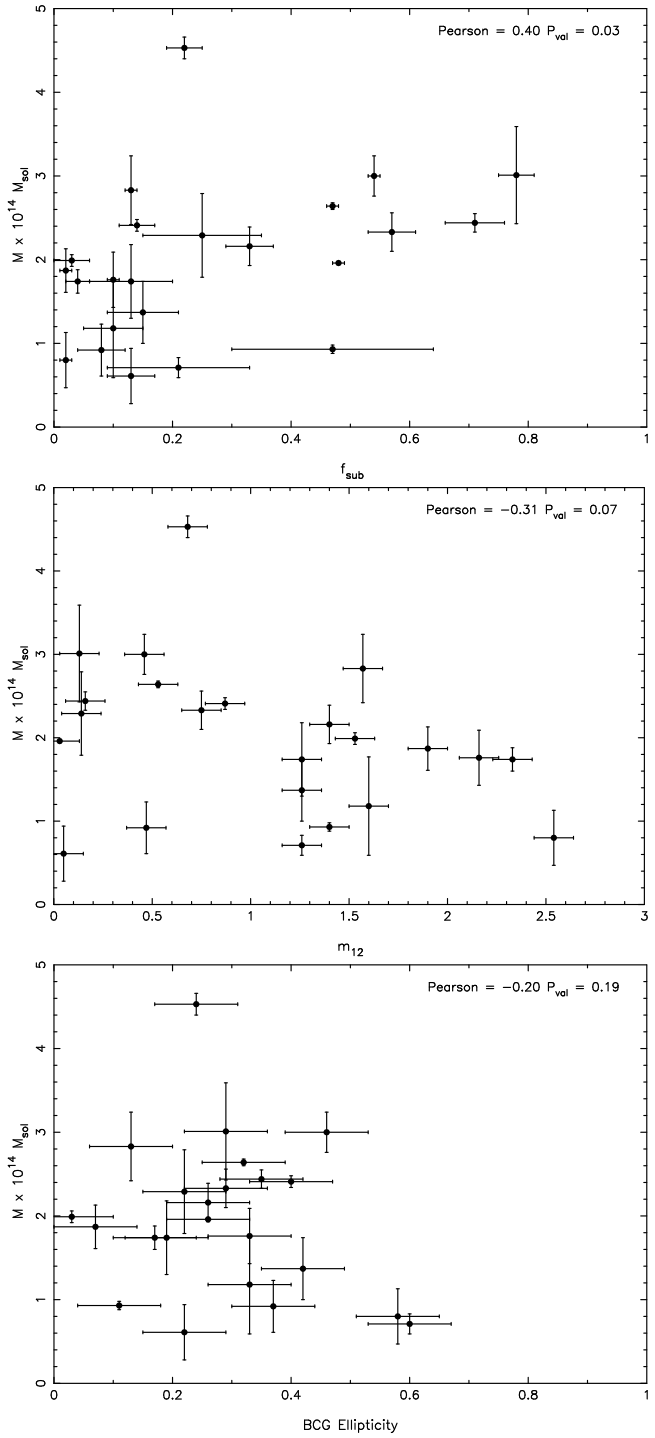
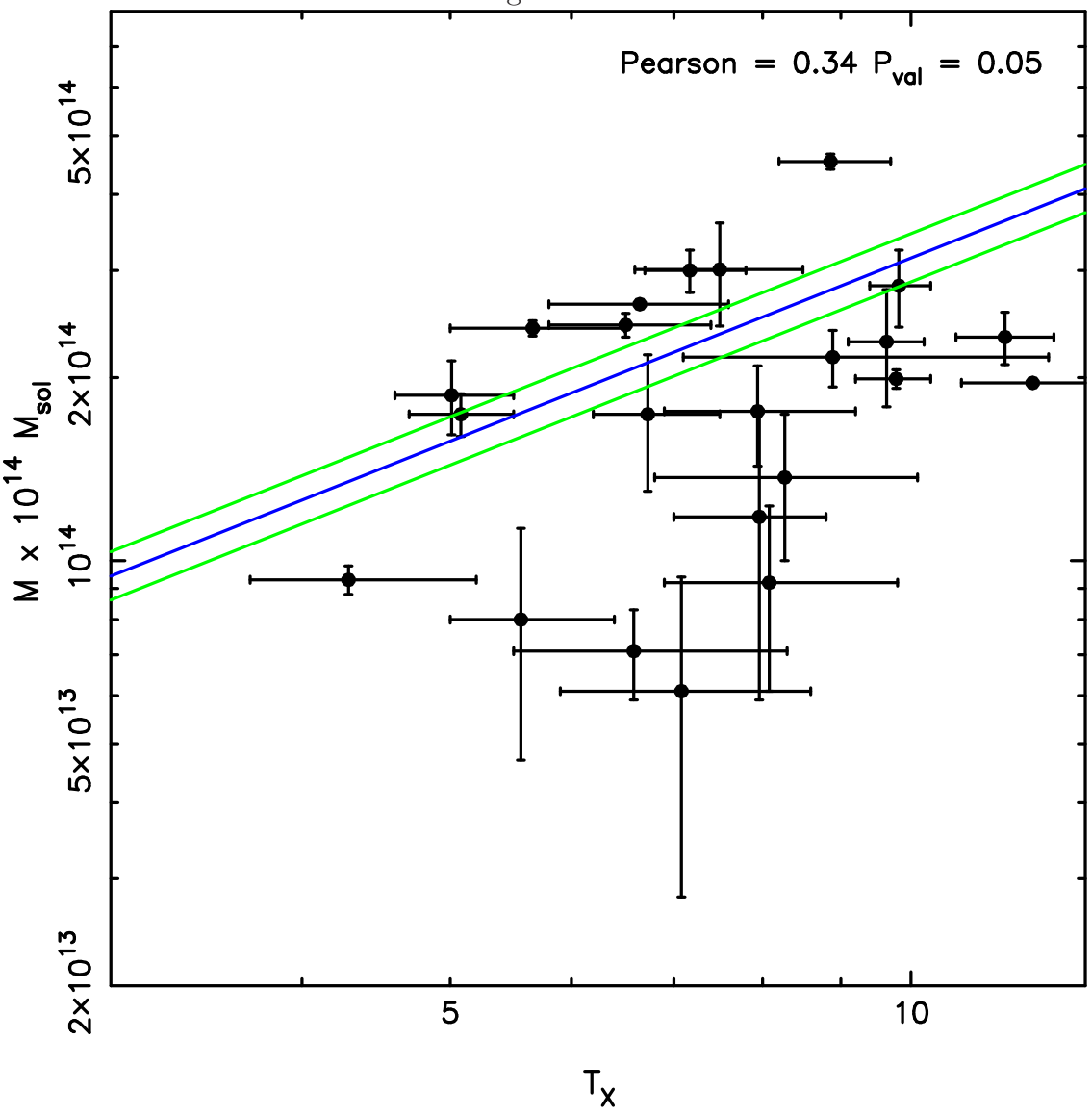


Figure 5.7: Top – Lensing mass plotted against substructure fraction, $f_{\rm sub}$. Middle – Mass plotted against luminosity gap, m_{12}. Bottom – Mass plotted against BCG ellipticity.

Figure 5.8



M_{SL} versus T_{X-Ray} in log space. Blue line shows the best fit line and the green lines the error lines after scaling the error bars.

fit value was plotted in log space in Fig. 5.8, the errors bars are from scaling the mass and temperature error bars (see following discussion). The lowest χ^2 value yielded a fit value of $\beta = 3.14^{+0.10}_{-0.09} \times 10^{13}$ (α was fixed to a value of 1, see §5.7). These errors were too low to be a realistic representation of the sample errors and, when also combined with a $\chi^2_{reduced} \approx 8$ (from the lowest value of $\chi^2_{total} = 160$), indicated large intrinsic scatter in the plot. This scatter was also reflected in the low Pearson value of 0.34. Press et al. (1992) suggested that by artificially inflating or deflating the sample errors a more realistic (68%) error line could be calculated. When this scaling was performed $\beta = 3.14^{+0.31}_{-0.27} \times 10^{13}$ (a jackknife error estimate returned $\sigma_\beta = 0.45 \times 10^{13}$) and can be seen in Fig. 5.8. Understanding the reasons for this scatter (its origins and possible causes) is considered in the following sections.

Other combinations of cluster values did not show any strong trends (very low Pearson values and high P_{values} were prevalent), suggesting either they had no correlation or other factors were more dominant.

5.5 Residuals

The intrinsic scatter in the mass versus temperature relation might be related to other factors that were not being compensated for. The next logical step was the ability to exclude outlying points (they might break the assumptions required for there to be a correlation) or to correct for deviation (mass or temperature was possibly being under- or over-estimated) from their true values, was the next logical step.

To do this the difference between the actual mass value and the value calculated using the fitted relation was calculated (ΔM). This was then reduced to a fraction of under- or over-mass estimation by dividing by the fitted line M_{fit}, to give the final value as $\Delta M/M_{fit} = (Mass - M_{fit})/M_{fit}$.

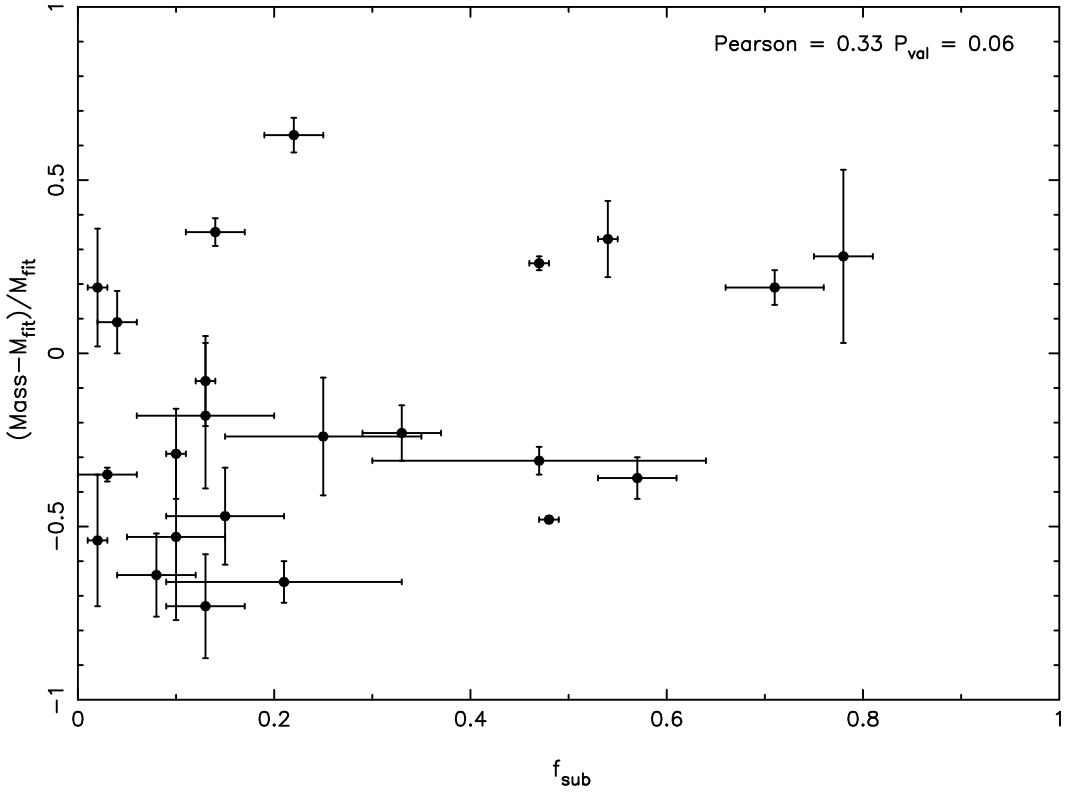


Figure 5.9: Mass difference from fitted mass as a fraction of fitted mass versus substructure fraction.

5.5.1 Substructure Fraction

The assumption of any possible fit would be that f_{sub} was an indicator of the cluster being disturbed and so the boost of temperature would create an artificial mass estimation that was higher than the true value, such that it would be under estimated. Fig. 5.9 plots $\Delta M/M_{fit}$ against substructure fraction f_{sub}, but showed no obvious relation between the two, with the Pearson correlation only yielding a "weak" value of 0.33. The assumption of any possible fit would appear to be too simplistic or non-existent in this case.

5.5.2 Luminosity Gap

Smith et al. (2010) showed a link between luminosity gap and the dominance of a cluster. If a cluster was relaxed and had time for the BCG to absorb infalling galaxies its luminosity should increase, creating a larger magnitude gap between the BCG and the next brightest cluster galaxy (m_{12}). The mass temperature relation worked on the assumption the cluster

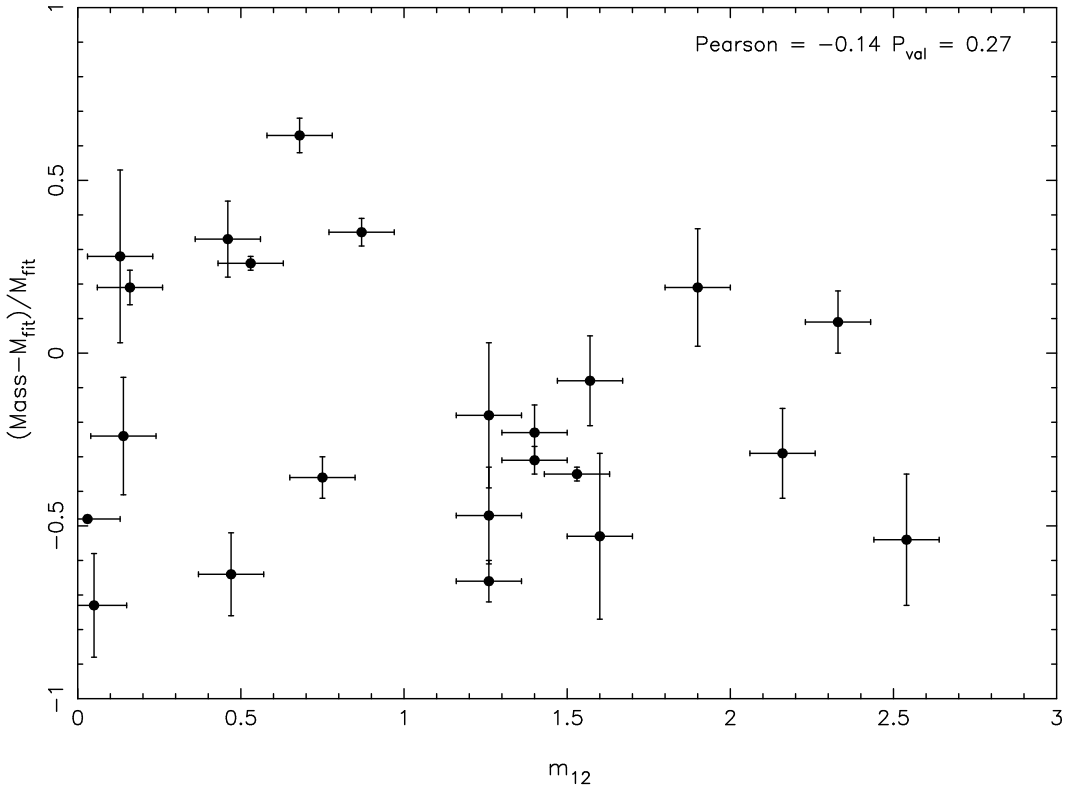


Figure 5.10: Mass difference from fitted mass as a fraction of fitted mass vs luminosity gap.

was relaxed and so any deviation from this would lead to an incorrect mass estimation. If luminosity gap was a possible indicator of cluster disturbance then plotting $\Delta m/m_{fit}$ against m_{12} in Fig. 5.10 might show a correlation. The resulting fit only yielded a negative correlation value of -0.21 ("weak" correlation) with a $P_{value} = 0.18$, which indicated that this correlation could be as a result of the sampling of the population, rather than a value from that population.

5.5.3 BCG Ellipticity

Observationally and via N-body simulations, evidence has been provided of DM halos not being axisymmetric (Frenk et al., 1988, Stadel et al., 2009, Vera-Ciro et al., 2011 and Salvador-Solé et al., 2012 all found evidence of asymmetry), but instead tri-axial, in a prolate, rather than oblate direction, and that elliptical galaxies, the BCG being an example, also possessed a tri-axial shape. This meant that if the long axis of the DM

halo was in the plane of the sky when observed, the same X-ray temperature would be observed but an underestimation of the mass would occur if tri-axiality was not taken into account. Conversely, looking at the DM halo along its long axis would read a higher mass for the same temperature. This is because, even though current DM profiles (such as PIEMD and NFW) may allow for ellipticity (only in the image plane) they are not truly tri-axial. As a result these DM profiles would not correctly approximate a tri-axial halo when viewed at various angles relative to the observer, leading to under and over estimation of masses when fitting observational constraints.

Determining tri-axiality in DM halos was difficult due to DM being dark and observable from one viewing direction, but other indicators on DM halo long axis orientation might exist for a cluster. One possibility hinted at in Marrone et al. (2009), and very recently further confirmed as an important effect to take into account in Marrone et al. (2012), was that of BCG ellipticity being a possible indicator of the direction of the long axis.s As a BCG tends to live at the centre of a cluster-scale potential, any prolateness in the potential would exert a torque on an elliptical BCG and cause it to align its own long axis with that of the DM potential if left to settle for a sufficient amount of time. By measuring the ellipticity of the BCG it should therefore be possible to determine if there was a relationship between BCG ellipticity and the under- or over-estimation of cluster mass.

Fig. 5.11 shows $\Delta M/M_{fit}$ plotted against BCG ellipticity. Taking the sample as a whole it only yielded a "very weak" (Pearson value of -0.18 for the whole sample) negative correlation. However, qualitatively there did appear to be a negatively correlated distribution of clusters, surrounded by many outliers that are known to be possess strong substructure (see Table 5.3). This correlation was empirically found when those clusters that were determined to have a small luminosity gap were removed. A small luminosity gap gave a strong indication that it was disturbed and not therefore in agreement with the mass temperature relation assumption of clusters being relaxed. This correlation was found due to many of these selected clusters being identified as outliers, resulting in a

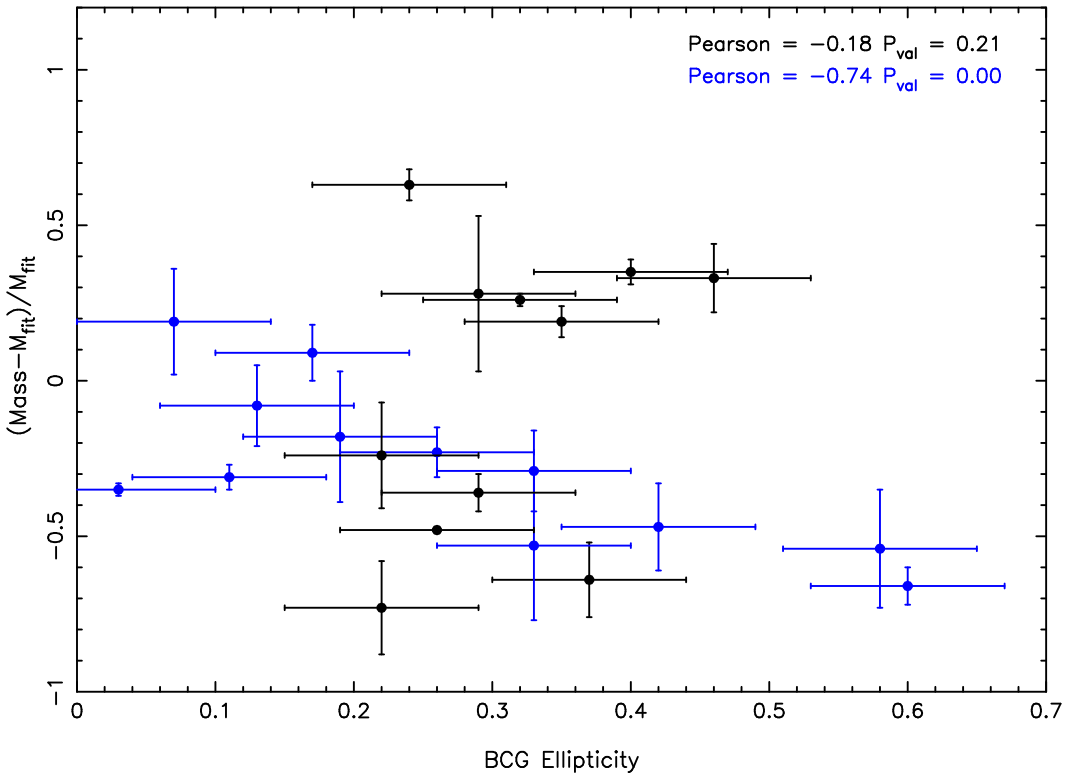


Figure 5.11: Mass difference from fitted mass as a fraction of fitted mass vs BCG ellipticity. Those clusters labelled in blue had a luminosity gap magnitude value greater than 1. The Pearson value of −0.18 was calculated for all points (blue and black).

"strong" negative correlation with a value of −0.71 and a $P_{value} = 0.00$, which strongly rejected the Null hypothesis of this being a sampling effect. In the following sections the method for outlier removal is described and explained. The correlation between mass over/under estimation and BCG ellipticity was found to persist with changing values for β in the strong lensing mass vs X-ray temperature relation. This meant the fitted line only changed the normalisation or y-axis of Fig. 5.11 and a different fitted value of β did not destroy the correlation found in this section.

5.6 Residual Discussion

BCG ellipticity tracing halo tri-axiality (see Fig. 5.11) showed promise as being a possible indicator and potentially correction factor for mass over- and under-estimations, but the ability to explain outliers to the relation were important in order to add any significance

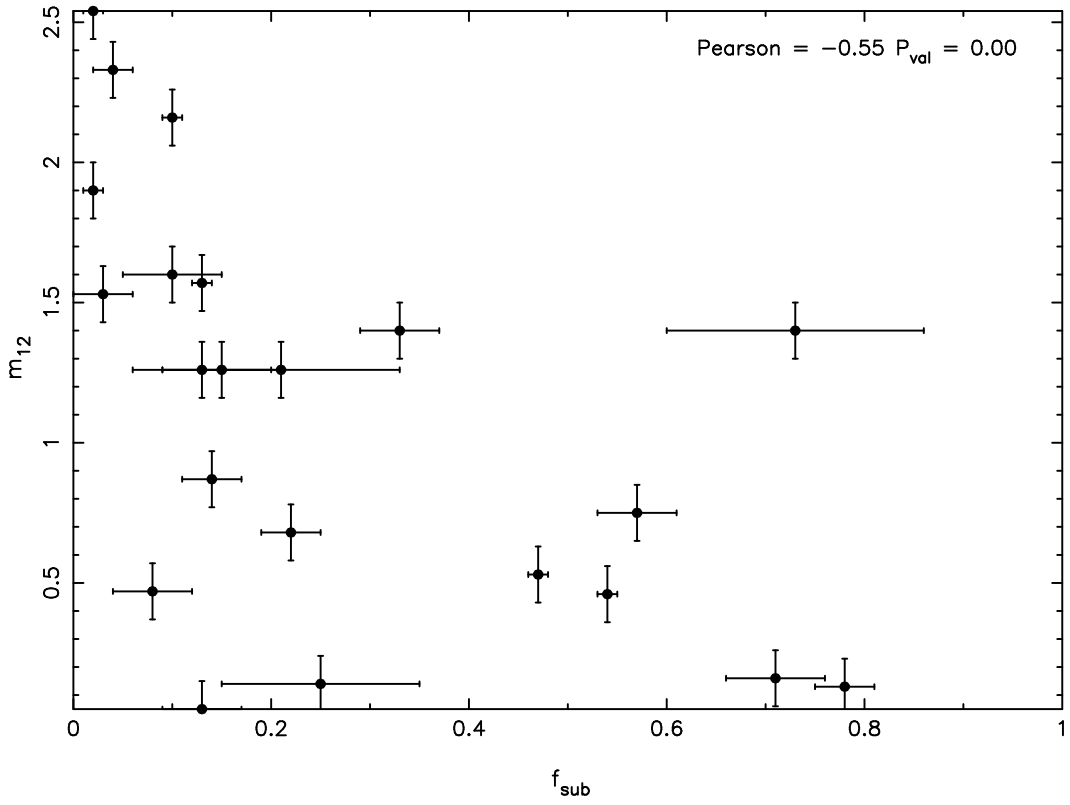


Figure 5.12: Luminosity gap versus substructure fraction.

to this.

As discussed earlier, the mass temperature relation only worked if the assumption of the cluster being relaxed was true. Smith et al. (2010) showed that dominant BCG's had larger luminosity gaps, but these only formed when clusters were settled; if the cluster was disturbed then the inter-mingled galaxies from the two merging systems would lessen this gap. Fig. 5.12 shows the inverse link between m_{12} and f_{sub} as the two values should be related. Indeed the Pearson correlation returned a "moderate" negative correlation (Pearson value was -0.55) between the two, and strongly rejected the null hypothesis ($P_{value} = 0.00$). A higher substructure fraction would indicate the cluster was dynamically disturbed and as a result a lower luminosity gap would be observed.

To remove disturbed clusters a luminosity cut of $m_{12} \geq 1$ was applied to the sample. Unlike fsub, this was not linked to mass measurement accuracy. This is justified from Smith et al. (2010), which showed that above this value the luminosity of first and second ranked galaxies started to become very large, suggesting an over bright BCG that had

had time to absorb luminous material. Applying a cut of $m_{12} \geq 1$ to the sample allowed a re-analysis of the correlation between mass over- or under-estimation against X-ray temperature, which resulted in a "strong" negative correlation (Fig. 5.11).

5.7 Chapter Discussion

Using the multiple image redshifts from chapter 3 enabled five strong lensing clusters to be modelled, with M_{SL} and f_{sub} measurements taken. Adding these to the existing sample (along with another from literature) enabled a sample increase from 17 to 23. Fitting this with X-ray temperature produced a fit shown in Fig. 5.8. The mass and temperature relation $M = \beta T_X^{\alpha}$, was fitted for β (α was fixed to 1), resulting in $\beta = 3.14^{+0.31}_{-0.27} \times 10^{13}$, with a large intrinsic scatter.

α was fixed to 1 following scaling relation theory. Assuming a self-similar scaling relation between mass and temperature, starting with the virial theorem and setting the specific energy of the ICM to be equal to the specific energy of the galaxies, yielded:

$$M_{vir} \propto T_X r_{vir} \tag{5.3}$$

The virial radius could be written as:

$$M_{vir} \propto r_{vir}^3 \tag{5.4}$$

This resulted in the predictions for a virialised cluster (Kaiser, 1986) and also other comparisons utilizing weak lensing data (Jee et al., 2011) of:

$$M_{vir} \propto T_X^{3/2} \tag{5.5}$$

Because all the temperature and mass measurements in this sample were taken inside a fixed sized aperture, Eqn.5.3 becomes $M_{vir} \propto T_X$ and hence $\alpha = 1$.

The scatter from the fit to mass temperature relation was better understood via two factors. Some of the points were disturbed clusters and hence broke the assumptions needed for a correlation to exist. The other factor was the tri-axiality of DM halos being observed at differing angles causing an under- or over-estimation of the mass with respect to the true mass for a given temperature. In a relaxed system the ellipticity of the BCG appeared to correlate well with this aspect and could be a useful corrective tool if the disturbed clusters can be isolated from the relaxed ones on the $\Delta M/M_{fit}$ (see Table 5.3) against BCG ellipticity fit (a cut of $m_{12} \geq 1$ was used). This was hinted at in the results of Richard et al. (2009) and was shown in Marrone et al. (2012) to be true for Sunyaev Zeldovitch and weak lensing mass comparison. Weak lensing by its statistical nature does not have the same resolution inside the cores of clusters that strong lensing (which only really probes the cores of clusters) does, but the agreement both have with BCG ellipticity and mass over- or under-estimation is a useful double confirmation of this effect in clusters.

A corrective factor to the mass relation fit was not applied due to the inherent dangers of calibrating and fitting data to itself. If after fitting a set of data, a correcting value for the scatter utilising the same data will normally produce an improvement in scatter of the fit regardless of whether it was true or not. Simulations of BCG ellipticity with dark matter halo orientation, and any resulting corrective factor, would be the next useful step (as discussed in §6.1).

5.8 Chapter Summary

The following conclusions can be drawn from the analyses:

(i) This work increased the number of LoCuSS SL multi-wavelength clusters from 17 to 23, with five new to this thesis and one newly published in Paraficz et al. (2012). An increase of 35%.

(ii) Fitting the mass to temperature provided a mass to temperature scaling relation,

but this could only fit the intercept $\beta = 3.14^{+0.31}_{-0.27} \times 10^{13}$ due to the concentration of points near the high mass end of the expected limits and the fixed measurement aperture.

(iii) Large scatter was found when fitting for strong lensing mass against X-ray temperature.

(iv) Examining residuals to the mass fitting, the closest match was found to BCG ellipticity, with outliers.

(v) Luminosity gap provided evidence for a link to substructure fraction and also towards whether a cluster was relaxed or not.

(vi) The assumption of BCG ellipticity following DM tri-axiality only worked for a cluster that was relaxed (same as for the mass versus temperature relation). This was in agreement with Marrone et al. (2012).

(vii) Removing clusters that showed signs of disturbances produced a "strong" negative correlation (Pearson value of 0.71) with BCG ellipticity and whether mass was over- or under-estimated when compared to the fitted mass vs temperature scaling relation.

(viii) Prior comparisons of X-ray and lensing values compared their calculated mass estimations. Fitting directly to T_X provided a more direct comparison utilising fewer assumptions.

CHAPTER 6

DISCUSSION OF WORK

This thesis used a range of data, both that gathered as part of this dissertation and that from collaborators. Despite this, there was still a paucity for some of the interpretation and this is discussed below. The work dealt with eight clusters, five with recorded strong lensing image constraints (A 2895, A 2813, A 368, A 3084 and A 2537) and three without (A 3364, A 3088 and RXCJ0528). Abell 3084 stood out as an unusual strong lensing cluster.

The redshift histograms (chapter 3) retrieved mean redshifts that within their 2σ errors were close to seven of the eight cluster literature BCG values (Guzzo et al., 2009). The exception from the eight was A 3364, which had an unusual BCG value (discussed in §3.7.7). In this case the value was close to the mean value for the other REFLEX galaxies of $z = 0.1489$, with $z_{hist} = 0.1484 \pm 0.0038$. This gave credence to the measured galaxy redshifts being galaxy members of the clusters.

Examination of the estimated velocity dispersion showed no strong link to the $\sigma - T_X$ relation from Xue & Wu (2000). However, this did not mean that there was not a link in our data (Xue & Wu found one in their work), it might be masked by the selection effects (a velocity dispersion was defined to make a redshift cut width to select cluster galaxy membership) and the simple method of calculating the velocity dispersion. Also the low numbers of galaxies measured might mean that the population was insufficiently sampled.

This sampling effect, coupled with uneven coverage on the sky meant that the Dressler Shectman (DS) tests might have suffered from biases when calculating. Any areas not covered by redshifts would render the DS tests blind to any substructure situated there but also, if the numbers were poorly distributed, then the nearest neighbours might be non-ideally selected. Reducing the nearest neighbour count increased the sensitivity to substructure, but below 20 cluster members the tests only provided a lower limit on the cluster substructure. Removing REFLEX galaxy members affected the DS statistics, but not in a consistent way, and any effects might be marginal when compared to the membership reduction and distribution.

Examination of BCG and X-ray centroid offsets in the clusters revealed no strong link between them and the DS tests. The DS tests might not show a link simply because the region that the centroid offsets were within ($\sim$ 1 arcmin radii), were devoid of many redshifts. Simply put, the DS test was blind within this region to cluster dynamical activity.

The results of this test comparison of dynamical status of cluster cores showed that the DS test could possibly provide a good view at the centre of clusters, provided there was sufficient cluster members in the centre of the cluster. Without information in the centre of the clusters the DS tests (and statistics) were blind to substructure situated there, in addition, clusters that had member numbers below 20 would only return a lower limit on the possible substructure present within the cluster (as discussed in section 3.3).

This redshift sample of eight clusters, possessed too low a number of cluster member galaxies per cluster to draw definite conclusions and so required expansion with more galaxies. The analysis of these redshifts (by DS tests) was useful as it provided information on possible priors when beginning mass model selection. Information such as possibly providing a useful guide as to the number and position for halos in many cases (as long as the DS tests were sensitive to those areas). Deciding on priors was one of the big problems when building lensing models.

The most unusual cluster (due to its lensing configuration) was parameterised first

using any multiple image redshifts (coming from the spectroscopic data reduction) to produce lensing models. The hyperbolic umbilic catastrophe (a rare image configuration) was spectroscopically confirmed at $z = 0.764$ (putting strong constraints on the mass within $\lesssim$ 20 kpc scales around the BCG). In addition a set of triple images arranged in a straight line (usually indicating a saddle point in the mass distribution) suggested the mass distribution was elongated and possible bi-modal. This elongation/bi-modality would have its major axis (as seen on the sky) aligned roughly North-East/South-West. When models were fitted, both single- and triple-cluster scale halo models were found to be a poorer fit than the double-cluster halo models. The preferred model had one cluster-scale halo on the BCG and another on the X-ray centroid. The two were offset from one another by 42 kpc suggesting the BCG had been disturbed and had been shifted from the ICM gas, with the BCG being coincident with a very compact DM halo possessing a $r_{core} \sim 7$ kpc. Comparison with the X-ray emission, which overall was relaxed and showed no obvious signs of fronts (though the shallow data would only be sensitive to very strong shocks or cold fronts), did not suggest a disturbed merging cluster. However, it might suggest a late cluster-cluster merger that was settling (discussed in later paragraphs).

Overall, of the two cluster-scale halos used in the mass model, the one centred on the X-ray centroid had more traditional cluster values, but was not as well constrained by the triple image system as the halo, coincident with the BCG, was constrained by the quintet system constraints. This BCG coincident halo had unusual values but they were well constrained by the quintet image system that laid very close to the centre of the cluster.

A Dressler Shectman plot was produced to investigate the dynamical structure of the cluster further and this showed the strongest evidence for departures from the global cluster kinematics to the South-East of the BCG. In fact, galaxies with the largest deviations from global kinematics dominated the lower of the two redshift histogram peaks, with the dynamically "colder" population presenting just one peak in the redshift distribution. Detailed interpretation was complicated by the incompleteness of the spectroscopic ob-

servations within 1arcmin of the BCG. Nevertheless, from the current data it could be concluded that the cluster was dynamically active along the line of sight, and that the BCG laid closer in redshift to the more prominent, higher redshift of two peaks (offset by $\sim 300\,\mathrm{km\,s^{-1}}$) than to the lower redshift peak (offset by $\sim 1000\,\mathrm{km\,s^{-1}}$).

One of the possible interpretations of what was occuring in A 3084 was that the BCG was not associated with the cluster and was in fact being observed in projection along the line of sight on its own (or two clusters superimposed). This interpretation had the problem of no other suitable "BCG" candidates near the cluster and so would imply it did not possess one (or only one of the two did in the twin cluster version). If true, this would be very interesting for cluster formation history as the lack of a BCG would mean that clusters could form without a BCG and the BCG falls in preformed rather than being created by the merging of smaller galaxies inside the cluster potential well. This would tie in with the other interpretation that the cluster had recently acquired its BCG and it was in the process of settling to the bottom of the potential well. This would explain the undisturbed indicators (smooth X-ray contours, small velocity dispersion from expected cluster values etc.) of other cluster measurements as the cluster was settled, but the BCG was not.

The simpler interpretation was that within the cluster the BCG was just that, the cluster BCG, and was simply settling after being displaced by a cluster-cluster merger. The gas already having settled would imply the BCG had taken longer to settle and, on a basic interpretation of galaxies being collisionless, they could be expected to take longer to settle than the gas (which interacted strongly with itself electromagnetically) and this was what was being seen. Whether the cluster was at the very early stages of a merger or at the very end of a merger having had a very long time to settle was the next question. If at the beginning of this merger, a second BCG associated with the other merging system was not visible, this invoked a similar argument that had been made before; that there was a cluster without a BCG. Also there was only one centroid in the X-ray, which appeared smooth and well described as being one not two lots of gas (although the two could be

superimposed). This suggested that the cluster was settling after a merger and the BCG had merged with the other cluster BCG forming one single BCG (this would also explain the large luminosity gap observed for this cluster as discussed below).

Four other clusters were reduced and added, along with another from the literature, to the members with strong lensing constraints. These were then added to an existing sample and increased its size. This expanded sample was used to construct a mass temperature scaling relation, with the slope fixed at the predicted and observed values ($\alpha = 1.0$, because the fixed aperture size changed the relation to $M \propto T$), a fit of the intercept was made at $\beta = 3.14^{+0.31}_{-0.27} \times 10^{13}$ $M_{\odot}$. The resulting fit had large intrinsic scatter and the work set out to understand this, finding a link between over- or under-estimation of the mass against predicted mass with the BCG ellipticity. The link was that BCG ellipticity traced the tri-axiality of the cluster-scale DM halo. This relation had large numbers of outliers preventing a firm quantitative link until those cluster which were disturbed were removed (using luminosity gap values as an indicator of dynamical activity). A cluster undergoing disturbance would no longer follow the predicted relation $M \propto T^{3/2}$, as the cluster undergoing shocks by the merger no longer traces the cluster potential. Looking at the dominance of the BCG by use of the luminosity gap showed a link (Smith et al., 2010) between relaxed and unrelaxed clusters, by having a greater or smaller luminosity gap respectively. Taking a cut of $m_{12} \geq 1$ removed many of the outliers and several of the remaining outliers could be explained by being bimodal clusters (A 3084 being one such outlier).

Comparing this to existing work was not straight forward due to the different mass measurements used (Arnaud et al., 2005 and Jee et al., 2011) used M_{2500} whereas here the measurements were $M_{2D,SL}(r < 250 \text{ kpc})$) and is suggested for future work.

This work highlighted how the remaining redshifts obtained by measuring redshifts for strong lensing image constraints could provide qualitative information when constructing gravitational lensing models. Since these redshifts were essentially "free" from trying to match and identify strong lensing features they could be a useful tool, although with

several caveats on their use regarding poor coverage and large blank areas with diameters of the order 1 arcminute. These were centred around the BCG's, due to the priority of mask slit placement.

This work also provided an in depth analysis of an unusual and predicted to be "rare" cluster, Abell 3084, that attempted to explain the clusters' infall history and what was observed in the epoch of observation. The shallow nature of the survey data could not provide answers to some important questions about this cluster, such as "Were there shocks?", "Was the X-ray peak nuclear activity or a cool core in the BCG?", but provided good evidence of the cluster being observed a very long time after suffering a cluster-cluster merger and so gave unique insights (discussed below) into this relatively short-lived event in a cluster's life.

Adding Abell 3084 and the other four strong lensing clusters to the existing LoCuSS sample of strong lensing clusters, provided a much expanded table of cluster observables all utilising the same methods. Examining links between other cluster observables and the strong lensing derived masses showed the same link, between BCG ellipticity and the deviation from a fitted mass and observed mass, as those between similar previous observations undertaken with weak lensing observations. There was also a link between BCG ellipticity and the deviation from a fitted mass and observed mass. This was in agreement with a very recent publication by Marrone et al. (2012) utilising weak and SZE data. Further work is required to confirm this link and the assumptions behind it but, if this relation is real, could provide important information in understanding the scatter in mass temperature correlations and even in correcting this observed scatter, possibly resulting in a robust mass temperature relation.

In short, this work has shown the usefulness of analysing cluster redshifts in order to provide useful priors for lensing constraints. For Abell 3084 this analysis showed that not only was it a rare cluster because of its "hyperbolic umbilic catastrophe" image arrangement (a useful analysis in its own right), but also could be a cluster in a rare epoch of cluster formation. The mass temperature relation with purely strong lensing clusters

showed large intrinsic scatter, which when investigated could be attributed to possible DM halo tri-axiality. The ability, in relaxed clusters, of the BCG ellipticity to trace this tri-axiality could become a useful tool in scaling masses with other cluster observables. This was in agreement with recent work utilising weak lensing data and suggested it was not a quirk of the strong lensing sample contained within this thesis.

From the questions outlined in § 1.8 this work has resulted in the following answers:

That A 3084 provided a unique insight into a cluster that was resettling after a merger: was this a rare event to observe? This question was raised because without the addition of multi-wavelength data to the analysis the cluster would appear as a simple disturbed cluster with two DM halos. The additional information brings the insight that the cluster appears relaxed in disagreement with a purely strong lensing interpretation (the high substructure fraction would indicate its dynamically active) and also an observationally well motivated placement for the second DM halo. This opened the question that other disturbed clusters might in fact be clusters resettling a long time after a merger and could explain the BCG offset observed in other clusters as the residual sloshing of the BCG around the cluster's potential well. This implied that further study of the activity within cluster cores utilising multi-wavelength data could reveal more about the history of a cluster's merger and the exact manner of the resettling post-merger. Did the BCG settle to the bottom of the potential well before the gas or was it the other way around, which would put constraints on numerical simulations of cluster-cluster mergers.

The strong lensing sample was one of the largest samples of clusters reduced consistently utilising similar data without using literature values and conversions. This provided a sample of clusters with multi-wavelength measurements. A correlation was found between strong lensing mass and temperature (as predicted) but that large scatter resulted. The multi-wavelength approach enabled an examination of why the large scatter occured. This resulted in two useful observable quantities, that of luminosity gap and that of BCG ellipticity. Both were measurable from purely optical based ground instrumentation. The luminosity gap correlated well with substructure fraction (and hence if a cluster was re-

laxed or not), providing an insight into whether a cluster was relaxed or not simply by the dominance of the BCG. Because not all clusters possess strong lensing constraints to measure substructure in the cores of clusters, luminosity gap was an easier measurement to take as all clusters has a luminosity. This was important as weak lensing measurements did not have the ability to probe the cores of clusters where residual substructures might lie. This made it blind to post-merger substructure after a certain period because, from hierarchical merging, substructure would slowly accrete to the centre of the potential well over time. For the mass measurements themselves, DM halos were believed to be tri-axial which when modelled using elliptical mass distributions would incur a mass over- or under-estimation depending on viewing angle. Having the BCG trace the long axis of the mass distribution would allow tri-axial parametrised models to be fitted to data, based on information from the BCG ellipticity. If found to be true, this would reduce the scatter in mass-observable relations allowing more accurate mass estimations utilising the fits to these relations.

Putting the results together, chapter 3 highlighted the need for spectroscopically complete samples to infer quantitative values from the data. Qualitatively they could help with mass model priors before parametrisation if no other multi-wavelength data were available. However, this can be insensitive to substructure and its usefulness is strictly limited without additional galaxy redshifts.

Chapter 4 showed a cluster in the late post-merger stage, while rare, it highlighted that BCG offset from X-ray centroid might provide an insight into the merger history of a cluster. Apart from strong lensing data and BCG offset from X-ray centroid, all other signs indicated the cluster was relaxed. This highlighted the possible need to look at other clusters that lacked multi-wavelength data to determine if some might be hidden late post-mergers. Such a result would show that they are more common than currently observed and require a multi-wavelength technique in order to detect them. If A 3084 was indeed in a late post-merger stage it gave information on what occurs inside the cluster cores at these epochs. Once the gas had settled to the centre of the potential well,

the BCG was still sloshing about and possessed a compact halo. Could this be due to the BCGs original DM halo being tidally stripped as it still moves about the cluster's potential well?

In chapter 5 a large sample of strong lensing clusters enabled the study of the M-T_X relation and of the residuals causing scatter. While the fit itself required lower end masses to help constrain it but this sample only populated the the extreme end of the relation. The mass measurement aperture forces the slope of the fit to be unity. The luminosity gap was found to indicate the dynamical state of the clusters, i.e. several outliers were known to require multi-halo models and, therefore, were not going to be relaxed. Utilising luminosity gap as a an indicator of cluster dynamical activity to remove outliers, showed that BCG ellipticity correlated well with the offset of measured mass from fitted mass. Comparing this result to simulations would provide a useful check as to the validity of utilising these two quantities (luminosity gap and BCG ellipticity) as a tool for the correction of mass measurements or as a prior before making mass-observable relations. It could also be used for estimating substructure content or for providing information about DM halo orientation. Combined this would reduce the scatter in the M-T_X relation.

6.1 Future Work

An important point for further work should be to provide more redshifts and deeper imaging for A 3084. Indeed time for further GMOS imaging was granted in order to take spectrum of galaxies in the core of the cluster which, at the time of this project, was under sampled. Also further deeper optical imaging might yield additional strong lensing images. This would not only provide tighter constraints on the mass distributed in the cluster, but would allow further investigation of the dynamical status in the centre of the cluster. X-ray data on A 3084 was only $\sim$ 20 ksec deep (shallow by X-ray standards) and deeper imaging would increase the 13 photons centred on the BCG and allow the re-centring of the X-ray model to be placed on the BCG for comparison and interpretation

with the SL models. It will also allow confirmation whether the sharp spike in the X-ray imaging coincident with the BCG is a point source or something else. Deeper X-ray data should allow examination of whether the data is sensitive enough to detect any cold fronts. Another useful task would be to look at how removing or adding in a low number of clusters from a DS plot changed the resulting statistics in order to understand the error range that can be expected from this.

Modelling and adding additional clusters to the LoCuSS sample would allow the temperature, mass relation to be more tightly refined and help verify if dominant BCG's trace DM halo tri-axiality. Using a non-fixed mass measurement aperture to measure the strong lensing derived massed, which combined with smaller clusters (from weak or strong lensing studies) would enable not only an intercept, but also a slope to be determined for the mass temperature relation. It would also allow further stacking of the DS statistics of any cluster galaxy redshifts that were obtained as a part of this modelling. Filling in the gaps around the centre of the SL clusters would allow the relationship (if any) between the Dressler Shectman measure of substructure and SL models to be compared, while also increasing the accuracy of the DS test on small scales.

To use a correction for the mass over or under estimation from the fitted mass temperature relation requires comparison with N body simulations. These would look into whether the assumption of BCG ellipticity orientation with halo tri-axiality was the effect being observed in Fig.5.11. If this turns out to be a correct assumption, then utilising the BCG ellipticity to correct for deviations due to tri-axial halo position would enable a tighter correlation (with smaller scatter and errors) for strong lensing masses versus X-ray temperatures to be made. Using m_{12} as an indicator of a cluster's dynamical state enabled some of the outliers in the relation to be removed. f_{sub} could also be used for this (it showed a "moderate" correlation with m_{12}), but required mass modelling and can be more observationally expensive than calculating m_{12} as a result. This made m_{12} a useful indicator of cluster dynamical state.

Other work fitting this mass temperature relation had previously been done with

weak lensing clusters and this strong lensing sample represents a larger sample than had been used previously (less than ten for cited works). Previous studies (Arnaud et al. 2005 and Jee et al. 2011) had utilised measurements with M_{2500}, whereas this work used $M_{2D,SL}(r < 250$ kpc). Re-measuring the sample to fit the same aperture sizes would enable a more direct comparison to previous work.

APPENDIX

ADDENDUM

APPENDIX

LIST OF REFERENCES

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

Abell, G. O., Corwin, Jr., H. G., & Olowin, R. P. 1989, APJS, 70, 1

Ackermann, M., Ajello, M., Albert, A., Atwood, W. B., Baldini, L., Ballet, J., Barbiellini, G., Bastieri, D., Bechtol, K., Bellazzini, R., Berenji, B., Blandford, R. D., Bloom, E. D., Bonamente, E., Borgland, A. W., Bregeon, J., Brigida, M., Bruel, P., Buehler, R., Burnett, T. H., Buson, S., Caliandro, G. A., Cameron, R. A., Cañadas, B., Caraveo, P. A., Casandjian, J. M., Cecchi, C., Charles, E., Chekhtman, A., Chiang, J., Ciprini, S., Claus, R., Cohen-Tanugi, J., Conrad, J., Cutini, S., de Angelis, A., de Palma, F., Dermer, C. D., Digel, S. W., Do Couto E Silva, E., Drell, P. S., Drlica-Wagner, A., Falletti, L., Favuzzi, C., Fegan, S. J., Ferrara, E. C., Fukazawa, Y., Funk, S., Fusco, P., Gargano, F., Gasparrini, D., Gehrels, N., Germani, S., Giglietto, N., Giordano, F., Giroletti, M., Glanzman, T., Godfrey, G., Grenier, I. A., Guiriec, S., Gustafsson, M., Hadasch, D., Hayashida, M., Hays, E., Hughes, R. E., Jeltema, T. E., Jóhannesson, G., Johnson, R. P., Johnson, A. S., Kamae, T., Katagiri, H., Kataoka, J., Knödlseder, J., Kuss, M., Lande, J., Latronico, L., Lionetto, A. M., Llena Garde, M., Longo, F., Loparco, F., Lott, B., Lovellette, M. N., Lubrano, P., Madejski, G. M., Mazziotta, M. N., McEnery, J. E., Mehault, J., Michelson, P. F., Mitthumsiri, W., Mizuno, T., Monte, C., Monzani, M. E., Morselli, A., Moskalenko, I. V., Murgia, S., Naumann-Godo, M., Norris, J. P., Nuss, E., Ohsugi, T., Okumura, A., Omodei, N., Orlando,

E., Ormes, J. F., Ozaki, M., Paneque, D., Parent, D., Pesce-Rollins, M., Pierbattista, M., Piron, F., Pivato, G., Porter, T. A., Profumo, S., Rainò, S., Razzano, M., Reimer, A., Reimer, O., Ritz, S., Roth, M., Sadrozinski, H. F.-W., Sbarra, C., Scargle, J. D., Schalk, T. L., Sgrò, C., Siskind, E. J., Spandre, G., Spinelli, P., Strigari, L., Suson, D. J., Tajima, H., Takahashi, H., Tanaka, T., Thayer, J. G., Thayer, J. B., Thompson, D. J., Tibaldo, L., Tinivella, M., Torres, D. F., Troja, E., Uchiyama, Y., Vandenbroucke, J., Vasileiou, V., Vianello, G., Vitale, V., Waite, A. P., Wang, P., Winer, B. L., Wood, K. S., Wood, M., Yang, Z., Zimmer, S., Kaplinghat, M., & Martinez, G. D. 2011, Physical Review Letters, 107, 241302

Allen, S. W. 1998, MNRAS, 296, 392

Appenzeller, I., Fricke, K., Fürtig, W., Gässler, W., Häfner, R., Harke, R., Hess, H.-J., Hummel, W., Jürgens, P., Kudritzki, R.-P., Mantel, K.-H., Meisl, W., Muschielok, B., Nicklas, H., Rupprecht, G., Seifert, W., Stahl, O., Szeifert, T., & Tarantik, K. 1998, The Messenger, 94, 1

Arnaud, M., Pointecouteau, E., & Pratt, G. W. 2005, A&A, 441, 893

Ascasibar, Y. & Diego, J. M. 2008, MNRAS, 383, 369

Ascasibar, Y. & Markevitch, M. 2006, ApJ, 650, 102

Babcock, H. W. 1939, Lick Observatory Bulletin, 19, 41

Bahcall, N. A. 1977, ARA&A, 15, 505

Barkana, R. & Loeb, A. 2010, MNRAS, 405, 1969

Bartelmann, M., Limousin, M., Meneghetti, M., & Schmidt, R. 2013, SSRv

Bautz, L. P. & Morgan, W. W. 1970, ApJL, 162, L149

Bennett, C. L., Boggess, N. W., Cheng, E. S., Hauser, M. G., Kelsall, T., Mather, J. C., Moseley, Jr., S. H., Murdock, T. L., Shafer, R. A., & Silverberg, R. F. 1993, Advances in Space Research, 13, 409

Bertschinger, E. 2006, Phys. Rev. D, 74, 063509

Birkinshaw, M., Hughes, J. P., & Arnaud, K. A. 1991, ApJ, 379, 466

Böhringer, H., Schuecker, P., Guzzo, L., Collins, C. A., Voges, W., Cruddace, R. G., Ortiz-Gil, A., Chincarini, G., De Grandi, S., Edge, A. C., MacGillivray, H. T., Neumann, D. M., Schindler, S., & Shaver, P. 2004, A&A, 425, 367

Bradač, M., Allen, S. W., Treu, T., Ebeling, H., Massey, R., Morris, R. G., von der Linden, A., & Applegate, D. 2008, ApJ, 687, 959

Bradač, M., Clowe, D., Gonzalez, A. H., Marshall, P., Forman, W., Jones, C., Markevitch, M., Randall, S., Schrabback, T., & Zaritsky, D. 2006, ApJ, 652, 937

Bradač, M., Lombardi, M., & Schneider, P. 2004, A&A, 424, 13

Bradley, L. D., Bouwens, R. J., Ford, H. C., Illingworth, G. D., Jee, M. J., Benítez, N., Broadhurst, T. J., Franx, M., Frye, B. L., Infante, L., Motta, V., Rosati, P., White, R. L., & Zheng, W. 2008, ApJ, 678, 647

Broadhurst, T., Benítez, N., Coe, D., Sharon, K., Zekser, K., White, R., Ford, H., Bouwens, R., Blakeslee, J., Clampin, M., Cross, N., Franx, M., Frye, B., Hartig, G., Illingworth, G., Infante, L., Menanteau, F., Meurer, G., Postman, M., Ardila, D. R., Bartko, F., Brown, R. A., Burrows, C. J., Cheng, E. S., Feldman, P. D., Golimowski, D. A., Goto, T., Gronwall, C., Herranz, D., Holden, B., Homeier, N., Krist, J. E., Lesser, M. P., Martel, A. R., Miley, G. K., Rosati, P., Sirianni, M., Sparks, W. B., Steindling, S., Tran, H. D., Tsvetanov, Z. I., & Zheng, W. 2005, ApJ, 621, 53

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

Burke, W. L. 1981, ApJL, 244, L1

Burns, J. O., Hallman, E. J., Gantner, B., Motl, P. M., & Norman, M. L. 2008, ApJ, 675, 1125

Calchi Novati, S., Mancini, L., Scarpetta, G., & Wyrzykowski, Ł. 2009, MNRAS, 400, 1625

Carlstrom, J. E., Holder, G. P., & Reese, E. D. 2002, ARA&A, 40, 643

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

Clowe, D., Luppino, G. A., Kaiser, N., & Gioia, I. M. 2000, ApJ, 539, 540

Combes, F., Boissé, P., Mazure, A., & Blanchard, A. 1995, Galaxies and cosmology (Springer)

Crane, P., Albrecht, R., Barbieri, C., Blades, J. C., Boksenburg, A., Deharveng, J., Disney, M. J., Jakobsen, P., Kamperman, T. M., King, I. R., Macchetto, F., Mackay, C. D., Paresce, F., Weigelt, G., Baxter, D., Greenfield, P., Jedrzejewski, R., Nota, A., & Sparks, W. B. 1990, in Bulletin of the American Astronomical Society, Vol. 22, Bulletin of the American Astronomical Society, 1280

Cruddace, R., Voges, W., Böhringer, H., Collins, C. A., Romer, A. K., MacGillivray, H., Yentis, D., Schuecker, P., Ebeling, H., & De Grandi, S. 2002, APJS, 140, 239

Dahle, H., Kaiser, N., Irgens, R. J., Lilje, P. B., & Maddox, S. J. 2002, APJS, 139, 313

Dressler, A. & Shectman, S. A. 1988, AJ, 95, 985

Ebeling, H., Edge, A. C., Allen, S. W., Crawford, C. S., Fabian, A. C., & Huchra, J. P. 2000, MNRAS, 318, 333

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

Eddington, A. S. 1919, The Observatory, 42, 119

Eddington, A. S., Jeans, J. H., Lodge, Sir, O., Larmor, Sir, J., Silberstein, L., Lindemann, F. A., & Jeffreys, H. 1919, MNRAS, 80, 96

Einstein, A. 1915, Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin), Seite 778-786., 0, 778

Elíasdóttir, Á., Limousin, M., Richard, J., Hjorth, J., Kneib, J.-P., Natarajan, P., Pedersen, K., Jullo, E., & Paraficz, D. 2007, ArXiv e-prints

Evans, J. D. 1996, Straightforward Statistics for the Behavioral Sciences (Brooks/Cole Pub. Co. (Pacific Grove))

Felten, J. E., Gould, R. J., Stein, W. A., & Woolf, N. J. 1966, ApJ, 146, 955

Fesenko, B. I. 1963, Soviet Ast., 6, 602

Fixsen, D. J. 2009, ApJ, 707, 916

Foster, A. R., Ji, L., Smith, R. K., & Brickhouse, N. S. 2012, The Astrophysical Journal, 756, 128

Franx, M., Illingworth, G. D., Kelson, D. D., van Dokkum, P. G., & Tran, K. 1997, ApJL, 486, L75+

Frenk, C. S., White, S. D. M., Davis, M., & Efstathiou, G. 1988, ApJ, 327, 507

Gamow, G. 1970, My world line; an informal autobiography (Viking Press)

Gregory, P. 2005, Bayesian Logical Data Analysis for the Physical Sciences: A Comparative Approach with Mathematica® Support (Cambridge University Press)

Grevesse, N. & Sauval, A. J. 1998, SSRv, 85, 161

Griest, K., Lehner, M. J., Cieplak, A. M., & Jain, B. 2011, Physical Review Letters, 107, 231101

Guzzo, L., Schuecker, P., Böhringer, H., Collins, C. A., Ortiz-Gil, A., de Grandi, S., Edge, A. C., Neumann, D. M., Schindler, S., Altucci, C., & Shaver, P. A. 2009, A&A, 499, 357

Harsono, D. & De Propris, R. 2009, AJ, 137, 3091

Hernquist, L. 1990, ApJ, 356, 359

Hou, A., Parker, L. C., Wilman, D. J., McGee, S. L., Harris, W. E., Connelly, J. L., Balogh, M. L., Mulchaey, J. S., & Bower, R. G. 2012, MNRAS, 421, 3594

Hubble, E. 1929, Proceedings of the National Academy of Science, 15, 168

Jee, M. J., Dawson, K. S., Hoekstra, H., Perlmutter, S., Rosati, P., Brodwin, M., Suzuki, N., Koester, B., Postman, M., Lubin, L., Meyers, J., Stanford, S. A., Barbary, K., Barrientos, F., Eisenhardt, P., Ford, H. C., Gilbank, D. G., Gladders, M. D., Gonzalez, A., Harris, D. W., Huang, X., Lidman, C., Rykoff, E. S., Rubin, D., & Spadafora, A. L. 2011, ApJ, 737, 59

Jee, M. J., Rosati, P., Ford, H. C., Dawson, K. S., Lidman, C., Perlmutter, S., Demarco, R., Strazzullo, V., Mullis, C., Böhringer, H., & Fassbender, R. 2009, ApJ, 704, 672

Jullo, E., Kneib, J.-P., Limousin, M., Elíasdóttir, Á., Marshall, P. J., & Verdugo, T. 2007, New Journal of Physics, 9, 447

Kaiser, N. 1986, MNRAS, 222, 323

Kalberla, P. M. W., Burton, W. B., Hartmann, D., Arnal, E. M., Bajaja, E., Morras, R., & Pöppel, W. G. L. 2005, A&A, 440, 775

Kassiola, A. & Kovner, I. 1993, ApJ, 417, 450

Kelson, D. D. 2003, PASP, 115, 688

Kitzbichler, M. G. & White, S. D. M. 2008, MNRAS, 391, 1489

Knebe, A. & Müller, V. 2000, A&A, 354, 761

Kneib, J., Ellis, R. S., Santos, M. R., & Richard, J. 2004, ApJ, 607, 697

Kneib, J., Ellis, R. S., Smail, I., Couch, W. J., & Sharples, R. M. 1996, ApJ, 471, 643

Kneib, J. P., Mellier, Y., Pello, R., Miralda-Escude, J., Le Borgne, J., Boehringer, H., & Picat, J. 1995, A&A, 303, 27

Koekemoer, A. M., Fruchter, A. S., Hook, R. N., Hack, W., & Hanley, C. 2006, in The 2005 HST Calibration Workshop: Hubble After the Transition to Two-Gyro Mode, ed. A. M. Koekemoer, P. Goudfrooij, & L. L. Dressel, 423

Kormann, R., Schneider, P., & Bartelmann, M. 1994, A&A, 284, 285

Kravtsov, A. V. & Borgani, S. 2012, ARA&A, 50, 353

Larson, D., Dunkley, J., Hinshaw, G., Komatsu, E., Nolta, M. R., Bennett, C. L., Gold, B., Halpern, M., Hill, R. S., Jarosik, N., Kogut, A., Limon, M., Meyer, S. S., Odegard, N., Page, L., Smith, K. M., Spergel, D. N., Tucker, G. S., Weiland, J. L., Wollack, E., & Wright, E. L. 2011, APJS, 192, 16

Limousin, M., Richard, J., Kneib, J.-P., Brink, H., Pelló, R., Jullo, E., Tu, H., Sommer-Larsen, J., Egami, E., Michałowski, M. J., Cabanac, R., & Stark, D. P. 2008, A&A, 489, 23

Loveday, J. 2000, MNRAS, 312, 557

Lynds, R. & Petrosian, V. 1989, ApJ, 336, 1

Mahdavi, A., Hoekstra, H., Babul, A., & Henry, J. P. 2008, MNRAS, 384, 1567

Mahdavi, A., Hoekstra, H., Babul, A., Sievers, J., Myers, S. T., & Henry, J. P. 2007, ApJ, 664, 162

Markevitch, M., Gonzalez, A. H., David, L., Vikhlinin, A., Murray, S., Forman, W., Jones, C., & Tucker, W. 2002, ApJL, 567, L27

Marrone, D. P., Smith, G. P., Okabe, N., Bonamente, M., Carlstrom, J. E., Culverhouse, T. L., Gralla, M., Greer, C. H., Hasler, N., Hawkins, D., Hennessy, R., Joy, M., Lamb, J. W., Leitch, E. M., Martino, R., Mazzotta, P., Miller, A., Mroczkowski, T., Muchovej, S., Plagge, T., Pryke, C., Sanderson, A. J. R., Takada, M., Woody, D., & Zhang, Y. 2012, ApJ, 754, 119

Marrone, D. P., Smith, G. P., Richard, J., Joy, M., Bonamente, M., Hasler, N., Hamilton-Morris, V., Kneib, J.-P., Culverhouse, T., Carlstrom, J. E., Greer, C., Hawkins, D., Hennessy, R., Lamb, J. W., Leitch, E. M., Loh, M., Miller, A., Mroczkowski, T., Muchovej, S., Pryke, C., Sharp, M. K., & Woody, D. 2009, ApJL, 701, L114

Martizzi, D., Teyssier, R., & Moore, B. 2012, MNRAS, 420, 2859

Massey, R., Kitching, T., & Nagai, D. 2011, MNRAS, 413, 1709

Massey, R., Kitching, T., & Richard, J. 2010, Reports on Progress in Physics, 73, 086901

Merrall, T. E. C. & Henriksen, R. N. 2003, ApJ, 595, 43

Meylan, G., Jetzer, P., North, P., Schneider, P., Kochanek, C. S., & Wambsganss, J., eds. 2006, Gravitational Lensing: Strong, Weak and Micro

Miralda-Escude, J. & Babul, A. 1995, ApJ, 449, 18

Narayan, R. & Bartelmann, M. 1996, ArXiv Astrophysics e-prints

Natarajan, P. & Kneib, J.-P. 1996, MNRAS, 283, 1031

Natarajan, P., Kneib, J.-P., & Smail, I. 2002, ApJL, 580, L11

Navarro, J. F., Frenk, C. S., & White, S. D. M. 1996, ApJ, 462, 563

—. 1997, ApJ, 490, 493

Okabe, N., Takada, M., Umetsu, K., Futamase, T., & Smith, G. P. 2010, PASJ, 62, 811

Olling, R. P. & Merrifield, M. R. 1998, in Astronomical Society of the Pacific Conference Series, Vol. 136, Galactic Halos, ed. D. Zaritsky, 219

Orban de Xivry, G. & Marshall, P. 2009, MNRAS, 399, 2

Ostriker, J. P. & Hausman, M. A. 1977, ApJL, 217, L125

Paraficz, D., Kneib, J.-P., Richard, J., Morandi, A., Limousin, M., & Jullo, E. 2012, ArXiv e-prints

Peterson, J. R. & Fabian, A. C. 2006, Phys. Rep., 427, 1

Phillips, A. C. 2003, The Physics of Stars, Second Edition (John Wiley & Sons ltd)

Pinkney, J., Roettiger, K., Burns, J. O., & Bird, C. M. 1996, APJS, 104, 1

Ponman, T. J., Allan, D. J., Jones, L. R., Merrifield, M., McHardy, I. M., Lehto, H. J., & Luppino, G. A. 1994, Nature, 369, 462

Poole, G. B., Babul, A., McCarthy, I. G., Sanderson, A. J. R., & Fardal, M. A. 2008, MNRAS, 391, 1163

Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. 1992, Numerical Recipes in FORTRAN. The Art of Scientific Computing, 2nd edn. (University Press, Cambridge)

Ragozzine, B., Clowe, D., Markevitch, M., Gonzalez, A. H., & Bradač, M. 2012, ApJ, 744, 94

Randall, S. W., Markevitch, M., Clowe, D., Gonzalez, A. H., & Bradač, M. 2008, ApJ, 679, 1173

Rasmussen, J., Mulchaey, J. S., Bai, L., Ponman, T. J., Raychaudhury, S., & Dariush, A. 2010, ApJ, 717, 958

Rau, S., Vegetti, S., & White, S. D. M. 2013, MNRAS, 430, 2232

Read, J. I. & Trentham, N. 2005, Royal Society of London Philosophical Transactions Series A, 363, 2693

Rey, M. 1999, Theoretical and observational cosmology (Kluwer Academic Publishers)

Richard, J., Smith, G., Kneib, J., Ellis, R., Sanderson, A., Pei, L., Targett, T., Sand, D., Swinbank, M., Dannerbauer, H., Mazzotta, P., Limousin, M., Egami, E., Jullo, E., Hamilton-Morris, V., & Moran, S. 2009, ArXiv e-prints

Richard, J., Smith, G. P., Kneib, J.-P., Ellis, R. S., Sanderson, A. J. R., Pei, L., Targett, T. A., Sand, D. J., Swinbank, A. M., Dannerbauer, H., Mazzotta, P., Limousin, M., Egami, E., Jullo, E., Hamilton-Morris, V., & Moran, S. M. 2010, MNRAS, 404, 325

Rott, C. 2012, ArXiv e-prints

Rykoff, E. S., Koester, B. P., Rozo, E., Annis, J., Evrard, A. E., Hansen, S. M., Hao, J., Johnston, D. E., McKay, T. A., & Wechsler, R. H. 2012, ApJ, 746, 178

Salvador-Solé, E., Serra, S., Manrique, A., & González-Casado, G. 2012, MNRAS, 424, 3129

Sand, D. J., Treu, T., Ellis, R. S., Smith, G. P., & Kneib, J. 2008, ApJ, 674, 711

Sanderson, A. J. R., Edge, A. C., & Smith, G. P. 2009, MNRAS, 398, 1698

Sanderson, A. J. R. & Ponman, T. J. 2010, MNRAS, 402, 65

Sanderson, A. J. R., Ponman, T. J., & O'Sullivan, E. 2006, MNRAS, 372, 1496

Santos, J. S., Fassbender, R., Nastasi, A., Böhringer, H., Rosati, P., Šuhada, R., Pierini, D., Nonino, M., Mühlegger, M., Quintana, H., Schwope, A. D., Lamer, G., de Hoon, A., & Strazzullo, V. 2011, A&A, 531, L15

Schneider, P., Kochanek, C., & Wambsganss, J. 2006, Gravitational Lensing: Strong, Weak and Micro: Saas-Fee Advanced Course 33 (Springer)

Shapiro, I. I. 1964, Physical Review Letters, 13, 789

Shu, C., Zhou, B., Bartelmann, M., Comerford, J. M., Huang, J.-S., & Mellier, Y. 2008, ApJ, 685, 70

Skrutskie, M. F., Cutri, R. M., Stiening, R., Weinberg, M. D., Schneider, S., Carpenter, J. M., Beichman, C., Capps, R., Chester, T., Elias, J., Huchra, J., Liebert, J., Lonsdale, C., Monet, D. G., Price, S., Seitzer, P., Jarrett, T., Kirkpatrick, J. D., Gizis, J. E., Howard, E., Evans, T., Fowler, J., Fullmer, L., Hurt, R., Light, R., Kopan, E. L., Marsh, K. A., McCallon, H. L., Tam, R., Van Dyk, S., & Wheelock, S. 2006, AJ, 131, 1163

Smith, G. P., Edge, A. C., Eke, V. R., Nichol, R. C., Smail, I., & Kneib, J. 2003, ApJL, 590, L79

Smith, G. P., Khosroshahi, H. G., Dariush, A., Sanderson, A. J. R., Ponman, T. J., Stott, J. P., Haines, C. P., Egami, E., & Stark, D. P. 2010, MNRAS, 409, 169

Smith, G. P., Kneib, J.-P., Smail, I., Mazzotta, P., Ebeling, H., & Czoske, O. 2005, MNRAS, 359, 417

Smith, G. P. & Taylor, J. E. 2008, ApJL, 682, L73

Soucail, G., Fort, B., Mellier, Y., & Picat, J. P. 1987a, A&A, 172, L14

Soucail, G., Mellier, Y., Fort, B., Mathez, G., & Hammer, F. 1987b, A&A, 184, L7

Spergel, D. N., Bean, R., Doré, O., Nolta, M. R., Bennett, C. L., Dunkley, J., Hinshaw, G., Jarosik, N., Komatsu, E., Page, L., Peiris, H. V., Verde, L., Halpern, M., Hill, R. S., Kogut, A., Limon, M., Meyer, S. S., Odegard, N., Tucker, G. S., Weiland, J. L., Wollack, E., & Wright, E. L. 2007, APJS, 170, 377

Spitler, L. R., Labbé, I., Glazebrook, K., Persson, S. E., Monson, A., Papovich, C., Tran, K.-V. H., Poole, G. B., Quadri, R., van Dokkum, P., Kelson, D. D., Kacprzak, G. G., McCarthy, P. J., Murphy, D., Straatman, C. M. S., & Tilvi, V. 2012, ApJL, 748, L21

Springel, V., White, S. D. M., Jenkins, A., Frenk, C. S., Yoshida, N., Gao, L., Navarro, J., Thacker, R., Croton, D., Helly, J., Peacock, J. A., Cole, S., Thomas, P., Couchman, H., Evrard, A., Colberg, J., & Pearce, F. 2005, Nature, 435, 629

Stadel, J., Potter, D., Moore, B., Diemand, J., Madau, P., Zemp, M., Kuhlen, M., & Quilis, V. 2009, MNRAS, 398, L21

Struble, M. F. & Rood, H. J. 1999, APJS, 125, 35

Sunyaev, R. A. & Zeldovich, Y. B. 1970, Comments on Astrophysics and Space Physics, 2, 66

Swinbank, A. M., Bower, R. G., Smith, G. P., Wilman, R. J., Smail, I., Ellis, R. S., Morris, S. L., & Kneib, J. 2007, MNRAS, 376, 479

Tsai, J. C. & Buote, D. A. 1996, MNRAS, 282, 77

Turner, E. L. 1990, ApJL, 365, L43

Vera-Ciro, C. A., Sales, L. V., Helmi, A., Frenk, C. S., Navarro, J. F., Springel, V., Vogelsberger, M., & White, S. D. M. 2011, MNRAS, 416, 1377

Vikhlinin, A. 1998, ApJL, 505, L123

Vikhlinin, A., Burenin, R., Forman, W. R., Jones, C., Hornstrup, A., Murray, S. S., & Quintana, H. 2007, in Heating versus Cooling in Galaxies and Clusters of Galaxies, ed. H. Böhringer, G. W. Pratt, A. Finoguenov, & P. Schuecker , 48

Voit, G. M. 2005, Reviews of Modern Physics, 77, 207

Walsh, D., Carswell, R. F., & Weymann, R. J. 1979, Nature, 279, 381

Wambsganss, J. 1998, Living Reviews in Relativity, 1

Weinberg, S. 1972, Gravitation and cosmology: principles and applications of the general theory of relativity (Wiley)

Xue, Y.-J. & Wu, X.-P. 2000, ApJ, 538, 65

Yahil, A. & Vidal, N. V. 1977, ApJ, 214, 347

Zitrin, A., Bartelmann, M., Umetsu, K., Oguri, M., & Broadhurst, T. 2012, MNRAS, 426, 2944

Zwicky, F. 1933, Helvetica Physica Acta, 6, 110

—. 1937, Phys. Rev., 51, 290