Title

Looking into the Cores of Galaxy Clusters and Dwarf Galaxies

Author

Andrade, Kevin Eugene

Publication Date

2021

UNIVERSITY OF CALIFORNIA,
IRVINE

Looking into the Cores of Galaxy Clusters and Dwarf Galaxies

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Physics

by

Kevin E. Andrade

Dissertation Committee:
Professor Manoj Kaplinghat, Chair
Professor James Bullock
Professor Michael Cooper

2021

DEDICATION

To my parents, who taught me to reach for the stars.

TABLE OF CONTENTS

Page

LIST OF FIGURES **vi**

LIST OF TABLES **x**

ACKNOWLEDGMENTS **xii**

CURRICULUM VITAE **xiii**

ABSTRACT OF THE DISSERTATION **xv**

1 Introduction **1**

2 Detecting Cores of Dark Matter Halos in Galaxy Clusters with Strong Lensing **5**
- 2.1 Introduction 5
- 2.2 Halo Models and Lensing Software 7
- 2.3 Mock Data Modeling 10
  - 2.3.1 Mock Data Preparation 12
  - 2.3.2 Mock Data Lens Model 15
  - 2.3.3 Mock Data Results 18
  - 2.3.4 Pseudo-Elliptical Approximation 20
  - 2.3.5 Model Dark Matter Halo Radial Density Profile 21
  - 2.3.6 Mock Data with Perturbers 24
  - 2.3.7 Summary of Mock Data Results 25
- 2.4 Modeling of Abell 611 25
  - 2.4.1 Abell 611 Data 25
  - 2.4.2 Abell 611 Lens Model 26
  - 2.4.3 Abell 611 Results 30
  - 2.4.4 Potential Systematic Errors 41
  - 2.4.5 Comparison to Other Works 42
- 2.5 Conclusions 44

3 A Stringent Upper Limit on Dark Matter Self-Interaction Cross Section from Cluster Strong Lensing **48**
- 3.1 Introduction 48

3.2 Strong Lens Modeling Methods 52
  3.2.1 Lens Profiles 55
  3.2.2 Photometric Measurement 56
  3.2.3 Central DM Halo 56
  3.2.4 BCG and Luminous Contribution of Member Galaxies 57
  3.2.5 DM Subhalos 58
  3.2.6 Use of Spectroscopic Redshift Data 60
  3.2.7 Use of Central Images 60
3.3 Models of Simulated Clusters 63
  3.3.1 Selection of Simulated Cluster Halos 63
  3.3.2 Simulated Image Production 64
  3.3.3 Model Details for Simulated Clusters 65
  3.3.4 Discussion 68
  3.3.5 Summary of Key Findings from Simulated Clusters 70
3.4 Models of Observed Clusters 72
  3.4.1 Abell 611 73
  3.4.2 Abell 2537 73
  3.4.3 RXC J2248.7-4431 74
  3.4.4 MS 2137.3-2353 74
  3.4.5 Abell 383 75
  3.4.6 Abell 2261 75
  3.4.7 MACS 2129.4-0741 76
  3.4.8 MACS 1720.3+3536 76
3.5 Strong Lensing Results 77
  3.5.1 Abell 611 79
  3.5.2 Abell 2537 81
  3.5.3 RXC J2248.7-4431 82
  3.5.4 MS 2137.3-2353 82
  3.5.5 Abell 383 83
  3.5.6 Abell 2261 83
  3.5.7 MACS 2129.4-0741 84
  3.5.8 MACS 1720.3+3536 84
  3.5.9 Halo and BCG masses 85
  3.5.10 Halo Concentrations 86
  3.5.11 Strong Lensing Systematic Errors 87
  3.5.12 BCG Offsets from DM Halo Centers 88
3.6 SIDM Halo Analysis 89
  3.6.1 SIDM Halo Model 89
  3.6.2 SIDM Halo Matching Inference Results 96
  3.6.3 Constraints on the Cross Section 104
3.7 Conclusions 107

4 Core Densities of Classical Milky Way Dwarf Galaxies: a Distribution Function Approach 111

4.1 Introduction . . . 111

4.2 Distribution Function Models . . . 113

4.2.1 DM Potentials . . . 116

4.2.2 Stellar DF Form . . . 118

4.2.3 Approximate Likelihood Function . . . 119

4.2.4 Derived Parameters . . . 120

4.3 Mock Data Modeling . . . 121

4.3.1 Mock Data Characteristics . . . 122

4.3.2 Mock Data Modeling Results . . . 123

4.3.3 Summary of Model Performance with Mock Data . . . 130

4.4 Bright Dwarf Spheroidal Models . . . 132

4.5 Conclusions . . . 139

Bibliography 142

A Appendices for Chapter 2 167

A.1 Relevant lensing formulas for cored halo models . . . 167

A.1.1 cNFW halo model . . . 167

A.1.2 Corecusp halo model . . . 169

A.2 Tables of Model Parameters . . . 173

A.3 Posterior Distributions for Mock Data and Abell 611 . . . 175

B Appendix for Chapter 3 180

B.1 Relevant lensing formulas for the cNFW halo model . . . 180

B.2 Stellar Mass–Halo Mass Relation . . . 182

B.3 Concentration Bias in Triaxial Halos . . . 182

B.4 Strong Lensing Error Residuals . . . 184

B.5 Surface Density Match between Strong Lensing and SIDM Models . . . 185

C Appendices for Chapter 4 187

C.1 Virial Shape Parameter . . . 187

C.2 Mock Data Characteristics . . . 189

C.3 Full Likelihood Function . . . 191

C.4 Core Parameter, Core Size, and Anisotropy Parameter Posteriors of the Observed Sample . . . 192

LIST OF FIGURES

Page

2.1 The location of the mock data image points, with representative caustic curves and critical curves. The representative curves shown correspond to a reference redshift of $z = 1.49$. [*Top Row*] Source plane plots for the cases that exclude central images. Cored is left, cuspy is right. [*Bottom Row*] Image plane plots for the base cored case (left) and base cuspy case (right). Central images are shown in cyan. 12

2.2 Example of a final fit image for the mock data. This fit is for the cuspy halo, without central images. The data points are shown in red, and the modeled images in cyan. The points appear as purple when the best fit model and data images overlay. The unmatched model images near (0, 0) are positive-parity “central" images, which are typically unobservable due to low magnification and/or obscuration by bright objects in the center of the cluster. 16

2.3 Core radius posterior distributions for the cored cases (upper) and cuspy cases (lower). The true value for core radius is 43.3 kpc ($10''$) in the cored case and zero in the cuspy case. Note the different horizontal scales. 17

2.4 Plots of scaled surface density (κ) versus radius for the mock data models. The use of the central images in the fitting enables a tighter constraint to the mass density in the inner region. The median (50th percentile) posterior value of the parameter set is shown as a solid red line, and the 16- to 84-percentile band is shown in gray. The true parameter value is shown as a dashed blue line. The radii bands in which images are located are highlighted in red. A reference redshift of $z = 1.49$ is used in the calculation of κ. The plots are slices that are averaged over 360°. The BCG can be observed as the bump at a radius of approximately $1''$. 19

2.5 Posterior distributions for halo mass, concentration and core radius parameters, showing the effect of using an approximate elliptical potential for highly elliptical systems. [*Top Row:*] full elliptical density profile used. [*Bottom Row:*] pseudo-elliptical approximation used. [*Left Column:*] cored case. [*Right Column:*] cuspy case. Orange lines and 'x' markers indicate the true parameter values. The units for m_{vir} are $10^{15} M_{\odot}$, and for core radius, arcseconds. 22

2.6 Scaled surface density (κ) versus radius for a fit of a cNFW halo lens on mock data generated with a Corecusp profile. The true density profile is shown as a dotted blue line. [*Top*]: Cored case, no central images. [*Bottom*]: Cuspy case, no central images. Note that the accuracy of the modeled profile declines at radii far from the image locations. 23
2.7 Image plane representation (left, with critical curves shown for z=1.49), and source plane representation (right, with caustic curves shown for z=1.49) of the best fit result for the Abell 611 cNFW model. The data points are shown in red, the modeled images in cyan, and show purple where they overlap. . . 30
2.8 Scaled surface density (κ) versus radius for Abell 611. cNFW and Corecusp models are shown, each subdivided into large-core and small-core solutions. . 32
2.9 Mass-to-light mismatch parameter α_{Salp} versus IMF slope Γ, where α_{salp} is defined as the ratio of M_*/L_V produced by this IMF over the value expected from SPS models (found in Newman et al. 2013) using a Salpeter IMF. Constraints from galaxy surveys in Cappellari+ 2013, Conroy+ 2017 and Leier+ 2016 are overlaid. The right panel shows the posterior probability densities in α_{salp} for the small-core and large-core solutions. 35
2.10 Velocity dispersion of the BCG, and corresponding fit to the observations of Newman et al. [133], assuming $\beta = 0$. 41

3.1 Schematic diagram of the analysis pipeline. Data sources are at the top. The final results are the parameter posteriors shown at the bottom. 53
3.2 Image plane plots for each cluster showing the data image positions (red) and modeling image positions (cyan). The points appear purple where they overlap. Critical curves are shown for a source redshift of 2.0. Redshift values for each family of points are indicated in the box to the right of each plot. North is up, East is to the left. Axes scales are in arc seconds. 61
3.2 , continued. 62
3.3 An example image of a simulated cluster (Sim. Cluster 3). The BCG is centered, and the red markers indicate 41 simulated point images from 16 source points. 66
3.4 Magnification maps for the simulated clusters. Regions of high magnification appear as yellow bands. The noise in the bands is due to the finite particle size of the simulation. The critical curves found from strong lensing are overlaid as cyan lines. Model image locations are shown as white circles for images matched to data, and red squares for extra images. Axes scales are in arc seconds. 69
3.5 Histograms for 2-dimensional BCG offset, BCG mass-to-light ratio and DM halo core radius posteriors. The histograms are normalized to unity area, and have been smoothed slightly with a Gaussian kernel for display purposes. . . 79
3.6 Surface density inferred from strong lensing ("data", in color) and the corresponding median SIDM modeled surface density ("model", in black) versus radius for each real cluster. The area shown for each is the 68% confidence interval. The clusters are offset by one decade each for display purposes. . . 80

3.7 Histogram plot for the line of sight axis ratio "s". The prior is shown as a dashed black line and is based on Vega-Ferrero et al. [193] for s<0.5 and is flat for s>0.5. 92
3.8 Comparison of the characteristic radii r_1 to the BCG radii and the radii of the innermost strong lensing data points. The dotted lines indicate 1:1 equality. Note that $r_1 > r_{\rm BCG}$ and $r_1 > r_{\rm inner}$ in all cases except A611, which allows a solution mode with low cross section and an r_1 value extending down to ~ 25 kpc at the 1σ level. 95
3.9 Posterior histograms for $\log_{10}(\sigma/m)$. The simulated clusters are shown as dashed lines. 98
3.10 Posteriors for relative velocity $V_{\rm rel} = 4\sigma_0/\sqrt{\pi}$ (horizontal axes, in km/s) versus SIDM cross section (vertical axes, in cm^2/g) from the SIDM halo matching model. The smaller boxes are for the clusters individually, and the larger box at the bottom is the composite of all 10 clusters. The 2σ regions are shaded, and the 1σ regions are shown in a darker shade. 99
3.11 Posteriors from the SIDM model for $r\,\rho(r)$ (vertical axes, in $M_\odot$ kpc^{-2}) versus radius (horizontal axes, in kpc). The median posterior and 2σ band are shown in orange. The dashed blue line shows the median concentration-mass relation found from Diemer and Joyce [45] using the redshift and the average virial mass of the specific cluster. For the simulated clusters, the actual $r\,\rho(r)$ is shown, but note that this includes subhalo density and is not exactly comparable to the other lines. 100
3.12 Comparison of the BCG mass inferred (top panel) and $c_{\rm vir}$ (bottom panel) from strong lensing with that from the SIDM model. The error bars indicate the 68% confidence interval. The diagonal lines represent equality. 102
3.13 *Top:* Concentration versus inferred cross section. The error bars indicate the 68% confidence interval. The square markers are the simulated clusters. *Bottom:* DM halo surface density at a radius of 28 kpc versus inferred cross section. 102

4.1 Typical fits to surface density (top), velocity dispersion (bottom) and vsp (bottom, inset), in this case for mock data set 10 (ID bcbO_2349). The best fit DF solution is shown in green. 124
4.2 Posteriors for $r_{\rm max}$ vs. $V_{\rm max}$ for the mock data sets. The 68% and 95% levels are shown, with the 68% level in a darker color. The black "x" indicates the true value. Top row: color-coded by data set. Bottom row: The same data, color-coded by their embeddedness in the DM halo. Right column: Cored profile. Left Column: NFW profile. 125
4.3 Predicted and true values for the half-light radius for the 24 mock data sets, using the same colors as Figure 4.2. The predictions are determined from the DF (see Equation 4.2) by calculating and integrating the surface density curve, finding the radius that yields have the total value (see Equation 4.18). The error bars indicate the 68% confidence interval. The true value is taken to be the median radius of the stars in the given data set. 126

4.4 Predicted and true values for the mass within the half-light radius for the 24 mock data sets, color-coded by their embeddedness in the DM halo. The error bars indicate the 68% confidence interval. The diagonal line indicates equality. 127
4.5 Predicted and true values for DM density at 150 pc for the 24 mock data sets, using the same colors as in Figure 4.2 . The error bars indicate the 95% confidence interval. 128
4.6 Posterior histograms of $\log_{10}[r_{\rm core}/{\rm kpc}]$ for the 24 mock data sets. The 12 cored sets are shown in blue; their true value is -0.585 (corresponding to $r_{\rm core} = 1$ kpc) and is indicated by the black dotted line. The 12 NFW sets are shown orange; their true value is $-\infty$ (corresponding to $r_{\rm core} = 0$ kpc), although we limit the parameter to -2 in log space (corresponding to $r_{\rm core} = 0.01$ kpc.) . . 129
4.7 True and predicted values of the core radius $r_{\rm core}$, in kpc, and color-coded by anisotropy profile. The group on the left is have NFW profiles, while the group on the right have cored profiles. The mock data sets with positive, rising β (blue) show a bias towards smaller cores. 130
4.8 True and predicted values for $\beta(r_{1/2})$, color-coded by embeddedness. 131
4.9 Posterior inferences for $r_{\rm max}$ vs. $V_{\rm max}$ for the observed sample. The 68% and 95% levels are shown, with the 68% level in a darker color. 132
4.10 Posteriors for the half-light radius and mass within that radius for the observed sample. The 68% and 95% levels are shown, with the 68% level in a darker color. Isodensity contours are shown as dotted lines, with the density value indicated, in units of $10^7 M_\odot$ kpc^{-3}. 134
4.11 The posterior inference for $r_{1/2}/r_s$, which indicates the degree to which the stellar population is embedded in its DM halo, for the observed sample. The 68% confidence intervals are shown. 136
4.12 DM density at 150 pc (ρ_{150}) inferred from the DF fits for the bright MW dSphs vs orbital pericenter distance (r_p). The error bars indicate the 68% confidence interval. 137

LIST OF TABLES

Page

2.1 The parameters, their true values, prior ranges and prior types for the mock data set. For all sampled parameters, the prior distribution is uniform over the parameter range. 11
2.2 Mock data calculated parameters. 11
2.3 The parameters and prior ranges for the dark matter halo, BCG and seven perturbing galaxies in the Abell 611 Lens Model. All priors are uniform over their range except for the parameter $r_{c,kpc}$, for which a log prior was used. . . 28
2.4 Key median posterior parameter values for the Abell 611 models. The bounds of the 68% confidence interval are also shown. The cNFW and Corecusp models exhibit bimodal solutions for some parameters, allowing either a near-zero core or a large core of $14''$to $16''$. Parameters labeled "sampled" were allowed to vary during the sampling run, while those labeled "derived" were derived from sampled parameters. 31

3.1 Key cluster halo parameters of simulated lensing clusters from Illustris and fitted values from strong lensing models. All simulated clusters were given a redshift of 0.3. The Illustris parameters are from 3D fits using Colossus. We were not able to obtain a meaningful strong lensng fit with Cluster 1 due to its unrelaxed nature. The strong lensing fit m_{200} and concentration values are for the main DM halo only and do not include that in subhalos. RMS position error is for the x and y coordinates, combined in quadrature. 63
3.2 Summary of the observed cluster sample. 72
3.3 Strong lensing results summary. Where confidence intervals are given, the 14th- and 86th-percentile values are indicated. The parameter r_c is the cNFW core radius. The BCG luminosity is measured using the closest HST filter to the rest frame wavelength, so as to be approximately comparable to V-band. Measurement uncertainties in luminosity are <2%. The halo offset from the BCG is the inferred projected offset in the plane of the sky. 78
3.4 SIDM parameter posterior summary. The columns are: (1) cluster name, (2) $\log_{10}$ cross section per unit mass, median and 68% C.L., (3) $\log_{10}$ cross section per unit mass, mean ± standard deviation, (4) characteristic radius, (5) central density, (6) central velocity dispersion, (7) halo maximum circular velocity, (8) BCG mass, (9) halo virial mass, and (10) concentration. The indicated ranges are the 68% confidence intervals. 98

4.1 Parameter limits in the MCMC model. Units are kpc for r_s and r_c, and km/s for v_s. The units for w are the number of stars in the population. The other parameters are dimensionless. The parameters $\tilde{E_c}$ and $\tilde{\Phi}_{\text{lim}}$ are made dimensionless by dividing by Φ_∞ for the distribution being used, and $\tilde{L_\beta}$ is made dimensionless by dividing by $r_s\sqrt{\Phi_\infty}$. 121
4.2 The Dwarf Galaxy Sample. The 2D half-light radius is quoted in this table. References: (1) Simon [177], (2) McConnachie [109] (3) The NASA/IPAC Extragalactic Database (NED), (4) MW 1 model from Patel et al. [139] (5) Light MW model from Battaglia et al. [6]. 131
4.3 Comparison of findings for r_{max} vs. V_{max} with the Jeans analysis of Kaplinghat et al. [81] for the observed sample. The Jeans analysis considered NFW and cored isothermal profiles as separate cases. 133
4.4 Comparison of 2D projected half-light radius ($R_{1/2}$), in kpc. The "DF" column is the $r_{1/2}$ posterior result from the DF, converted to 2D projected $R_{1/2}$ by dividing by 1.33. The "Plummer Fit" is the result from the best fitting 2-parameter Plummer profile, as applied to the (one-dimensional) sphericalized surface density data. The rightmost column is the half-light radius reported by Muñoz et al. [122] for a Sersic profile fit to 2D surface density maps, sphericalized as described in the text. 134
4.5 Comparison of findings for DM density at 150 pc (ρ_{150}), in units of $10^7 M_\odot$ kpc^{-3}. The mean posterior value and the 1σ uncertainties are indicated. The references for comparison are (A) Read et al. [152], (B) Kaplinghat et al. [81], isothermal case, and (C) Kaplinghat et al. [81], NFW case. Note that CVn I was not studied in reference A. 137

ACKNOWLEDGMENTS

I would like to give heartfelt thanks my advisor, Manoj Kaplinghat, for his patient tutelage during my graduate studies at UC Irvine. I thank my committee members James Bullock, Michael Cooper and Asantha Cooray (advancement committee) for their time and guidance. I am grateful for the many discussions with fellow graduate students and postdocs as well, including Nick Timmons, Derek Wilson, Milad Pourrahmani, Hooshang Nayyeri, Hasitha Eranda, Quinn Minor, Andrew Pace and Mauro Valli. Thanks are due to my co-authors for their hard work in preparing these works: Manoj Kaplinghat, Quinn Minor, Anna Nierenberg, Mauro Valli, Jackson Fuson, M. Grant Roberts, Demao Kong and Sophia Gad-Nasr. Finally, I thank my wife and children for enduring my passionate pursuit of astrophysics.

CURRICULUM VITAE

Kevin E. Andrade

EDUCATION

Doctor of Philosophy in Physics **2021**
University of California, Irvine *Irvine, CA*

Master of Science in Physics **2018**
University of California, Irvine *Irvine, CA*

Master of Business Administration **1991**
University of California, Berkeley *Berkeley, CA*

Master of Science in Electrical Engineering **1988**
Stanford University *Stanford, CA*

Bachelor of Science in Electrical Engineering **1985**
California State University, Fresno *Fresno, CA*

RESEARCH EXPERIENCE

Graduate Research Assistant **2014–2021**
University of California, Irvine *Irvine, California*

TEACHING EXPERIENCE

Teaching Assistant **2017**
University of California, Irvine *Irvine, CA*

REFEREED JOURNAL PUBLICATIONS

Detecting Cores of Dark Matter Halos in Galaxy Clusters with Strong Lensing **2019**
Monthly Notices of the Royal Astronomical Society, 487, 2, 1905

A Stringent Upper Limit on Dark Matter Self-Interaction Cross Section from Cluster Strong Lensing **2021**
Accepted for Publication in Monthly Notices of the Royal Astronomical Society

Core Densities of Classical Milky Way Dwarf Galaxies: a DistributionFunction Approach
in prep., Monthly Notices of the Royal Astronomical Society

ABSTRACT OF THE DISSERTATION

Looking into the Cores of Galaxy Clusters and Dwarf Galaxies

By

Kevin E. Andrade

Doctor of Philosophy in Physics

University of California, Irvine, 2021

Professor Manoj Kaplinghat, Chair

Here are three studies of dark matter ("DM") dominated objects, with the aim of constraining the shapes of the inner regions of their DM halos. The first chapter is an introduction to the dissertation. Using galaxy cluster Abell 611 as a prototype, in Chapters 2 and 3 we model the strong lensing of eight galaxy clusters to infer the mass distribution in the range of 10 kpc to 150 kpc from the centers of the clusters, then subsequently derive constraints on self-interaction cross-section over mass σ/m of DM particles. We infer the mass profiles of the central DM halos, bright central galaxies, key member galaxies, and DM subhalos for the member galaxies for all 8 clusters using the `QLens` code. The inferred DM halo surface densities are fit to a self-interacting dark matter ("SIDM") model, which allows us to constrain the self-interaction cross section over mass σ/m. When our full method is applied to mock data generated from two clusters in the Illustris-TNG simulation, we find results consistent with no DM self-interactions as expected. For the eight observed clusters with average relative velocities of 1458^{+80}_{-81} km/s, we infer $\sigma/m = 0.082^{+0.027}_{-0.021}\mathrm{cm}^2/\mathrm{g}$ and $\sigma/m < 0.13$ cm^2/g at the 95% confidence level. In Chapter 4 we examine the bright dwarf spheroidal galaxies of the Milky Way, using a novel approach not yet completed in literature on these 9 objects: the use of phase space distribution functions. We use as data the observed surface density, velocity dispersion and fourth-order velocity moment. We use mock data from the Gaia Challenge project to show that the model can infer important characteristics such as $r_{\max}$, $v_{\max}$, the half-light radius, the density at 150 pc and the core radius. We find that the accuracy of the predictions is highest for data sets in which the stellar component is not too deeply embedded within the DM halo. For the observed sample, we infer these same parameters with accuracy comparable to those using a Jeans analysis approach. We confirm the wide diversity of inner densities in these objects, in particular that the Draco, Leo I, Leo II and Ursa Minor dwarf spheroidal galaxies are approximately four times more dense than Carina, Sextans and Fornax.

Chapter 1

Introduction

Dark Matter ("DM") is thought to make up a substantial fraction of the mass of the universe, but its scientific description is still vague. Zwicky [213] first used the phrase in studying the redshifts and velocity dispersions of galaxies, and concluded that "dunkle Materie" (dark matter) is present in much larger quantities than luminous matter. Today, the prevailing theory for the formation of the universe is ΛCDM, in which the Λ denotes Einstein's cosmological constant and CDM denotes cold dark matter. In ΛCDM, the mass budget of the universe is approximately 85% DM [1]. Scientists today are engaged in a great many searches for DM, some attempting to detect DM particles directly and others using indirect methods. But yet we still know little about this mysterious substance, especially its particle nature (and whether it is a particle at all).

We see evidence for its existence in many ways, including the velocity dispersion of galaxy clusters, the rotation of spiral galaxies, the shapes of galaxy cluster mergers, the cosmic microwave background, gravitational lenses, the filament structure of the universe and galaxy formation/evolution history [174]. ΛCDM is successful at explaining all of these, although some discrepancies remain [15]. There are two primary places where we try to examine DM:

laboratories on Earth, and astrophysical "laboratories" in the universe. Here I focus on the astrophysical approach. DM interacts with baryonic matter primarly (or perhaps even exclusively) through gravity. We can ask "Where should we look in the universe to best observe DM's gravitational effects?" Ideal places would have high concentrations of DM, but low concentrations of baryons, so that we can mostly ignore the baryons and not have to disentangle their effects from that of the DM. Still, we need *some* baryons, since baryons are what we ultimately observe.

Because DM shows its astrophysical existence primarily via gravity, we need it to be present in large quantities in order to study it. At present, that means scales of at least several orders of magnitude more than the mass of the sun. Behroozi et al. [7] and Read et al. [151] examine the DM fraction of at scales ranging from $10^8 M_\odot$ to $10^{15} M_\odot$, and conclude that the DM fraction is highest at either the bottom end of that range (i.e., dwarf galaxy scales) or the top end of the range (galaxy cluster scales). At those scales, we expect baryonic matter to comprise less than 0.1% of the total mass of astronomical objects. In the middle ground, containing galaxies such as our own Milky Way (approximately $10^{12} M_\odot$), the baryon fraction can be up to a few percent of the total. So here are two sets of astrophysical targets that are prime laboratories for studying DM: galaxy clusters and dwarf galaxies.

The Standard Model of particle physics does an excellent job at explaining most of the phenomena that we observe in the Universe, but it does not explain DM, so we need to look beyond the Standard Model. One of the main frameworks for doing so is the dark sector (or "hidden" sector) framework. This theory introduces new particles and force carriers beyond the Standard Model. A natural result of these new particles would be self-interaction, in which one DM particle occasionally collides with another. Self-interactions do not occur under the ΛCDM paradigm. Such self-interacting dark matter ("SIDM"), [182, 150, 195, 163, 80, 155] might explain some potential departures of ΛCDM from observation. Two of these potential discrepancies are (1) that the observed density profiles in the cores of DM halos is not as steep or "cuspy" as ΛCDM predicts [20], and (2) that there are fewer dwarf galaxies observed than are predicted by ΛCDM N-body simulations [15, 81].

In Chapter 2, we examine the galaxy cluster Abell 611. This cluster is a beautiful example of a gravitational lens, with distant source galaxies at various redshifts making multiple lensed images. We use the gravitational lens software `QLens` to model the DM and baryonic components of the cluster. We explore different functional forms for the DM density profile, and show that the model can put robust constraints on the DM halo core size and core density, and consequently can help constrain the self-interaction cross-section for SIDM.

In Chapter 3, we use the analysis procedure from Chapter 2 and expand its application to three simulated clusters from the Illustris-TNG simulation [128] as well as 8 observed clusters, which include Abell 611 and 7 similar ones. Models of the simulated clusters accurately recover the critical curves and image positions of the gravitational lens. More importantly, the models recover the density profiles of the simulated halos, including their cuspy inner profile. We apply the model to the 8 observed clusters to obtain their density profiles. We next use a halo-matching model that allows the DM halos to assume an inner profile consistent with self-interaction, which allows us to infer the self-interaction cross-section for each. Finally, using the SIDM cross-section results of all 8 clusters, we are able to constrain the self-interaction cross-section to be less than 0.13 cm^2/g at the 95% confidence level.

In Chapter 4 we shift to the dwarf galaxy scale to look into the cores of the bright dwarf spheroidal galaxies of the Milky Way. These galaxies have been analyzed before using Jeans analysis, but we take a different approach here. We use phase-space distribution functions to model 24 simulated dwarf galaxies based on those of the Gaia Challenge project [154] and the 9 bright dwarf spheroidal galaxies of the Milky Way, using as data their surface density, velocity dispersion and fourth-order velocity moment. Using this approach, we are able to infer key characteristics of the DM halos such as the halo radius scale, velocity scale, the mass within the half-light radius and the density at 150 pc from the galaxy centers. These inferences have accuracy comparable to that of the Jeans analysis method, and confirm that the DM halos of these Milky Way satellites are very diverse. These characteristics will be useful in further studies that explore small-scale puzzles in DM such as the ones mentioned above, and whether SIDM or other physics beyond ΛCDM can help explain them at the scale of dwarf galaxies.

Chapter 2

Detecting Cores of Dark Matter Halos in Galaxy Clusters with Strong Lensing

2.1 Introduction

Galaxy clusters provide a critical test of dark matter theories if their inner dark matter density profile can be measured. Hierarchical structure formation models make concrete predictions about cluster density profiles. For example, in the cold dark matter (CDM) paradigm, dark-matter-only simulations show that hierarchical structure formation leads to cuspy dark matter halos with a Navarro-Frenk-White (NFW) density profile [125, 126, 60], with the 3-D density profile $\rho \propto r^{-1}$ in the inner region. However, self-interactions or the feedback effects from baryons can potentially modify the inner slope. For example, Active Galactic Nuclei (AGN) may potentially cause flattening of central cusps of cluster mass halos [141, 106, 110]. Self-interacting dark matter (SIDM) models [182] and simulations thereof [179, 51, 163] predict shallower slopes for radial dark matter density profiles.

The presence or absence of dark matter cores in clusters is an open question. Sand et al. [168]

found an inner logarithmic slope of ~ -0.5 in two relaxed clusters using a combination of lensing and kinematic data. Newman et al. [133, 132] found a similar result for Abell 611 and similar clusters using a combination of strong lensing, weak lensing and stellar kinematics. Del Popolo [40, 41] came to the same conclusion analyzing a group of clusters that include Abell 611. Annunziatella et al. [4] dissect the cluster MACS J0416 and find that the inner dark matter profile is shallower than NFW. In contrast, Caminha et al. [23] finds an inner log slope of the density profile close to the canonical NFW value of -1 for the cluster MACS J1206, although they do not separate out the baryonic component. While simulations have been steadily advancing in scope and resolution, there is still no consensus in the question of cores in clusters (see Schaller et al. 170, Martizzi et al. 106).

The mass profiles of galaxy clusters can be probed by several methods, each having a distinct range of radii at which it can yield insight. These methods include stellar kinematics, strong lensing, X-ray emission and weak lensing, which cover the full range from 10 kpc to 1 Mpc scales [190, 130, 74].

Strong lensing refers to multiple images of a background source, with the image positions and magnifications determined by the mass distribution of the deflector. Multiple images can be exploited to provide strong constraints on the distribution of the matter in the lens, since the mass distribution must simultaneously satisfy the lens equations for all images [e.g., 88]. Images appear near the Einstein radius of the lens, which is typically tens of kpc for galaxy cluster lenses, well within the cluster scale radius, and also where the effects of baryons and dark matter self-interactions are strongest. Separating the two effects is critical.

Abell 611 is dynamically relaxed [47]. It has been studied before, but usually in conjunction with X-ray images, weak lensing and/or kinematic data (see Donnarumma et al. [47], Schmidt and Allen [172], Romano et al. [164], Richard et al. [156], Newman et al. [130, 133, 132], Del Popolo [40, 41], Monna et al. [118], Zitrin et al. [212]). Abell 611 is an excellent test case for gravitational lensing analysis, as it images multiple sources at multiple redshifts, contains radial arcs at various locations, and has images at a range of radii, from 30 kpc to more than 100 kpc.

In this Chapter and the following one we constrain the dark matter distribution using strong lensing alone. This allows us to characterize the strengths and limitations of strong lensing separately from other techniques. Our primary goal is to determine how well one can constrain the size of a central core in a lensing cluster, i.e., determine the radius (if any) below which the density profile becomes relatively flat.

We adopt a flat cosmology with $\Omega_\Lambda = 0.7$, $\Omega_m = 0.3$, and $H_0 = 70\, km\ s^{-1} Mpc^{-1}$. At the adopted lens redshift of 0.288, the distance to the lens is 893.8 Mpc, and $1''$is equal to 4.329 kpc. We define halo mass as M_{200}, the mass enclosed by a sphere of radius R_{200}, which we define in turn as the radius at which the halo density is 200 times the critical density of the universe at the redshift of the halo.

This Chapter is organized as follows. In Section 2.2 we discuss the new lensing code and lens mass profiles. In Section 2.3 we discuss the mock data sets and lens models used in our analysis, and the results of those models. In Section 2.4 we describe the data and lens model for Abell 611, and discuss the results of the analysis for that cluster. Our conclusions are summarized in Section 2.5.

2.2 Halo Models and Lensing Software

We choose a flexible mass model for our tests which has a core, and for which the mass distribution approaches NFW as the core radius goes to zero. We require a fast method to calculate the magnification and deflection at each point. These requirements are in a new lensing software, `QLens` , which allows for both pixel image modeling (using pixelated source reconstruction) and point image modeling (with option to include fluxes, time delays, and multiple sources at different redshifts). `QLens` includes 14 different analytic lens models to use for model fitting, including 10 different density profiles where ellipticity can be introduced into either the projected density or the lensing potential (the remaining models include a general multipole term in either the kappa or the potential, point mass, external shear or mass sheet). In addition, a built-in nested sampler is included, along with an adaptive MCMC algorithm called T-Walk [30]; however the code can also be compiled with MultiNest [56] or PolyChord [65]. The `QLens` package is now available on GitHub by request and includes a student-friendly tutorial to get users started with point image modeling. Here, we describe the novel features implemented in `QLens` that are critical to the results of this paper.

We consider two types of cored halo models for modeling cluster halos. The first model is a cored NFW profile (cNFW), for which the density profile is defined as

$$\rho = \frac{\rho_s r_s^3}{(r_c + r)\,(r_s + r)^2}. \tag{2.1}$$

This will be the primary lens model we use in this work for the cluster halo. Note that as the core radius $r_c \to 0$, the density profile reduces to the standard NFW form. This profile has been used in lens modeling in Newman et al. [133, 132], and was found by [143] to provide a reasonably good fit to cored DM halos found in hydrodynamical simulations of [64]. Analytic formulas for the kappa profile and deflection of the corresponding spherical model are given in Appendix A.1.

The second cored halo model, which we call the *Corecusp* model, is defined as

$$\rho = \frac{\rho_0 r_s^n}{(r^2 + r_c^2)^{\gamma/2}\,(r^2 + r_s^2)^{(n-\gamma)/2}}. \tag{2.2}$$

This is an extension of the "cuspy halo model" of [121]. This model also allows for a core in addition to a scale radius, where $r_c < r_s$. In the limit of large r, the log-slope is given by n, whereas in the limit of small r and zero core, the log-slope is given by γ. Note that in this model, both the core and scale radius are added in quadrature to r, resulting in a more rapid turnover compared to the cNFW model at the scale and core radii. As a result, the profile does not reduce exactly to NFW in the limit $r_c \to 0$. On the other hand, the greater flexibility afforded by the variable inner and outer log-slopes may become useful when combining strong lensing with data that probe the density profile on larger scales, e.g. weak lensing or X-ray data. For the purposes of this paper, however, this profile provides a comparison of how sensitive the core constraints are to the exact nature of the turnover behavior of the density profile near r_c. As we are primarily interested in the behavior in the region interior to r_s, we fix n to 3 in this work, to match that of an NFW profile. The relevant lensing formulas are given in Appendix A.1.

For each of the above lens models, we add ellipticity by making the replacement $R^2 \to qx^2 + y^2/q$ in the projected density profile. The deflection and Hessian of the lens mapping must be calculated by numerical integration (see Schramm 173, Keeton 85), which is computationally expensive. While the integrals can be done relatively quickly using Gaussian quadrature, it is not known *a priori* how many points will be required for the integral to converge beyond a specified tolerance. This can be solved by an adaptive quadrature scheme where the integration is done at successively higher orders and an error estimate is obtained after each iteration, then stopping when the error falls below a specified tolerance. To implement this, we employ a modification of Gaussian quadrature known as Gauss-Patterson quadrature [38], which consists of nested quadrature rules whereby a given order of integration retains the function evaluations from the lower orders, thus ensuring they are not wasted (at the cost of allowing up to a maximum order of 511 points).[1] For lensing calculations, we find this

---

[1] The algorithm described above is quite similar to adaptive Clenshaw-Curtis quadrature [38] except it is an open interval quadrature rule, thus dodging the issue of having to evaluate the projected density or its derivative at $r = 0$.

adaptive quadrature scheme requires nearly an order of magnitude fewer function evaluations compared to Romberg integration [38] for a tolerance $\sim 10^{-3}$. This reduces the expense of lensing calculations enormously for elliptical projected density profiles.

The mass distribution of the dark matter halo in strong lenses has been shown to be consistent with elliptical isodensity contours in several studies [205, 206, 90]. Hence, when modeling the cluster halo, there is strong motivation for introducing ellipticity into the projected density profile as described above. However, because of the computational burden of performing the integrations for elliptical density profiles, it is common to instead use the "pseudo-elliptical" model in which the halo ellipticity is incorporated into the gravitational potential rather than the projected density [62]. Here, we consider both approaches, and will compare the pseudo-elliptical approximation to the full elliptical density approach.

2.3 Mock Data Modeling

We are interested in determining the capabilities of strong lens modeling for inferring the cluster dark matter halo properties. The following questions guided our choice of mock data sets.

1. Is it possible to distinguish between a core and a cusp with strong lensing alone?
2. To what extent can central images help in determining the inner density profile?
3. How do inaccuracies in the outer density profile affect the result? Do they lead to a spurious detection of a core or cusp?
4. Does the use of an elliptical *potential* rather than using a elliptical *density* profile lead to inaccurate results?

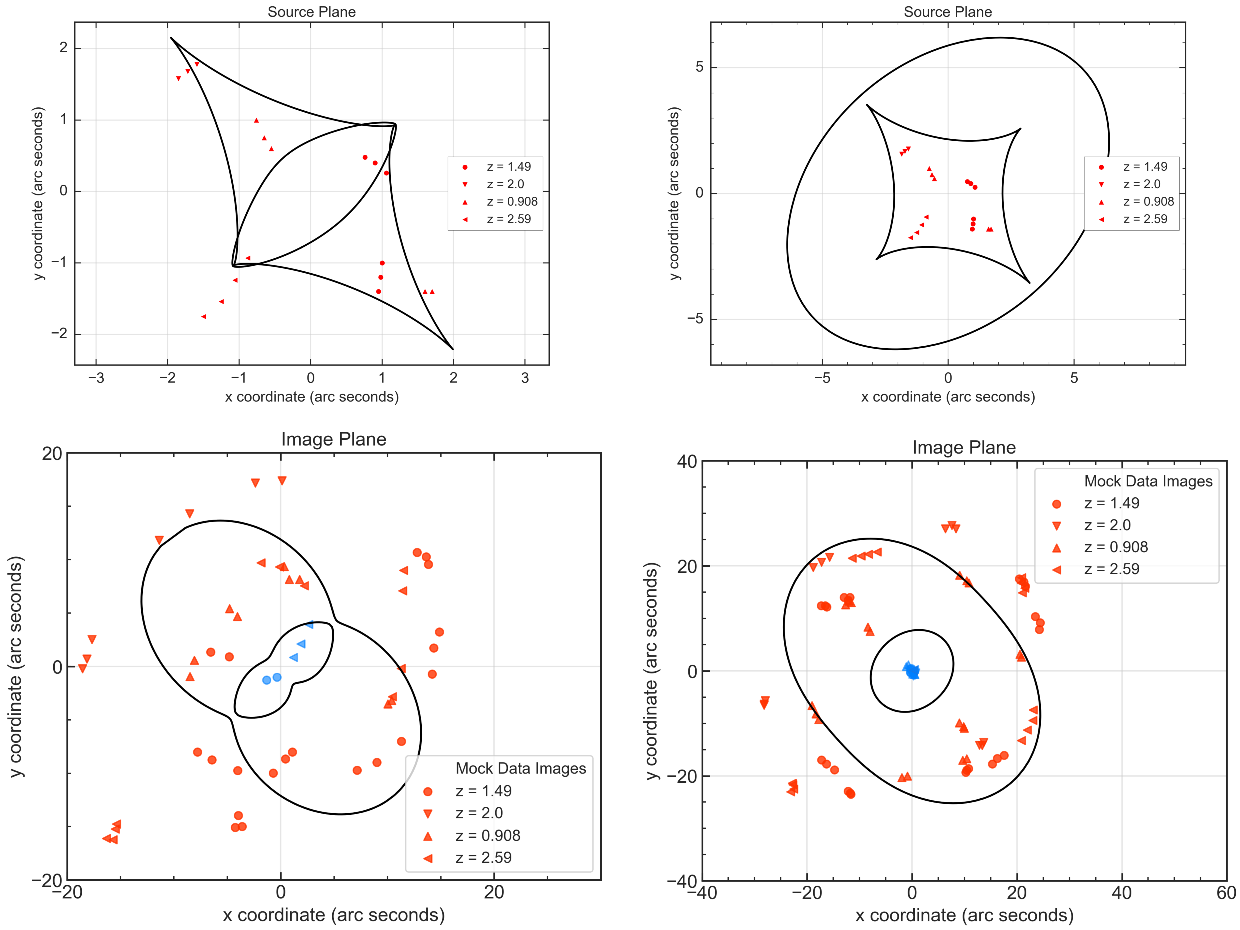
Table 2.1: The parameters, their true values, prior ranges and prior types for the mock data set. For all sampled parameters, the prior distribution is uniform over the parameter range.

Parameter Name	Description	Units	True Value	Prior Range
DM Halo (cNFW profile)				
M_{200}	halo mass	$M_\odot$	1.1×10^{15}	$2 \times 10^{14} - 5 \times 10^{15}$
c_{200}	concentration	-	7.0	1 - 20
β_c	core ratio (r_c/r_s)	-	0.157	0.0 - 0.96
q	axis ratio	-	0.8	0.3 - 1.0
θ	orientation	degrees	132.5	120 - 150
x-center	x coordinate of center	$''$	0.	-5 - 5
y-center	y coordinate of center	$''$	0.	-5 - 5
BCG (dPIE profile)				
b	mass parameter	$''$	0.60	0.1 - 10.0
a	scale radius	$''$	15	(fixed)
s	core radius	$''$	0.05	(fixed)
q	axis ratio	-	0.75	(fixed)
θ	orientation	degrees	72.5	(fixed)
x-center	x coordinate of center	$''$	0.5	(fixed)
y-center	y coordinate of center	$''$	-0.9	(fixed)

Table 2.2: Mock data calculated parameters.

Parameter Name	Description	Cored System	Cuspy System
D_{lens}	angular dia. dist. to lens	894 Mpc	894 Mpc
$\Sigma_{crit}(z_{src,ref} = 1.49)$	critical surface density	$2.53 \times 10^9 M_\odot\ /kpc^2$	$2.53 \times 10^9 M_\odot\ /kpc^2$
DM Halo (cNFW profile)			
r_s	scale radius	$63''.7$	$63''.7$
r_c	core radius	$10''$	$0''$
r_{200}	halo radius	$445''.7$	$445''.7$
BCG (dPIE profile)			
$r_{E,BCG}$	Einstein radius	$0''.57$	$0''.57$
M_{BCG}	total mass	$1.34 \times 10^{12} M_\odot$	$1.34 \times 10^{12} M_\odot$

Figure 2.1: The location of the mock data image points, with representative caustic curves and critical curves. The representative curves shown correspond to a reference redshift of $z = 1.49$. [*Top Row*] Source plane plots for the cases that exclude central images. Cored is left, cuspy is right. [*Bottom Row*] Image plane plots for the base cored case (left) and base cuspy case (right). Central images are shown in cyan.



Modeling the mock data allows us to test the power of models to constrain relevant parameters in a controlled way.

2.3.1 Mock Data Preparation

To test the ability of the lens models and software to constrain the relevant halo and Brightest Cluster Galaxy ('BCG') parameters, six sets of mock data were created, as follows:

- Cuspy, no central images
- Cored, no central images
- Cuspy, with central images
- Cored, with central images
- Cored, no central images, highly elliptical halo (axis ratio = 0.5)
- Cuspy, no central images, highly elliptical halo (axis ratio = 0.5)

The mock data sets were constructed to be similar in nature to Abell 611 in most respects, including mass, redshift, position angle, offset and ellipticity. To examine the usefulness of central images in constraining system parameters, two image sets were generated for each of the cored and cuspy cases; one image set included positive parity central images and one did not. The input parameters of the mock data objects are summarized in Table 2.1, and calculated parameters are shown in Table 2.2.

The mock data sets consisted of a dark matter halo and a bright central galaxy, offset by $\sim 1''$, each at a redshift of 0.288, matching the redshift and inferred offset of Abell 611 [D11]. The dark matter halo was modeled by a cNFW profile. The scale radius ("r_s") was chosen to be $50''$and the halo mass to be $1.1 \times 10^{15} M_\odot$, similar in magnitude to Abell 611 and other galaxy clusters. The dark matter halo is oriented 132.5° counterclockwise from the x-axis. In the cored cases only, a uniform density core is modeled with a transition radius of $10''$.

We chose an axis ratio of 0.8 for primary mock data sets, as that is similar to that of typical clusters. But since galaxy clusters can sometimes have highly elliptical structure (see Richard et al. [156], Table 7), we constructed two separate mock data sets with a highly elliptical dark matter halo (axis ratio = 0.5) to investigate the effects of more severe ellipticity on the model inferences, which we discuss in Section 2.3.4.

The BCG was modeled with a dual pseudo-isothermal ellipsoid (dPIE) profile (Elíasdóttir et al. [52], and defined in Appendix A.1.2) that allows separate specification of the tidal break radius and core radius. The `QLens` mass parameter for dPIE profiles,"b", can be expressed as

$$b = \frac{\sigma_0^2 r_{cut}}{2G\Sigma_{crit}(r_{cut} - r_c)} \tag{2.3}$$

where σ_0 is the central velocity dispersion, G is the gravitational constant, Σ_{crit} is the critical surface density of the lens at the relevant redshift, r_{cut} is the tidal cutoff radius and r_c is the core radius. Therefore, the mass parameter "b" roughly corresponds to the Einstein radius (and reduces exactly to the Einstein radius as $r_c \to 0$), and $b \propto \sigma_0^2$. The BCG was given a stellar mass of $1.34 \times 10^{12} M_\odot$, which corresponds to $b = 0.60$, and a small core of uniform density, with a radius of $0.05''$, similar to that noted in Abell 611 [D11].

The tidal break radius of the BCG in Abell 611 found in recent literature is between 10 and $20''$(Newman et al. [130] and [D11]); a value of $15''$was chosen for this mock data analysis. The axis ratio was chosen at 0.75, as might be typical for an elliptical galaxy, and close to that measured for the BCG in Abell 611 in other strong lensing analyses. The BCG was oriented at an angle of 72.5° and was positioned at ($0\rlap{.}''5$, -$0\rlap{.}''9$) in the image plane. This corresponds to an offset from the dark matter halo by $\sim$ 4 kpc, consistent with the offset found for Abell 611 in Donnarumma et al. [47] and Newman et al. [130], and similar to that of other clusters [133].

Sources for the mock data were chosen in four redshift groups: 0.908, 2.00, 2.06 and 2.59, respectively, similar to source redshifts Abell 611 [D11]. For each redshift, one or two compact sources were created within an area of approximately $1''$, with two to three source points each. Simulated position errors with a standard deviation of $0\rlap{.}''2$ were incorporated into the image positions. Figure 2.1 shows the source plane and image plane representations for the cored and cuspy data sets.

2.3.2 Mock Data Lens Model

A lens system model was constructed with two lens objects, one to represent the halo and one to represent the BCG. As in the mock data preparation, the halo lens was modeled with a cNFW profile and the BCG lens with a dPIE profile.

The systems were analyzed using a nested sampling algorithm with 1,000 live points. There are 8 sampled parameters (7 for the dark matter halo and one for the BCG). For purposes of computational speed, source plane χ^2 evaluations were used in the beginning of each run, with a switch to image plane evaluations occurring mid-run, producing the image-plane result but in less time. Spot checks of the posterior distributions were made to verify that the results are nearly identical to employing the image-plane χ^2 for the entire run.

The image plane χ^2 is calculated as follows:

$$\chi^2_{img} = \sum_i \frac{\left(\mathbf{x}_{obs,i} - \mathbf{x}_{mod,i}\right)^2}{\sigma_i^2} \tag{2.4}$$

where i is the image index, σ_i are the image position uncertainties, $\mathbf{x}_{obs,i}$ is the observed image position and $\mathbf{x}_{mod,i}$ is the modeled image position. A similar χ^2 can be calculated in the source plane, as described in Keeton [84].

We used uniform priors for the parameters as shown in Table 2.1. All of the models showed a close match between the data images and the model images. An example final fit image is shown in Figure 2.2.

Figure 2.2: Example of a final fit image for the mock data. This fit is for the cuspy halo, without central images. The data points are shown in red, and the modeled images in cyan. The points appear as purple when the best fit model and data images overlay. The unmatched model images near (0, 0) are positive-parity "central" images, which are typically unobservable due to low magnification and/or obscuration by bright objects in the center of the cluster.

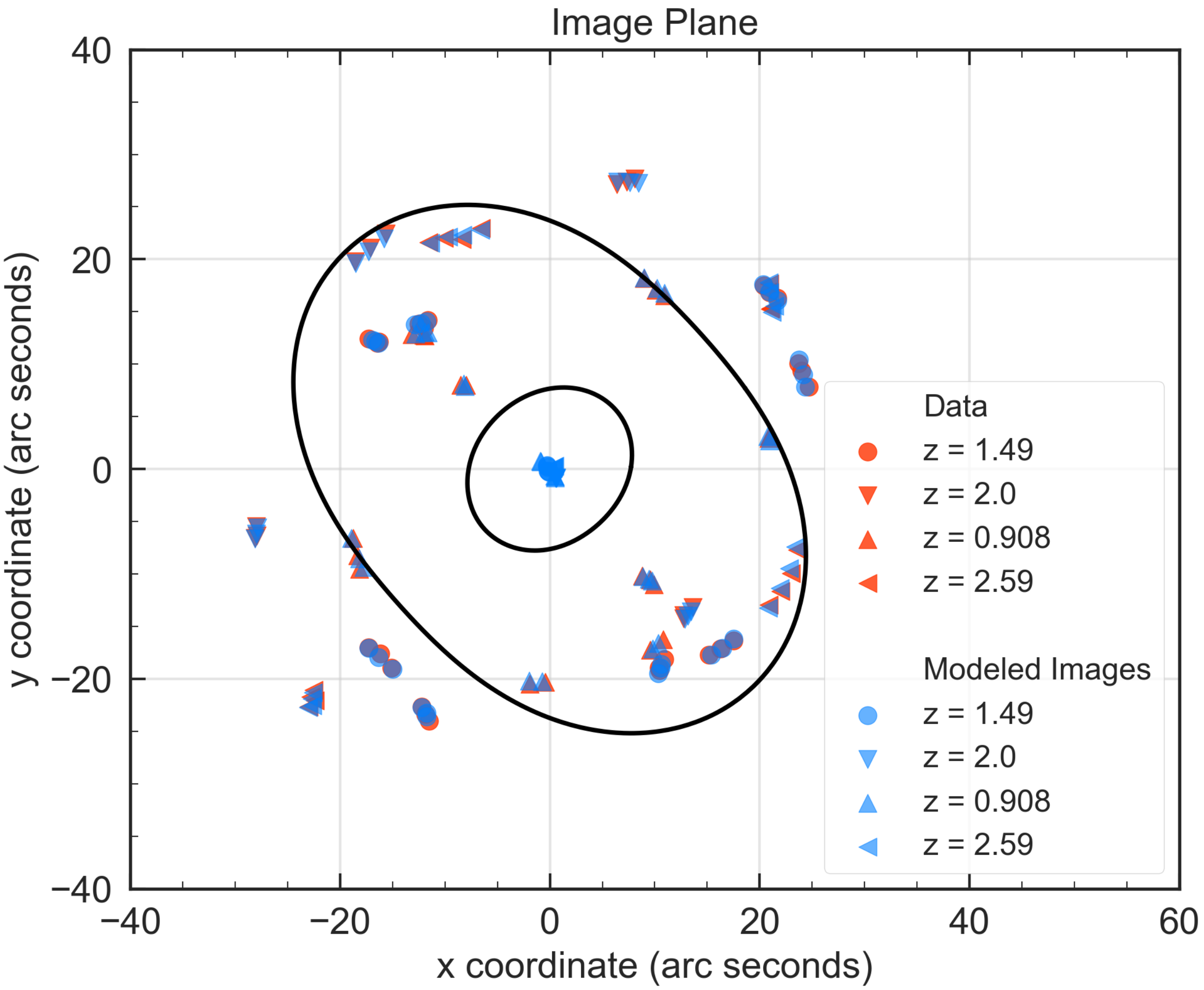
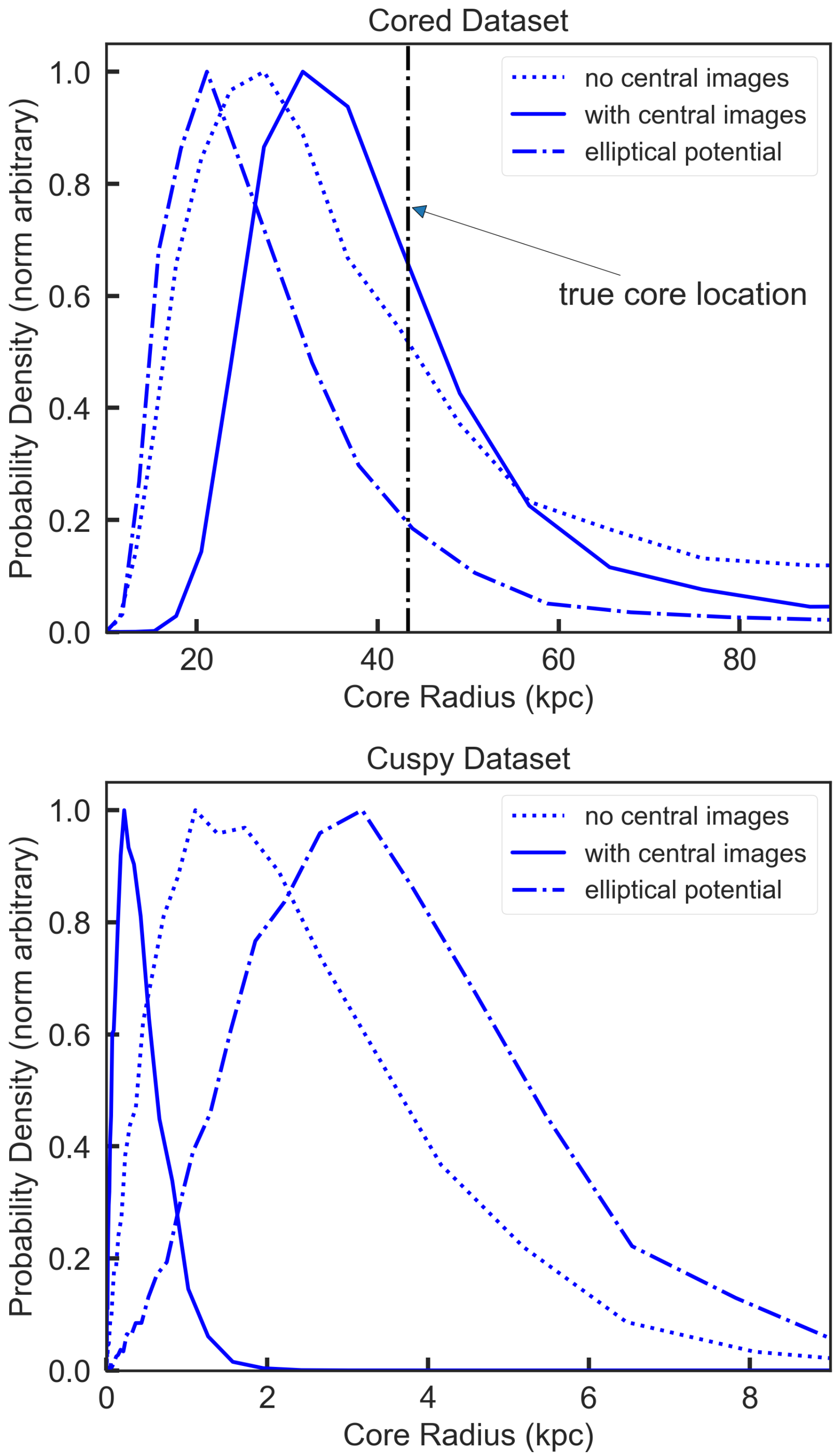


Figure 2.3: Core radius posterior distributions for the cored cases (upper) and cuspy cases (lower). The true value for core radius is 43.3 kpc (10″) in the cored case and zero in the cuspy case. Note the different horizontal scales.



2.3.3 Mock Data Results

2.3.3.1 Data without Central Images

We first consider as our baseline a dataset with no central images included, as might be expected for a cluster system with a bright object near the center that would obscure such central images. Figure 2.3 shows the posterior probability profiles for the cNFW "r_c" parameter (core break radius) for the six cases. For the cored data set with no central images (dotted curve), the median value of core radius (true value of 10″.0) is 7″.0, with a 1-σ lower bound of 4″.8. For the cuspy data set (true core radius of zero), the median fit value is 0″.3, with a 1-σ upper bound of 0″.6. The model is able to accurately distinguish between the cored and cuspy cases.

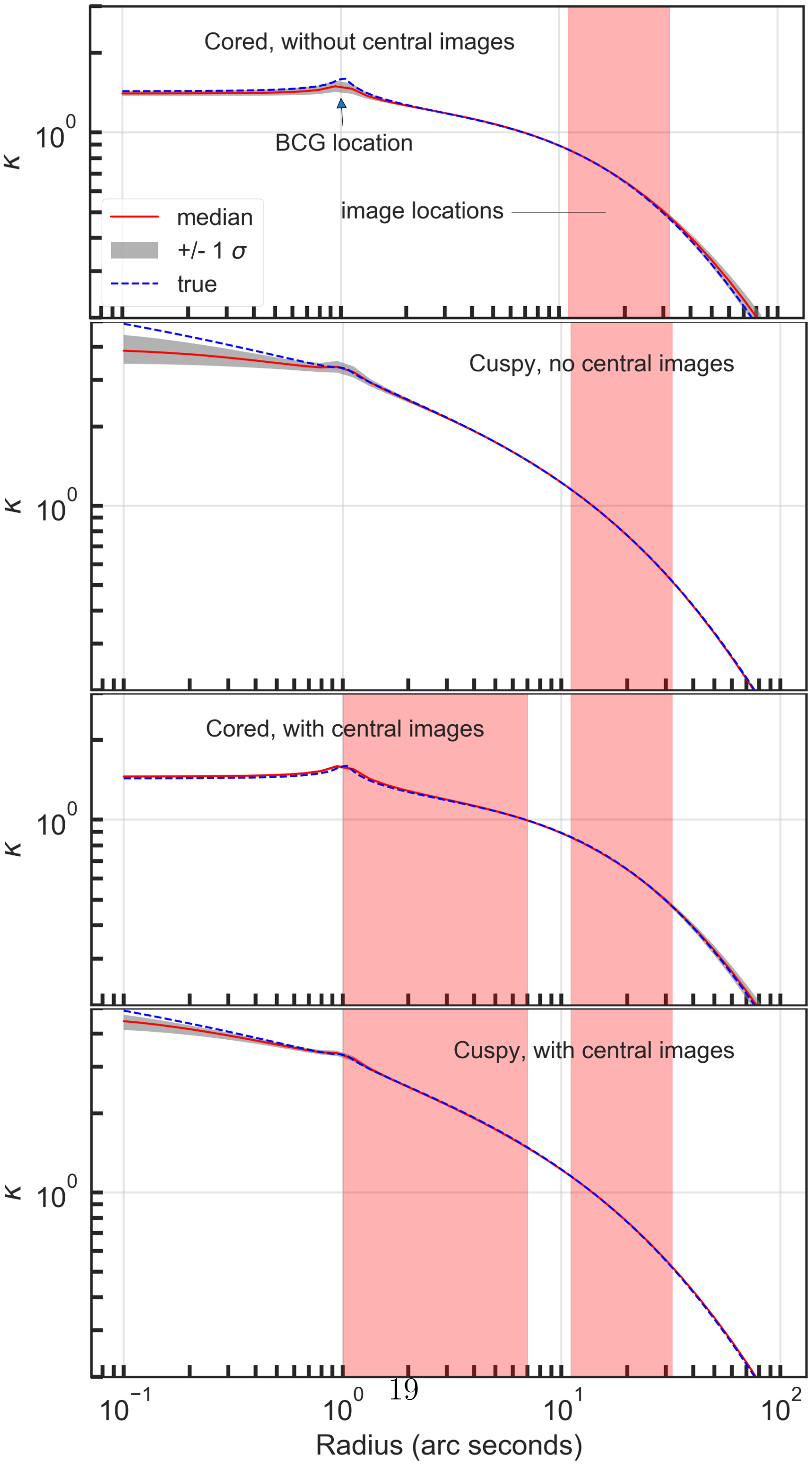
Triangle plots showing the posterior distributions of the parameters can be found in Figures A.1 and A.2 in Appendix A.3. The parameters are successfully recovered, with all of the true values of the parameters within ~ 1 standard deviation of the best-fit posterior value.

2.3.3.2 Data with Central Images

It is interesting to look at the same analysis with the central images included to see how useful the central images are in constraining parameter values. The solid curves of Figure 2.3 provide an illustration of this. For the cored data set, the best fit value of core radius is 8″.5, with a 1-σ lower bound of 6″.5. For the cuspy data set, the best fit value is 0″.05, with a 1-σ upper bound of 0″.12. Clearly, the central images greatly enhance the ability of the model to accurately constrain the core radius. The other parameters follow a similar pattern, with the mass and scale radius parameters of the dark matter halo determined with some uncertainty (but with more certainty than in the cases without central images), while the axis ratio, position angle and centroid coordinates are determined with high certainty.

Figure 2.4: Plots of scaled surface density (κ) versus radius for the mock data models. The use of the central images in the fitting enables a tighter constraint to the mass density in the inner region. The median (50th percentile) posterior value of the parameter set is shown as a solid red line, and the 16- to 84-percentile band is shown in gray. The true parameter value is shown as a dashed blue line. The radii bands in which images are located are highlighted in red. A reference redshift of $z = 1.49$ is used in the calculation of κ. The plots are slices that are averaged over 360°. The BCG can be observed as the bump at a radius of approximately 1″.

[*From top*]: Cored without central images, cupsy without central images, cored with central images, cuspy with central images.



2.3.3.3 Constraints on Surface Density

One measure of the utility of the model is the ability to accurately reproduce the (2-dimensional) surface density of the cluster. Here we examine scaled surface density, $\kappa(r) \equiv \frac{\Sigma(r)}{\Sigma_{crit}}$, where $\Sigma(r)$ is the surface density and Σ_{crit} is the critical surface density for the pertinent lens and source redshifts. Figure 2.4 shows the circularly-averaged κ versus radius for the cored and cuspy data sets, both with and without central images. Surface density is very accurately determined in the region where images are present, and the presence of central images enhances the accuracy of the predictions in the inner regions. Only in the radial regions far from the images does the predicted κ deviate significantly from its true value.

2.3.4 Pseudo-Elliptical Approximation

As discussed in Section 2.2, an often-employed approximation is to use an elliptical form for the gravitational potential of the object rather than the density itself [62]. The limits of validity for that approximation is given in Golse and Kneib [62] to be for the range of ellipticities $\epsilon \lesssim 0.25$, which corresponds to an axis ratio $q \gtrsim 0.75$. In Dúmet-Montoya et al. [49], they analyze the effect of the pseudo-elliptical approximation on an NFW profile (i.e., with no core), and conclude that ellipticities up to $\epsilon \lesssim 0.5$ can be used for low values of surface density and in the region of lensing arcs. Here we compare the results of such an approximation to that of using the true elliptical density for both the cuspy (NFW) and cored (cNFW) profiles. The dot-dashed lines in Figure 2.3 show the cored and cuspy cases (without central images) but using the pseudo-elliptical approximation, for the mock data set with axis ratio $q = 0.8$. The pseudo-elliptical model has somewhat less power to resolve the different cases than the model that uses the full elliptical density (solid lines in the figure). Specifically, in the cored case, where the true value of the core radius is 43.3 kpc, the core radius posterior for pseudo-elliptical model peaks at ∼20 kpc, while the true elliptical model shows a peak at ∼30 kpc. In the cuspy case, where the true core radius is zero, the pseudo-elliptical model produces a peak posterior at ∼3 kpc, whereas the true elliptical model has a peak at ∼1 kpc.

Turning now to the mock data sets with high ellipticity (i.e., axis ratio $q = 0.5$), Figure 2.5 illustrates the posterior distribution results for the mass, concentration and core radius parameters. For this very elliptical halo, the models using the full elliptical density profile recover the parameters well, with the true value of all parameters located within the 1-σ posterior contours. In contrast, the pseudo-elliptical approximation does not accurately recover the input parameters. In the cuspy case, the true value for the halo mass is outside the 2-σ contour of the posterior. In the cored case, the true values for halo mass and concentration are both well outside the 2-σ contours of the posteriors. As an example of how this could bias inference of core size, if weak lensing or X-ray constraints were used that constrain halo mass and concentration to be close to their proper values, this will in turn cause r_c to be biased low. The lower left posterior in Figure 2.5 demonstrates that if the halo mass were fixed to its (correct) value of $1.1 \times 10^{15} M_\odot$, the value of r_c would be inferred at approximately $5''$, half as large as the true value ($10''$). This illustrates the dangers of combining different probes to obtain core constraints if systematic errors are present in the lens model.

2.3.5 Model Dark Matter Halo Radial Density Profile

In order to explore the importance of the shape of the assumed halo profile in detecting cores, an alternative set of mock data was constructed using a Corecusp profile for the dark matter halo rather than a cNFW profile. The Corecusp profile is very similar to cNFW but has a faster transition between the inner and outer slopes. (Refer to Section 2.2 for definitions

Figure 2.5: Posterior distributions for halo mass, concentration and core radius parameters, showing the effect of using an approximate elliptical potential for highly elliptical systems. [*Top Row:*] full elliptical density profile used. [*Bottom Row:*] pseudo-elliptical approximation used. [*Left Column:*] cored case. [*Right Column:*] cuspy case. Orange lines and 'x' markers indicate the true parameter values. The units for m_{vir} are $10^{15} M_\odot$, and for core radius, arcseconds.

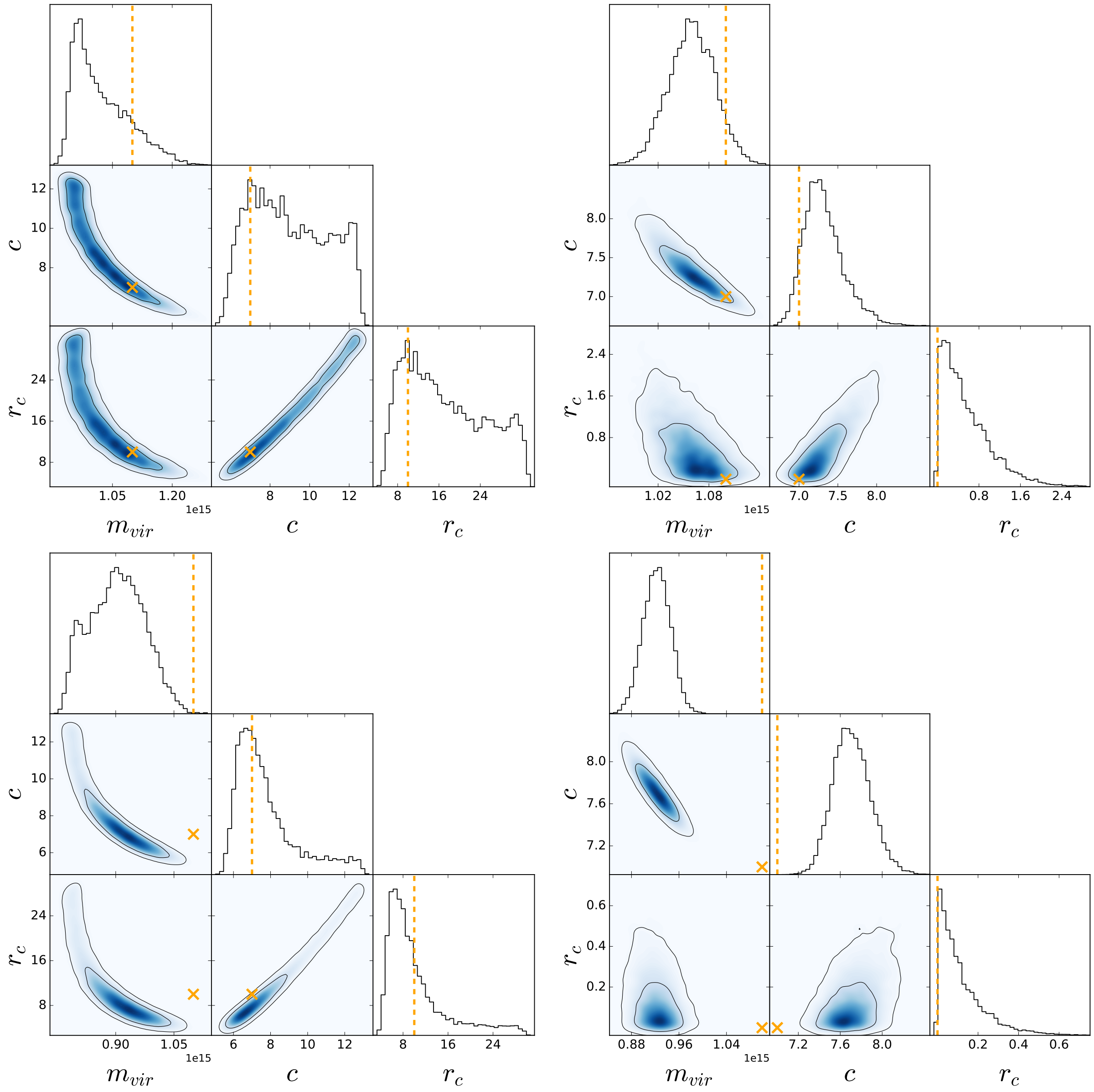
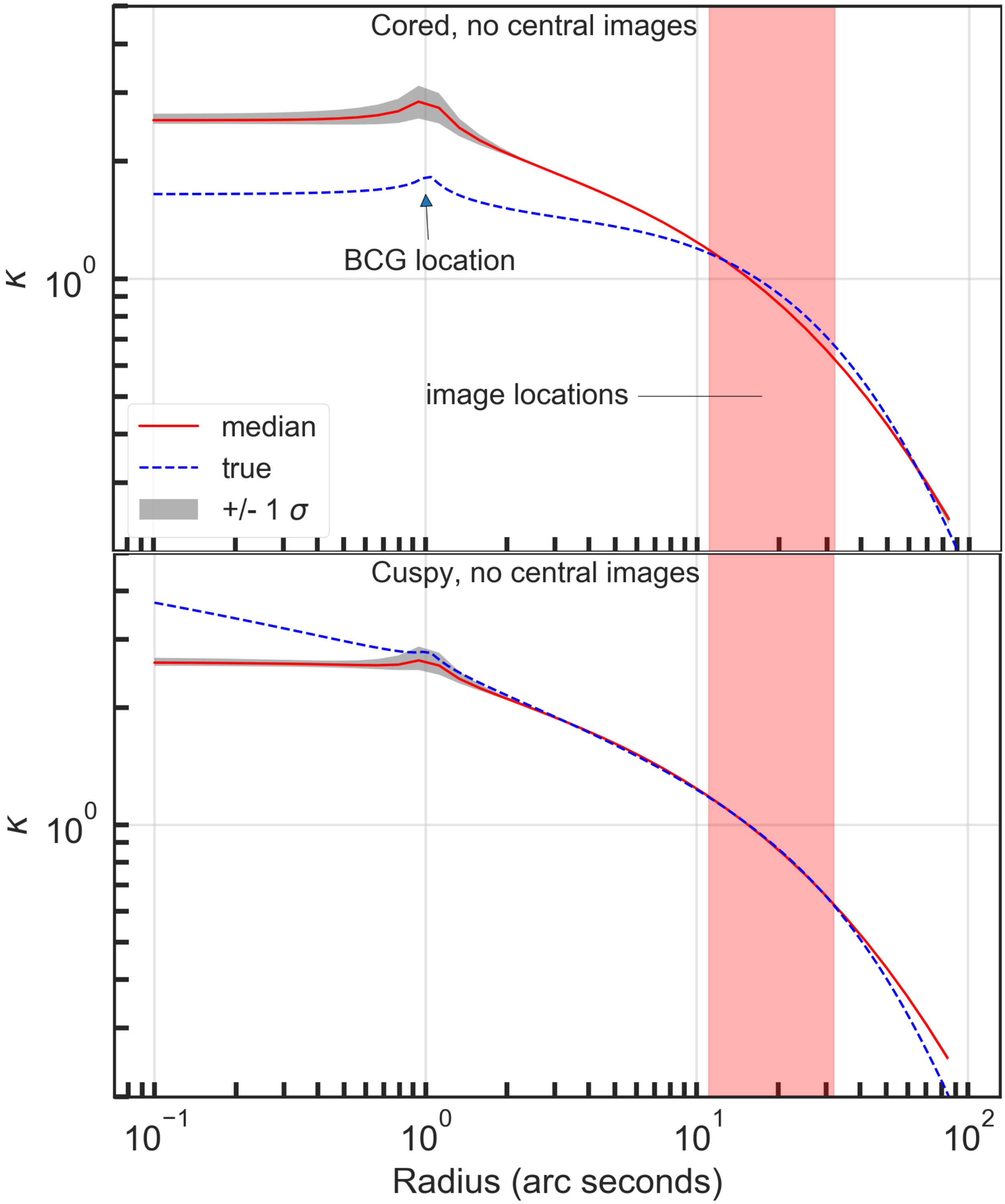


Figure 2.6: Scaled surface density (κ) versus radius for a fit of a cNFW halo lens on mock data generated with a Corecusp profile. The true density profile is shown as a dotted blue line. [*Top*]: Cored case, no central images. [*Bottom*]: Cuspy case, no central images. Note that the accuracy of the modeled profile declines at radii far from the image locations.



of the various density profiles.) The Corecusp parameters for scale radius and core radius were the same as those used in the cNFW mock data, i.e., $63''.7$ and $10''$, respectively. The slope parameters for inner and outer logarithmic slope were set to -1 and -3, respectively, matching those of a cNFW profile.

Figure 2.6 shows the results of fitting a cNFW dark matter halo lens to mock data constructed with a Corecusp dark matter halo, with no central images. The fit is very good in the region of radius values of $10''$to $50''$, close to where the images are located. However, in attempting to fit that area as well as possible, the model predicts a cored halo for the region interior to approximately $1''$, when in fact the halo is cuspy. In the case of the cored halo, the model does predict a core but overestimates the surface density in the core by approximately 40%. Without data in the inner regions to guide it, predictions become unreliable there.

2.3.6 Mock Data with Perturbers

The mock data set described above does not contain perturbing galaxies, and it is possible that the inclusion of such perturbers might make the model parameters more difficult to determine. In addition, the mock data set has four unique redshift groups, and while this is typical of data sets for galaxy clusters, the analysis of Abell 611 in this work relies on data at only two unique redshifts (see Section 2.4.1). To ameliorate the concern that these differences between the mock data and the real data might have a strong effect on the results, we constructed another mock data set for the cNFW lens (only), using only two redshift groups and including two perturbing galaxies. Both cored and cuspy cases were examined.

The results are qualitatively similar to that of the original mock data, i.e., the models faithfully reproduce the data images and recover the model parameters fairly well, with some mild degeneracy between the mass components. The core size is clearly different between the cored and cuspy data sets, allowing the existence of a core to be discerned. Posterior distributions for selected parameters are shown in Appendix A.5.

2.3.7 Summary of Mock Data Results

By modeling these mock data, it is clear that the surface density of the cluster, and thus cored or cuspy characteristics, can be well-predicted in the radial regions having image data points. Central images provide data in the inner regions ($1''$to $8''$in our model) and thus enhance accuracy there. Using a profile shape that sufficiently approximates the true halo profile is important, as mismatches can lead to inaccurate predictions in regions devoid of image data. The use of the pseudo-elliptical approximation can lead to inaccuracies in parameter recovery exceeding 2-σ for highly elliptical halos.

2.4 Modeling of Abell 611

Having tested the sensitivity of our model to varying mass profiles using strong lensing alone, we now turn to the real cluster data. We test for the presence for a core by fitting the data to two different profiles: cNFW and Corecusp.

2.4.1 Abell 611 Data

The data for the modeling of Abell 611 were taken from sources A and B in Table A.1 of Donnarumma et al. [47]. Their originally reported redshifts were 0.908 and 2.06, respectively, however subsequent analysis [see 133, 8] indicate that the correct redshift for source A is 1.49. We have adopted that value. It is interesting to note that Monna et al. [118] included an additional, demagnified central image for Source B in their analysis, in the BCG light at a location of (J2000:120.236 680°, 36.056 140°). We verified that our cNFW and Corecusp models both predict a central image within $\sim 1''$of that location. While we did not include this image in our baseline analysis, the effects of its inclusion are discussed in Section 2.4.3.1.

Source C consists of two points at a reported redshift of 2.59, however the inclusion of this source in our model led to a large shift of the centroid of the dark matter halo, which was inconsistent with the findings from Donnarumma et al. [47] and Newman et al. [130]. This source also has the weakest photometry in the data set, with HST F606W magnitude fainter than 27.0 [156]. Therefore, we decided not to include Source C in the model. Source D is a 4-image source with no confirmed spectroscopic redshift. Donnarumma et al. [47] elected not to include this source in their models, and we also exclude it. We did test the inclusion of sources C and D, and found a substantial increase in the resulting χ^2 of the model. We set the origin to be the coordinates of the BCG as given in Table 4 of Donnarumma et al. [47] (J2000: 120.236 78°, 36.056 572°). The resulting data set contained 25 images in set A and 24 in set B, for a total of 49 images of 13 source points. We adopted $0.''2$ as the position error value, as did Donnarumma et al. [47] (but see Section 2.4.5.1 for a discussion of the importance of that assumption). The images are located in a range of 7 to $28''$from the BCG center.

2.4.2 Abell 611 Lens Model

Following Donnarumma et al. [47], we constructed a lens model with a dark matter halo, BCG, and seven other perturbing lenses. The model parameters are described in Table 2.3. Since we are using various density profiles for the dark matter halo, we need a consistent way to compare concentration and core radius, and have therefore adopted the following definitions. We define "core radius" as the radius at which the logarithmic slope of the density is -1, i.e.,

$$\left.\frac{dlog(\rho(r))}{dlog(r)}\right|_{r=r_{core}} = -1. \tag{2.5}$$

We define the concentration as

$$\hat{c}_{200} \equiv \frac{r_{200}}{r_{-2}} \tag{2.6}$$

where r_{200} is the radius at which the density is 200 times the critical density of the universe at the redshift of the lens, and r_{-2} is the radius at which the logarithmic slope of the density profile is -2. Note that for the cNFW profile, in the limit $r_c \rightarrow 0$ where the profile reduces to NFW, we have $r_{-2} = r_s$.

The Abell 611 system was analyzed using the MultiNest sampling algorithm [56] with 4,000 live points. There were 14, 15 and 13 sampled parameters for the cNFW, Corecusp and NFW profiles, respectively. Considering the 49 image points being generated from 13 sources, we then have 57 to 59 degrees of freedom depending on the lens model. We tested both source plane and image plane chi-square evaluations, and found very similar results for each method. Source plane chi-square evaluations were used for all runs, with additional image plane evaluations made as needed to verify the correct reproduction of multiple images.

2.4.2.1 Cluster Halo and BCG Models

We studied three mass profiles for the cluster halo: cNFW, Corecusp and NFW. Uniform priors were used on all sampled halo parameters except the core scale parameter $r_{c,kpc}$, for which a log prior was used. The BCG was modeled as a dPIE profile with mass parameter b sampled, and other parameters fixed at the values given by Newman et al. [133]. Table 2.3 summarizes the parameter values and ranges.

Table 2.3: The parameters and prior ranges for the dark matter halo, BCG and seven perturbing galaxies in the Abell 611 Lens Model. All priors are uniform over their range except for the parameter $r_{c,kpc}$, for which a log prior was used.

Parameter	Description	Units	Prior Range
DM Halo (cNFW, Corecusp and NFW)			
M_{200}	halo mass	$10^{14} M_{\odot}$	3 - 20
c_{200}	concentration	-	1 - 40
$r_{c,kpc}$ (*)	core scale	kpc	0.001 - 500
γ (**)	inner log slope	-	0 - 2.99
q	axis ratio	-	0.5 - 0.95
θ	orientation	deg.	120 - 160
x-center	x coord. of center	$''$	-5 to 5
y-center	y coord. of center	$''$	-5 to 5
BCG (dPIE profile)			
b	mass parameter	$''$	0.5 - 10.0
a	scale radius	$''$	fixed: 10.7
s	core radius	$''$	fixed: 0.277
q	axis ratio	-	fixed: 0.73
θ	orientation	deg.	fixed: 132.3
x-center	x coord. of center	$''$	fixed: 0.0
y-center	y coord. of center	$''$	fixed: 0.0
Cluster Members (dPIE profile)			
b	mass parameter	$''$	0.05 - 10.0
a	scale radius	$''$	0.05 - 10.0
s	core radius	$''$	fixed (†)
q	axis ratio	-	fixed (†)
θ	orientation	deg.	fixed (†)
x-center	x coord. of center	$''$	fixed (†)
y-center	y coord. of center	$''$	fixed (†)

* cNFW and Corecusp only

** Corecusp only

† See Table A.1 for these values.

2.4.2.2 Cluster Member Models

The seven perturbing lens elements were modeled with dPIE profiles, allowing for separate specification of their mass, core size, cutoff radius, axis ratio, orientation angle and centroid. Perturbers 1 and 2 are quite close to image groups B.4 and B.5, allowing a stronger constraint on their Einstein radii. As such, the mass and cutoff radius parameters for those perturbers were sampled. To avoid a proliferation of parameters, the mass and cutoff parameters for perturbers 3 through 7 were anchored together, allowing two parameters to specify the mass and scale for that group.

The "b" parameter of the dPIE lens is proportional to the lens mass and varies as the square of velocity dispersion (see Equation 2.3). Faber and Jackson [54] show that velocity dispersion scales as $L^{1/4}$, so the relevant scaling relation is

$$b' = b\left(\frac{L'}{L}\right)^{\frac{1}{2}} \tag{2.7}$$

The mass parameters for perturbers 3 through 7 are anchored together according to this relation. The luminosities are shown in Table A.1.

Similarly, the cutoff radii and core radii were scaled using

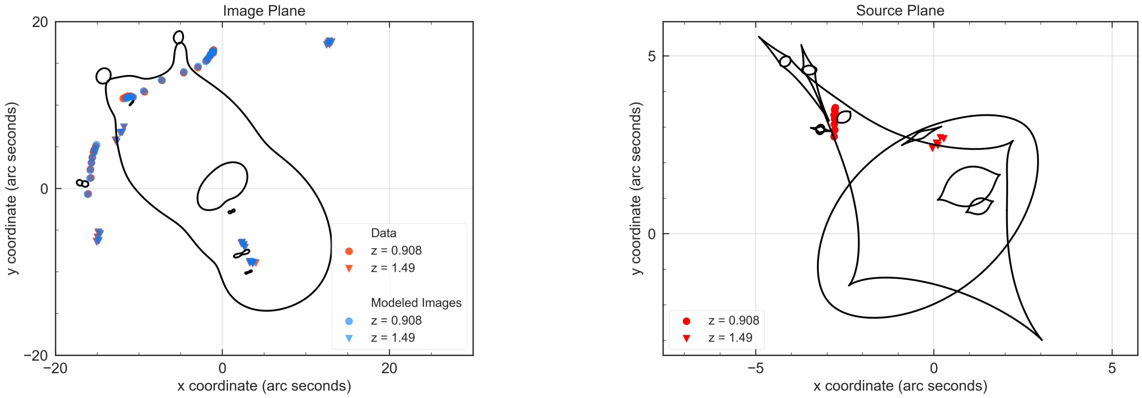
$$r'_{cut} = r_{cut}\left(\frac{L'}{L}\right)^{\frac{1}{2}} \tag{2.8}$$

and

$$r'_c = r_c\left(\frac{L'}{L}\right)^{\frac{1}{2}} \tag{2.9}$$

The exponent 1/2 in the two equations above correspond to a constant mass-to-light ratio among perturbers 3 through 7. The cutoff radius normalization for those perturbers was a

Figure 2.7: Image plane representation (left, with critical curves shown for z=1.49), and source plane representation (right, with caustic curves shown for z=1.49) of the best fit result for the Abell 611 cNFW model. The data points are shown in red, the modeled images in cyan, and show purple where they overlap.



sampled parameter. As the core radii are difficult to constrain without visible images near the core of the cluster member, they were fixed according to a normalized core radius of 0″.035 for an ST magnitude of 18.0, matching the assumption of Donnarumma et al. [47]. For the the centroid locations, axis ratios, and orientations of the perturbers, Donnarumma et al. [47] used GALFIT to determine those values, and we adopt them. They are shown in Appendix Table A.1.

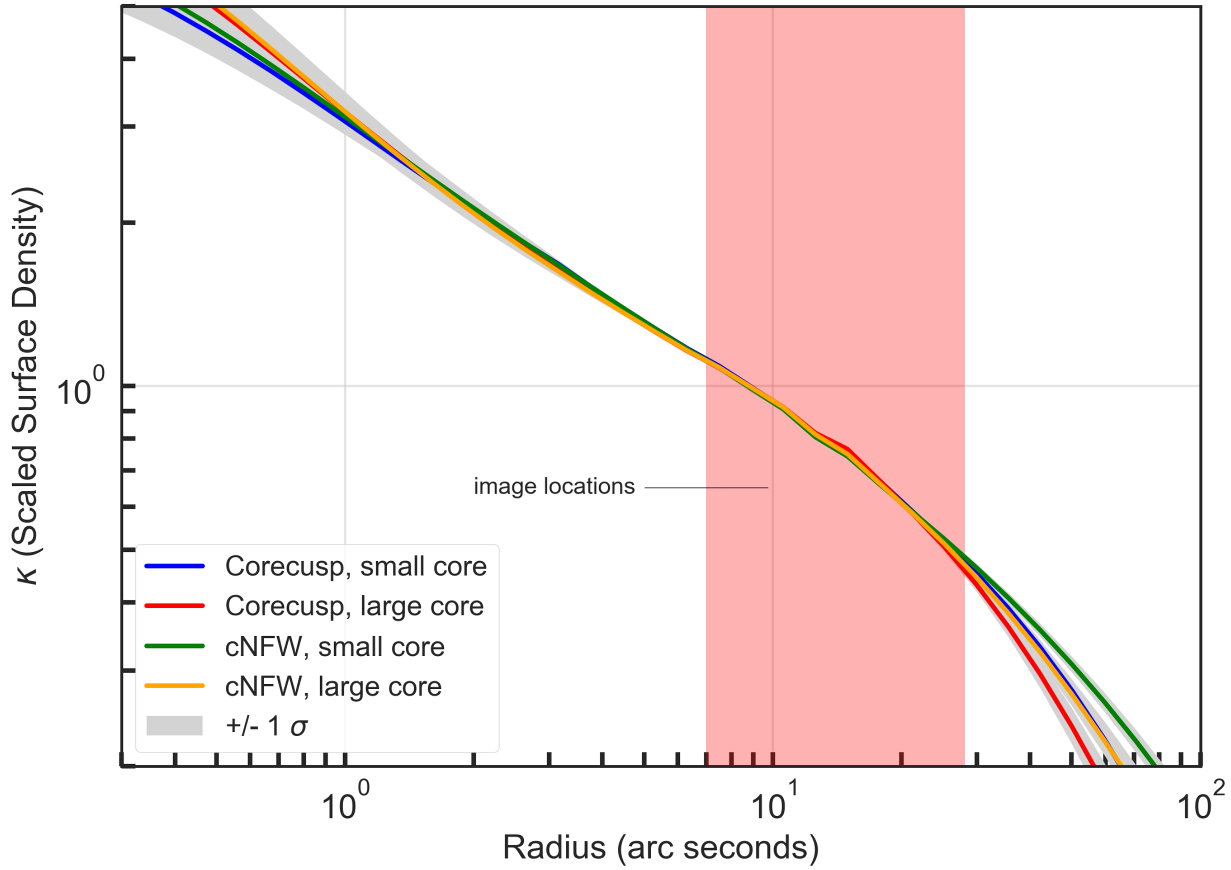
2.4.3 Abell 611 Results

The resulting fits were very good for all the profiles modeled. The reduced χ^2 for the fits range from 0.29 and 0.31, and the RMS errors between the modeled image and data positions range from 0″.12 to 0″.13. The cNFW is our baseline model and is marginally favored by the Bayesian evidence (ln(Z) = -44.8) over the Corecusp model (ln(Z) = -46.6). The resulting best-fit images for the cNFW profile model are shown in Figure 2.7, and the key best-fit parameters for all models are shown in Table 2.4.

Table 2.4: Key median posterior parameter values for the Abell 611 models. The bounds of the 68% confidence interval are also shown. The cNFW and Corecusp models exhibit bimodal solutions for some parameters, allowing either a near-zero core or a large core of 14″to 16″. Parameters labeled "sampled" were allowed to vary during the sampling run, while those labeled "derived" were derived from sampled parameters.

Halo profile			cNFW		Corecusp		NFW
χ^2 [d.o.f.]			17.2 [58]		17.9 [57]		17.0 [59]
Reduced χ^2			0.30		0.31		0.29
ln(Bayesian Evidence)			-44.4		-52.7		-44.7
RMS position error (″)			0.12		0.12		0.12
Mode			Sm. Core	Lg. Core	Sm. Core	Lg. Core	(no core)
Bimodal Parameters	Sampled or derived	Units					
M_{200} (cNFW, NFW)	sampled	$10^{14}M_{\odot}$	$12.7^{+1.4}_{-1.1}$	$8.4^{+1.7}_{-0.9}$			$12.8^{+1.6}_{-1.3}$
M_{200} (Corecusp)	derived	$10^{14}M_{\odot}$			$7.2^{+0.7}_{-0.6}$	$5.4^{+0.8}_{-0.6}$	
$\hat{c}_{200}$	derived	-	$4.1^{+0.3}_{-0.4}$	$6.1^{+0.6}_{-0.9}$	$6.3^{+0.4}_{-0.4}$	$7.6^{+0.6}_{-0.6}$	$4.1^{+0.4}_{-0.4}$
r_{core}	derived	″	$0.3^{+1.2}_{-0.2}$	$13.3^{+1.3}_{-2.0}$	$0.01^{+0.6}_{-0.01}$	$15.6^{+1.5}_{-2.6}$	(0 by def.)
M_{BCG}	derived	$10^{12}M_{\odot}$	$4.9^{+1.2}_{-1.0}$	$7.2^{+1.2}_{-1.0}$	$4.7^{+1.0}_{-0.9}$	$6.8^{+1.2}_{-1.0}$	$4.9^{+1.3}_{-1.1}$
$\kappa_{tot}(5'')$	derived	-	$1.32^{0.01}_{-0.01}$	$1.30^{+0.01}_{-0.01}$	$1.32^{+0.01}_{-0.01}$	$1.29^{+0.01}_{-0.01}$	$1.32^{+0.01}_{-0.01}$
$\kappa_{DM}(5'')$	derived	-	$1.12^{+0.04}_{-0.04}$	$1.02^{+0.04}_{-0.04}$	$1.13^{+0.03}_{-0.03}$	$1.03^{+0.04}_{-0.04}$	$1.12^{+0.04}_{-0.04}$
$\kappa_{DM}(20'')$	derived	-	$0.60^{+0.01}_{-0.01}$	$0.59^{+0.01}_{-0.01}$	$0.60^{+0.01}_{-0.01}$	$0.59^{+0.01}_{-0.01}$	$0.60^{+0.01}_{-0.01}$
γ (Corecusp only)	sampled	-			$1.04^{+0.05}_{-0.03}$	$0.71^{+0.16}_{-0.19}$	
Unimodal Parameters	sampled or derived	Units					
position angle	sampled	degrees	$133.3^{+0.2}_{-0.2}$		$133.3^{+0.3}_{-0.2}$		$133.3^{+0.3}_{-0.2}$
axis ratio	sampled	-	$0.67^{+0.01}_{-0.01}$		$0.67^{+0.01}_{-0.01}$		$0.67^{+0.01}_{-0.01}$
x-center	sampled	″	$-0.2^{+0.3}_{-0.3}$		$-0.3^{+0.3}_{-0.3}$		$-0.2^{+0.2}_{-0.2}$
y-center	sampled	″	$0.7^{+0.3}_{-0.3}$		$1.0^{+0.3}_{-0.3}$		$0.7^{+0.3}_{-0.2}$

Figure 2.8: Scaled surface density (κ) versus radius for Abell 611. cNFW and Corecusp models are shown, each subdivided into large-core and small-core solutions.



The posterior distributions for the cNFW and Corecusp models both exhibit a bimodal solution for the lens model parameters. There is a clear "small-core" mode, with a core radius $< 1''$, and a "large-core" mode, with core radius $\sim 15''$. Two-dimensional posterior distribution plots for selected parameters for the cNFW and Corecusp cases are included in Appendix Figure A.3 and Figure A.4, respectively. The χ^2 for the best fit points of each of the two modes are very similar: 16.5 for the small-core cNFW mode and 17.1 for the large-core cNFW mode. The small-core mode is associated with a higher halo mass and a lower BCG mass, while the large-core mode is reversed in that regard. In the discussion that follows, we divide the posterior parameter space for bimodal parameters into large-core (i.e., $r_{core} > 10''$) and small-core ($r_{core} < 3''$) modes. Table 2.4 shows those posterior values.

There is clearly a degeneracy between the halo core and BCG mass, as the halo and BCG are nearly co-centered (approximately 1″apart), and it is the combination of their masses that determines the surface density and hence deflection angles. In evaluating these two modes, we can ask whether the resulting BCG mass is consistent with prior constraints on early-type galaxies. The luminosity of the BCG was found to be $5.47 \times 10^{11} M_{\odot}$ in V-band [133], which, given the median BCG masses in Table 2.4 would imply best-fit stellar mass-to-light ratios of 9 and 13 for the small-core and large-core modes, respectively. At the low-mass end, we infer the small-core BCG mass $\gtrsim 3.0 \times 10^{12} M_{\odot}$ at 95% CL, equating to a minimum stellar mass-to-light ratio of 5.5; for comparison, the large-core mode requires a BCG mass $\gtrsim 5.6 \times 10^{12} M_{\odot}$ at 95% CL, equating to a stellar mass-to-light ratio of 10.3. Such high mass-to-light ratios imply that both solutions require a steep stellar initial mass function (IMF). However, as we will show in Section 2.4.3.2, the IMF slope required by the small-core mode is consistent with recent constraints from high-mass early-type galaxies, whereas the large-core mode is inconsistent with these constraints.

As might be expected, the NFW halo model produces nearly identical posteriors to the small-core mode cNFW solution, albeit with a slightly higher χ^2 (17.0 versus 16.3 for cNFW). The concentration $\hat{c}_{200}$ is 4.1 for the small core solution, consistent with previously studied mass-concentration relations [129], although it should be noted that those relations were created for systems with NFW profiles, and may not be easily compared to other forms of profiles that have cores. Interestingly, the small-core mode of the Corecusp model prefers an inner slope of 1.05, very similar to an NFW inner slope. However, the Corecusp solution is more concentrated ($\hat{c}_{200} = 6.2$) and has a much smaller halo mass, as can be seen in Table 2.4.

The resulting posterior distributions for many parameters are similar between the cNFW profile and the Corecusp profile, and are unimodal. These include orientation angle (θ), centroid location (x_c, y_c), axis ratio (q) and scaled surface density (κ). In particular, κ_{tot} at 5″ is very well constrained and is remarkably consistent between the models, varying between

1.28 and 1.32. A plot of κ versus radius is shown in Figure 2.8, with cNFW and Corecusp posteriors separated into their large-core and small-core components. Their median values are in close agreement in the range of radii between $1''$and $30''$.

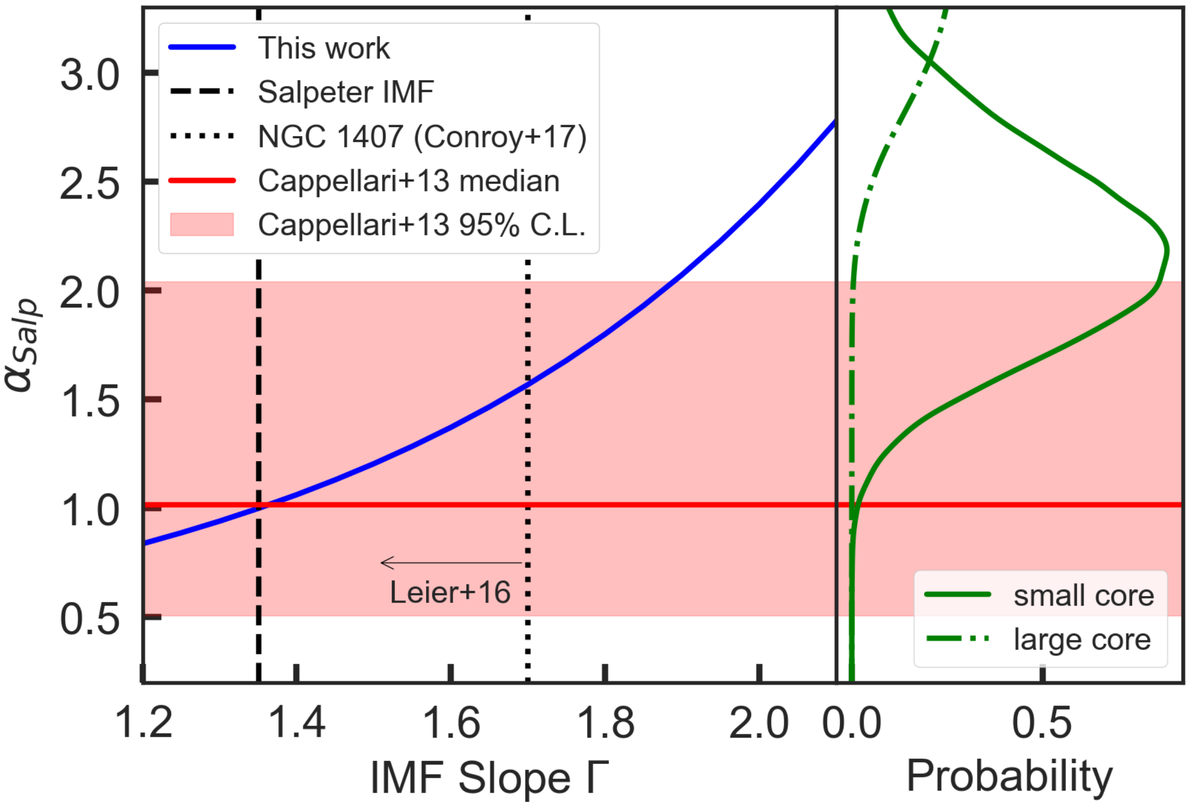
2.4.3.1 Consideration of a Central Image

Monna et al. [118] included a demagnified, positive-parity central image in their data. While our results above do not include this image, we did examine the effect of including such an image in the data as a part of our image set B. Image set B consists of 5 points along 5 tangential arcs, for a total of 25 data points. We inserted a sixth trial data point in each image, at a location of (J2000:120.236 680°, 36.056 140°), which corresponds to coordinates of ($0\overset{''}{.}29$, -$1\overset{''}{.}56$) relative to the BCG center. We assumed a position uncertainty of $1''$ for each of the x and y coordinates. With the additional image, the cNFW model yielded a small-core-only solution, with $\chi^2 = 25.5$ and Bayesian evidence $ln(Z) = -50.9$. While this χ^2 and Bayesian evidence are marginally poorer than the baseline, the fact that it selects the small-core mode strengthens the case for the selecting the small-core solution to the apparent core size degeneracy.

2.4.3.2 Implications for the Stellar Initial Mass Function of the BCG

The high stellar mass-to-light ratio we have inferred for the BCG would suggest a fairly steep stellar initial mass function. Given the high central dispersion of the BCG, this is not suprising: many authors in recent years have inferred a bottom-heavy IMF in early-type galaxies, using either spectral lines [93, 24, 36, 102, 35] or strong lensing [95, 133]. In [93] and [24], spectra from a large sample of early-type galaxies (SPIDER and ATLAS-3D, respectively) were analyzed, revealing a trend in the IMF log-slope: galaxies with low central dispersions show a shallow slope consistent with a Chabrier/Kroupa IMF [27, 92], whereas

Figure 2.9: Mass-to-light mismatch parameter α_{Salp} versus IMF slope Γ, where α_{salp} is defined as the ratio of M_*/L_V produced by this IMF over the value expected from SPS models (found in Newman et al. 2013) using a Salpeter IMF. Constraints from galaxy surveys in Cappellari+ 2013, Conroy+ 2017 and Leier+ 2016 are overlaid. The right panel shows the posterior probability densities in α_{salp} for the small-core and large-core solutions.



galaxies with higher dispersions show a steeper slope, comparable to or even steeper than that of a Salpeter IMF [166].

This begs the question, are either our small- or large-core solutions compatible with constraints on the IMF in early-type galaxies? These solutions require stellar mass-to-light ratios of at least $M_*/L_V \gtrsim 5.5$ and $\gtrsim 10.3$, respectively (at 95% CL). Since M_*/L_V depends on other factors besides the IMF (e.g. metallicity, stellar ages), one way to compare IMF constraints is to define the "IMF mismatch" parameter $\alpha_{salp} = (M_*/L_V)/(M_*/L_V)_{salp}$, where $(M_*/L_V)_{salp}$ is the mass-to-light ratio generated by a Salpeter IMF. Positive α_{salp} values then would imply an IMF that is more bottom-heavy compared to Salpeter. In [133], the stellar mass-to-light ratio M_*/L_V for Abell 611 was estimated using the BCG colors and a stellar population synthesis model. Under the assumption of a Salpeter IMF, they infer $(M_*/L_V)_{salp} = 3.98$. Using this, our small- and large- core solutions require α_{salp} values of at least 1.38 and 2.59, respectively, at 95% CL.

In [24] a trend line is fit to $\log \alpha_{salp}$ as a function of dispersion (see their Figure 13), with corresponding lines for 2.6σ scatter. Using the fact that the luminosity-weighted dispersion of the BCG within its half-light radius is 306 km/s, we estimate the median value for $\alpha_{salp} \approx 1.0$ with the $\pm 2.6\sigma$ bounds at 0.5 and 2.0. [102] find a similar range using galaxies in the CALIFA survey, for which α_{salp} lies in the approximate range 0.6-1.5, while [95] find a similar range 0.5-1.5 in the SLACS lens sample. Given these constraints, it is evident that our small-core solution ($\alpha_{salp} \gtrsim 1.4$) is compatible with current constraints, whereas the $\alpha_{salp} \gtrsim 2.6$ required by the large-core solution lies beyond the upper bound for all the surveys mentioned here.

Next, we go further and estimate the constraint on the slope of the IMF of the BCG in Abell 611 from our lensing analysis. We will model the IMF using a double power-law model, $\xi(M) \propto M^{1-\Gamma}$ where $\Gamma = 1.35$ (Salpeter) for $M > M_\odot$ while Γ for $M < M_\odot$ will be freely varied. This is identical to one of the models used in [95] and one of the parametric models used in [36]. To relate the mass-to-light ratio to the IMF slope Γ, we have

$$\frac{M_*}{L_V} = \frac{\int_{M_{low}}^{M_{high}} M^{2-\Gamma}}{\int_{M_{low}}^{M_{high}} L_V(M) M^{1-\Gamma}} \tag{2.10}$$

where $L_V(M)$ is the V-band stellar luminosity-mass relation. For the lower mass cutoff we adopt the usual convention $M_{low} = 0.1 M_\odot$, and the high mass cutoff M_{high} will be determined by the particular isochrone used. For the mass-luminosity relation we use the Padova isochrones [61] assuming metallicity $Z = Z_\odot$ (the same choice was adopted in [133]), and consider a few different stellar ages. To account for the fact that the stellar ages inferred for Abell 611 may differ from our choices here (along with possible slight differences in the SPS model used), we will write $L_V(M) = \lambda L_{V,0}(M)$ where $L_{V,0}(M)$ is generated from the isochrone, and λ is a correction factor which we expect to be close to 1 if the correct median stellar age is assumed. For a given assumed stellar age, we perform the integration in Eq. 2.10 by interpolating over a table of values in $L_{V,0}(M)$ generated from the isochrone, then solve for λ using the [133] values ($M_*/L_V = 3.98$, $\Gamma = 1.35$). With λ in hand, we can then use Eq. 2.10 to plot the IMF slope Γ needed to produce a given stellar mass-to-light ratio. In practice, we find that the results are nearly identical regardless of stellar age (we tried ages in the range of 6-10 Gyr), since the luminosity is sensitive to the IMF slope at high stellar mass which is still Salpeter in our model; since $\lambda \approx 1$ for 8 Gyr, we adopt this stellar age in the following.

In Figure 2.9 we plot α_{salp} as a function of IMF slope Γ, while on the right side is plotted the posteriors of the small- and large-core solutions in α_{salp}. Note that the curve equals 1 at $\Gamma = 1.35$, since the above procedure enforces consistency with the [133] results. From this figure we see that an IMF slope $\Gamma \approx 1.5$ is required to produce a $\alpha_{salp} \approx 1.4$, which is close to the 95% CL lower limit required by our small-core solution. By contrast, a slope $\Gamma \gtrsim 2$ is required to be consistent with the large-core solution. Among the SLACS lenses, [95] found that for the double power law IMF, a slope of 1.7 implies a dark matter fraction of zero, and hence steeper slopes are ruled out. Indeed, [27] argue that in the most extreme starburst conditions, the IMF "saturates" at a slope $\Gamma \approx 1.7$. Recently [36] investigated a galaxy with a similar dispersion to ours (NGC 1407) and found $\Gamma = 1.7$, possibly reaching the saturation limit. While the small-core solution is consistent with these constraints, the large-core solution clearly is not, painting a consistent picture with the above constraints on α_{salp}.

Our inference of a stellar IMF slope $\Gamma \gtrsim 1.5$ carries some important caveats. First, any inference about the IMF slope depends on the form of the IMF used. A popular variant is the "bimodal" IMF [192], favored in several recent studies [102, 95, 94], which uses a variable slope Γ_b at $M > M_\odot$ whereas the slope tapers to zero at low masses. As an additional check, we repeated the above analysis using the bimodal IMF and found that a slope $\Gamma_b \gtrsim 3$ is required for the small-core mode. This is near the upper limit of the ranges observed in surveys (Lyubenova et al. 102 find a maximum $\Gamma_b \approx 3.1$, while La Barbera et al. 93 infer $\Gamma_b \sim 3.0$ for ~ 300 km/s dispersions); again, the large-core mode requires a much higher slope and hence is likely ruled out.

Another caveat is that the IMF slope may vary with radius, as recent studies have suggested [94, 209, 191]. This is important because the presence of the BCG in the total projected density is only distinct out to ~ 0.3 times the effective radius, as can be seen from the size of the "bump" in Figure 2.8 (note that the effective radius is ~ 10 arcsec). Thus, we may only be sensitive to the stellar mass in the inner regions, where the mass-to-light ratio is high. If the IMF indeed becomes shallower further out, then the *total* M_*/L_V may potentially be lower than we infer from the strong lens modeling. It would also imply that the stellar mass profile becomes steeper than the light profile at larger radii, which could be an important systematic when inferring stellar masses from lensing. Allowing for a possible steepening of the stellar mass profile relative to the light profile when doing the lens modeling (to account for this systematic) is left to future work.

2.4.3.3 Performance of the Pseudo-Elliptical Model

The pseudo-elliptical cNFW model yielded a best-fit χ^2 that was slightly higher than the corresponding elliptical cNFW fit, and the resulting posteriors were generally similar for most parameters. Interestingly, the median posterior value for the BCG mass was approximately 50% higher when using the pseudo-elliptical approximation. This is remarkable given that the ellipticity is not extreme: the inferred ellipticity of the potential contours is $\epsilon \approx 0.19$ (where ϵ is defined the same as in Golse and Kneib 62), which is low enough that it might appear “safe” to use the pseudo-elliptical approximation. By contrast, the cNFW fit inferred an axis ratio $q \approx 0.67$ for the density contours, markedly lower than one might have naively expected from the pseudo-elliptical fit. (We note that in our mock data runs, there did not appear to be systematic difference in the BCG mass when using the pseudo-elliptical approximation, even in the case of high ellipticities.)

The inferred BCG mass using the pseudo-elliptical model makes the best-fit stellar mass-to-light ratio even higher (≈ 13 for the small-core mode), making it much harder to reconcile with IMF constraints for early-type galaxies. We conclude that the pseudo-elliptical model can bias the results significantly even if the inferred ellipticity is not extreme, and hence modeling lenses with true elliptical density contours is strongly preferred.

2.4.3.4 Consideration of Stellar Kinematic Data

Newman et al. [133, 132] used long-slit spectroscopic observations of velocity dispersion in the BCG and spherical Jeans equation analysis to find a χ^2 for those stellar kinematic data, which is then incorporated into their overall fit. The attraction to this approach is that it incorporates constraints from the inner region of the cluster, where there are no strong lensing images due to the bright BCG image. They assume that the BCG is centered at the same location as the dark matter halo and that the system is spherical. For their fiducial case they assume an isotropic system, i.e., $\beta_{aniso} = 1 - (\sigma_\theta^2/\sigma_r^2) = 0$, but they also ran models for β_{aniso} values between -0.2 and +0.2, with constant values of β in all cases. We adopt their dispersion observations and error values. However, we excluded the innermost point from the analysis, as that point is subject to systematic error from slit and seeing effects that are greater than the observational error (A. B. Newman, personal communication, December 7, 2018). We then apply the spherical Jeans analysis.

As a starting point, we used a cNFW model similar to our baseline but adopt a fixed BCG mass of $1.5 \times 10^{12} M_{sun}$, which is similar to the mass found in Newman et al. [133] and Donnarumma et al. [47]. We then produced a "large core" chain, with $r_{core} > 10''$, and a "small core" chain, with $r_{core} < 3''$. We analyzed an isotropic case with $\beta = 0$ as well as mildly radially and tangentially biased cases with $\beta = \pm 0.2$.

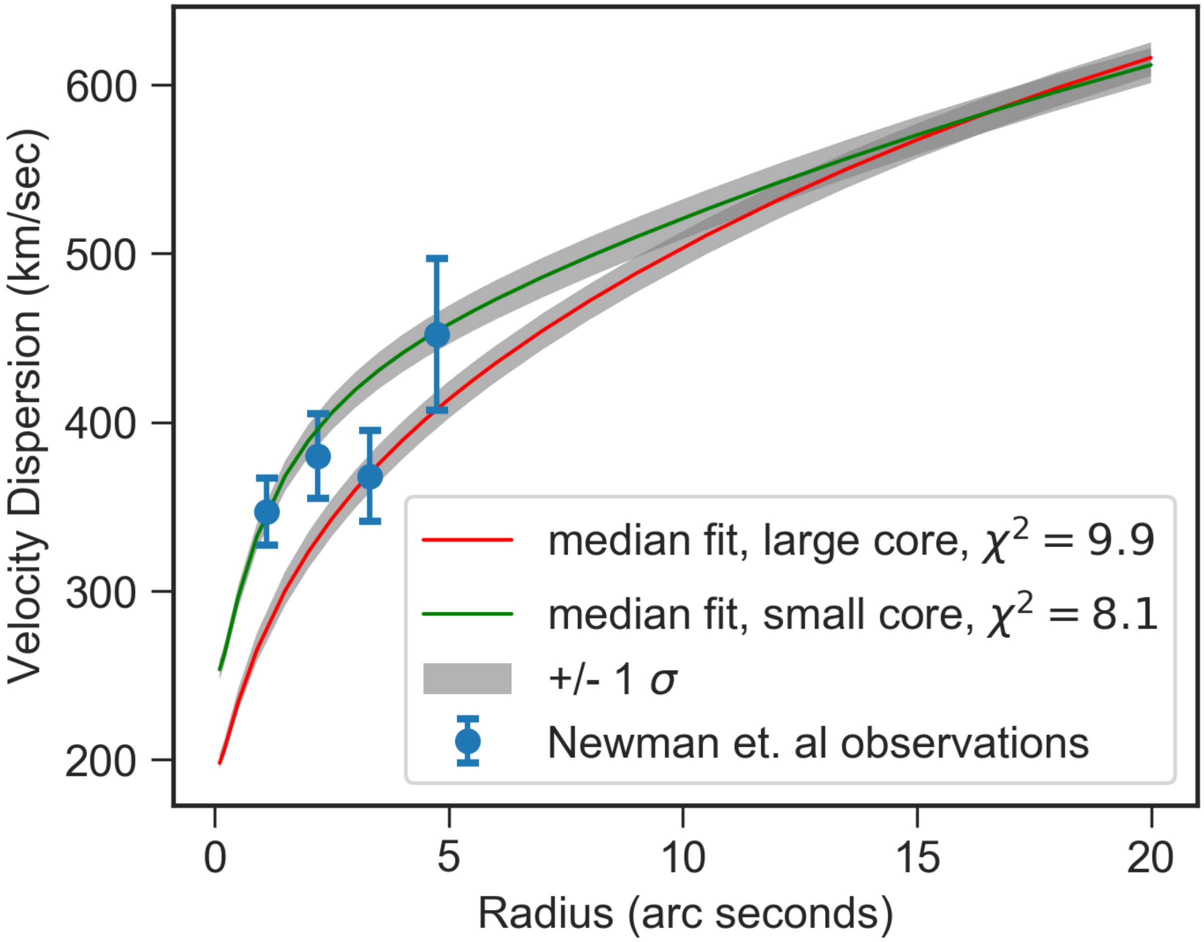
Following Cappellari [25], the velocity dispersion over the line of sight can be found from

$$\sigma^2_{BCG,LOS}(R) = \frac{2G}{\Sigma_{BCG}(R)} \int_R^\infty \frac{\mathcal{F}(r, R, \beta) \rho_{BCG}(r) M(r)}{r^{2-2\beta}} dr, \tag{2.11}$$

where Σ_{BCG} is the BCG surface density (derived in our case from the 3D dPIE profile), ρ_{BCG} is the dPIE density profile, $M(r)$ is the mass of all components generating the potential and $\mathcal{F}(r, R, \beta)$ is an analytic function derived in Cappellari [25].

The velocity dispersions assuming $\beta = 0$ at all radii is shown in Figure 2.10. The small core and large core cases both provide plausible fits to the stellar kinematic data. As β is varied over a modest range of -0.2 to $+0.2$, the fits change from favoring a small core to favoring a large core. These represent a constant value of β at all radii; if β were allowed to vary with radius, even within these modest bounds, a wide variety of solutions could be accommodated. In Schaller et al. [170], they examined six simulated clusters similar in size and character to Abell 611 and found that β did indeed vary beyond the range of -0.2 to $+0.2$, and could vary significantly over the radius of a cluster. We conclude that the velocity dispersion data does not provide a meaningful constraint for our purpose of discerning core size, because a wide range of core sizes can be fit by the data with only minor variations in anisotropy, and assuming $\beta = 0$ fits both large and small cores equally well. We note that the data in this case extends only to $\sim 5''$, whereas the half-light radius of the BCG is $10''.7$, limiting the influence of the data. However, as data becomes available at larger radii and with less noise, it will offer more constraining power. Also, with more data, it may be possible to use two-dimensional kinematic analysis and/or higher order moments of the

Figure 2.10: Velocity dispersion of the BCG, and corresponding fit to the observations of Newman et al. [133], assuming $\beta = 0$.



velocity dispersion to constrain the anisotropy.

2.4.4 Potential Systematic Errors

Variations on the baseline models were created in order to examine the possible effects of systematic errors. One such possible source is external shear from perturbing galaxies that are close in projection to the line of sight to the lens. We found that including external shear had the effect of changing the posteriors for the halo orientation angle θ by as much as

$\sim 10^\circ$, as well as the centroid coordinates x_c and y_c, in some cases by more than $1''$. However, we did not observe significant effects on the posteriors of other parameters, and including external shear did not significantly improve the fit.

The type of prior distribution can also impact the modeling results. As our baseline, we used uniform priors with wide ranges (see Table 2.3), except for the core radius parameter, for which we used a log prior. We tested the impact of using log priors for the mass parameters of the BCG and perturbers. These did not have a significant impact on the resulting model posteriors or fit metrics.

Another source of systematic error is the triaxiality of the cluster, since lensing is only sensitive to the projected mass. Depending on the projection and axes ratios, the projected ellipticity could vary significantly. We do not yet have the models in `QLens` to take this complication into account. This issue also becomes important when comparing to other probes such as weak lensing (sensitive to projected outer halo shape) and velocity dispersion measurements (sensitive to the 3d mass profile in the inner region) [131].

2.4.5 Comparison to Other Works

Abell 611 has been studied by numerous other groups, utilizing a variety of techniques, including strong lensing, weak lensing, X-rays and stellar kinematics. Our emphasis is strong lensing, so here we focus our comparison with strong lensing results of others where possible. The work of Donnarumma et al. [47], Newman et al. [130, 133, 132], Monna et al. [118] are particularly relevant.

The predicted value for core size varies significantly in the literature. Monna et al. [118] use velocity dispersion measurements of 17 cluster members in their strong lensing analysis, and infer a core size of $5.8^{+2.0}_{-1.6}{}''$, although they assume a dPIE profile for their halo. Their result has a reduced χ^2 of 0.7 and they assume position errors of $1''$. Donnarumma et al. [47] uses an NFW halo and so does not examine core size. Newman et al. [132] find a core size of $\sim 0.7''$ in their cNFW model, in which they have a reduced χ^2 of ~ 1 and they assumed position errors of $0.5''$. Our preferred solution (i.e., cNFW, small-core mode) is for a core size of $<1''$, with a reduced χ^2 of 0.28 and assumed position errors of $0''.2$.

In our preferred model (cNFW, small-core), the median posterior value of the dark matter halo mass M_{200} is $1.2 \times 10^{15} M_\odot$. As several of the prior analyses [47, 130, 156] used an incorrect value for the redshift of one of the sources, their strong lensing mass results are not directly comparable, so instead we compare to their weak lensing and X-ray results. The X-ray analysis of Donnarumma et al. [47] found a mass of $\sim 1 \times 10^{15} M_\odot$. Newman et al. [133], which did use the correct source redshifts, found a halo mass of $8.3 \times 10^{14} M_\odot$ in their combined analysis. Romano et al. [164] used two weak lensing techniques and various model profiles, and found M_{200} to be in the range of $5.3 \times 10^{14} M_\odot - 5.9 \times 10^{14} M_\odot$, with moderate uncertainties. In Richard et al. [156], their X-ray analysis puts the 2-D projected mass within $R < 250$ kpc as $2.06 \times 10^{14} M_\odot$, while the same statistic for our model is $2.12 \times 10^{14} M_\odot$, similar to theirs.

2.4.5.1 The Importance of Position Errors

The magnitude of assumed positional errors σ_{pos} of the observed image positions directly impacts the χ^2 of the strong lensing model, as σ_{pos}^2 appears in the denominator of the equation for χ^2. This in turn impacts comparisons with other modeling methods. When combining strong lensing analysis with other approaches such as weak lensing, stellar kinematics or X-ray analysis, authors often assume a strong lensing positional error that accommodates possible deficiencies in the lens models [132, 212]. In our case, this is not a consideration since we are only employing one type of analysis. In addition, systematic errors (see discussion in Section 2.4.4) are often difficult to quantify, and an attempt is sometimes made to account for those errors by increasing the assumed positional error, sometimes dramatically.

Donnarumma et al. [47] assumed positional errors of $0\farcs2$, while Newman et al. [132] used $0\farcs5$, Monna et al. [118] used $1\farcs0$ and Zitrin et al. [212] used $1\farcs4$. We made a model run with the positional error as a sampled parameter, resulting in a best-fit value of $0\farcs18$. We ultimately adopted a position error value of $0\farcs2$. Nevertheless, the reduced χ^2 of our model is quite low at 0.30 (although we did exclude images that would have raised that, as discussed in Section 2.4.1). Had we used higher values of position error such as $0\farcs5$ or $1\farcs4$, the reduced χ^2 would have been 0.048 or 0.0061, respectively.

2.5 Conclusions

Our main aim in this paper was to put robust constraints on the dark matter densities in the central regions of clusters using strong lensing alone. Constraints on the central dark matter density of clusters is critical for constraining the particle physics of self-interacting dark matter models.

We used simulated cluster data to test whether strong lensing data could distinguish between cuspy and cored data sets (with a core radius of $10''$), both with and without central images present. The non-central images were in the 10" − 30" range, in agreement with observed images. Our main findings from the analysis of mock data are as follows.

- It is possible to distinguish between the cored and cuspy data sets, even in absence of central images, provided the density profile and shape of the density contours are accurately modeled. For the cored halo mock data with core radius of $10''$, we infer a core radius greater than $3\farcs89$ at 95% confidence level. For the cuspy data set, we infer a core radius less than $1\farcs01$ at 95% confidence level.

- Approximating the potential with a pseudo-elliptical model rather than using a true elliptical density can degrade parameter recovery for strongly elliptical halos. In the case of a dark matter halo with axis ratio $q = 0.5$, the halo mass and concentration parameters were both outside their 2-σ contours. Although the inferred core size was not significantly biased in this case, combining these results with other probes such as weak lensing to better constrain the mass distribution would likely bias the inferred core size significantly, illustrating a specific danger of combining multiple probes when modeling systematics are present.

- The use of a radial density profile with a different shape than that of the mock data caused the inferred surface density (hence, core size) in the regions void of images (i.e., either near the center or on the outskirts of the cluster) to be biased. We find this effect can be severe enough to make a cored halo appear cuspy and vice versa, even though the profile remains well-fit in the range of radii where the images are located. This systematic can be alleviated if visible central images are present in the data.

With these lessons, we modeled Abell 611 with two halo profiles (“cNFW" and “Corecusp") that allow for a variable core size, a model for the BCG and seven cluster members (see Section 2.4.2). Our main findings are the following.

- Both the cNFW and Corecusp models found similar solutions. The cNFW model has the lower χ^2 and is the preferred model with higher Bayesian evidence. Reduced χ^2 values of 0.28 and 0.30 were obtained for the cNFW and Corecusp models, respectively, even with a small value of assumed position error of $0\farcs2$.

- A bimodal solution was found for key parameters such as core size, halo mass and BCG mass. The large-core solution did not allow for reasonable values of BCG stellar mass-to-light ratios, with $(M_*/L_V) > 10.3$ at 95% confidence level. For the small-core solution, we found $(M_*/L_V) > 5.5$ at 95% confidence level. The required (M_*/L_V)

for the large core solution is not consistent with the measurement of stellar mass-to-light ratios in ATLAS3D early-type galaxies. The required slope of the IMF for the large-core solution is also inconsistent with various inferences [36, 95], as summarized in Figure 2.9. In addition, a cNFW model that included a central image for source B as was used by Monna et al. [118] strongly favors the small-core solution. This evidence points to the small core as the reasonable solution for Abell 611, consistent with the finding of Newman et al. [132].

- We infer a bottom-heavy IMF for the BCG, with IMF log-slope $\Gamma \gtrsim 1.5$ (per logarithmic interval) for stellar mass below $M_\odot$, at 95% C.L. Since the lensing data are most sensitive to the BCG mass within $\lesssim 0.2$ times the half-light radius, this result is consistent with recent studies that find an extreme bottom-heavy IMF at the centers of massive elliptical galaxies.

- Fitting the pseudo-elliptical halo model to Abell 611 results in an inferred BCG mass that is 50% larger compared to using the true elliptical density. This inflates the stellar mass-to-light ratio significantly, despite yielding a low inferred ellipticity ($\epsilon \approx 0.19$), and illustrates the danger of fitting the pseudo-elliptical model even in cases where the inferred ellipticity may not be extreme.

- The scaled surface density (κ) at $5''$ is found to be 1.32 ± 0.01, and is a particularly well-constrained parameter in all models. We expect this to be a key constraint on models of self-interacting dark matter. The inferred core density and core size (for the preferred small-core solution) are consistent with those found previously by Newman et al. [132], whose results were used by Kaplinghat et al. [78] to argue that Abell 611 prefers a self-interaction cross section over mass of about $0.06^{+0.07}_{-0.03}\mathrm{cm}^2/\mathrm{g}$ at a relative velocity of about 1500 km/s. Our robust inference of the core size in Abell 611 underscores the promise of density profile measurements in galaxy clusters to measure the self-interaction cross section of dark matter with a precision of 0.1 cm^2/g or better.

- The existing kinematic data for Abell 611 only go out to about half of the half-light radius and thus are highly sensitive to the unknown velocity dispersion anisotropy parameter β_{aniso}. With more data that provide constraints on velocity dispersion to 2-3 half-light radii, we expect that the velocity dispersion constraints will play an important role in constraining the mass-to-light ratio of the BCG and hence the underlying dark matter halo profile.

We have shown how gravitational strong lensing can be used to put robust constraints on the dark matter halo core size and core density in galaxy clusters. Our results for Abell 611 prefer a high central density and small core size. The corresponding constraint on the self-interaction cross section at velocities of about 1500 km/s is expected to be at the 0.1 cm^2/g level, which would be the tightest constraint on the dark matter self-interaction cross section.

Acknowledgements

We would like to thank Drew Newman for providing valuable feedback on the manuscript and for insightful discussions during the course of the project. We gratefully acknowledge a grant of computer time from XSEDE allocation TG-AST130007. QM was supported by NSF grant AST-1615306 and MK was supported by NFW grant PHY-1620638.

Chapter 3

A Stringent Upper Limit on Dark Matter Self-Interaction Cross Section from Cluster Strong Lensing

3.1 Introduction

The currently-favored theory of the formation of galaxies in the Universe, dark energy plus cold dark matter or ΛCDM, is remarkably successful in explaining many of the observations of large-scale structures [194, 171]. Dark-matter-only simulations show that structures form in a heirarchical manner, and result in dark matter (DM) halos that are well-approximated by a Navarro-Frenk-White (NFW) density profile [125, 60]. Statistically significant departures from the NFW halo profile have been reported for galaxy clusters [168, 146, 107, 133, 132, 40, 41, 4], which provides a fertile ground for the exploration of DM physics. For galaxy clusters, there are a number of studies that point to the existence of DM cores in some galaxy clusters [168, 133, 132, 40, 41, 4]. One frequently proposed explanation for galaxy clusters with cores is that of Active Galactic Nuclei (AGN) feedback, where a black hole accreting gas can cause mass blowout and heating via gravitational interactions, leading to cores [48, 186, 149, 106, 107, 142]. In Martizzi et al. [107], they show that AGN feedback can potentially cause cores in their simulated halo of mass $1.42 \times 10^{13} M_{\odot}$. On the other hand, [170] used the hydrodynamical EAGLE simulations of six galaxy clusters with $m_{200} > 10^{14} M_{\odot}$, which employ weaker AGN feedback, and found instead that feedback does not produce cores. Numerical cosmological simulations have advanced steadily in resolution and accuracy over the past two decades, but there is still no consensus about the presence of cores in large clusters of galaxies [170, 106, 107].

Self-interacting dark matter (SIDM), in which DM has a nonzero self-interaction cross section, is another possible explanation for cored halos. SIDM has been discussed extensively in astro/particle physics literature as another possible model that can produce cores [182, 150, 195, 163, 80, 155, as key examples]. The basic mechanism of core formation in SIDM is that self-interactions between DM particles in regions with high density cause the DM to thermalize, which results in the transfer of energy to the high density inner region of a halo, thus lowering the core density. One frequently employed measure of self-interaction strength is the scattering cross-section per unit mass, σ/m, although that does not account for the dependence of scattering cross-section on particle velocity. Constant cross-section SIDM models cannot simultaneously have appreciable effects in dwarf galaxies, while being consistent with densities measured in clusters of galaxies [207, 155, 80]. A more complete model that accounts for the velocity-dependence of cross section is necessary because of the larger range of velocities probed in going from dwarf galaxies to clusters of galaxies (about 50 to 2000 km/s).

In this Chapter, we analyze 8 observed galaxy clusters and two simulated clusters in three distinct analysis stages: (1) strong lensing, to determine DM and baryonic density profiles, (2) SIDM halo matching, to determine SIDM cross sections and relative particle speeds for each cluster, and (3) constraints on the cross section, to finally infer the SIDM cross section and systematic error. Our analysis pipeline is shown schematically in Figure 3.1. In the first stage we extend the strong lensing analysis of Andrade et al. [3], in which one cluster (Abell 611) was examined, to include 8 relaxed clusters. The clusters we examine here are in the range of $4 \times 10^{14} M_{\odot}$ to $2 \times 10^{15} M_{\odot}$. These are the among largest bound structures in the universe [42]. Their centers have dense concentrations of DM, and if such particles self-interact, cores with densities less than that predicted by ΛCDM could form. Our aim is to put a stronger constraint on the SIDM self-interaction cross-section by inferring the inner distribution of DM in these 8 clusters, using strong lensing alone. In cluster strong lensing, images typically appear near the Einstein radii of the clusters, which usually range from 20 to 100 kpc, well within the scale radii of the clusters [156]. The DM density profile is constrained with highest accuracy near these image locations. This is also the region where SIDM thermalization would be expected to occur [80]. Baryonic effects are also strong in this region, as the brightest cluster galaxy (BCG) is located there, with high baryonic mass. It is therefore important to accurately characterize the BCG's contribution to the total mass profile in the region of interest. We use photometry to measure cluster member characteristics such as radius, position, ellipticity, orientation angle and luminosity, but leave the stellar mass-to-light ratio as a parameter to be fit for each cluster.

Other techniques that could be employed to determine cluster DM profiles include weak lensing, x-ray temperature analysis and stellar kinematics [115, 124, 169, 89, 167, 19, 103, 137, 131, 190, 34, 133, 132]. Strong lensing directly probes the profiles at the radii at which we are interested, while weak lensing and x-ray analysis probe much larger radii, where the ellipticity and profile may be significantly different than that of the inner region. Stellar kinematics of the BCG stars can help significantly in constraining the BCG mass and the slope of the dark matter density profile in the center [132, as an example]. This measurement depends on having the correct priors for radial variation of the stellar velocity dispersion anisotropy [170] in spherical or axisymmetric Jeans analysis.

He et al. [69] examined the ability of stellar kinematics, strong and weak lensing data sets to predict dark matter densities in the inner regions of simulated clusters, and found them to be accurate. This validates the approach undertaken by Newman et al. [133, 132] in combining these three data sets to constrain the dark matter density profile of clusters. While adding other data sets is well motivated, it is of interest to ask what constraints are possible with just one data set given that biases may creep in because of the assumption of a common density profile with a fixed ellipticity that is inherent in a joint analysis. Somewhat surprisingly, we find that it is possible to infer strong constraints on the SIDM cross section with just strong lensing data. In this paper, we develop the method to infer the SIDM cross section using strong lensing data, and leave the analysis including other data sets to future work. He et al. [69] also discussed possible ways to reconcile the differences between the inferred dark matter density profile slopes in Newman et al. [132] and their simulated clusters in CDM cosmology, and highlighted a possible role for bias resulting from degeneracy with r_s, which is constrained in part by the weak lensing data. We will discuss the inference of r_s from strong lensing data, and argue that our results are consistent with expectations from simulations.

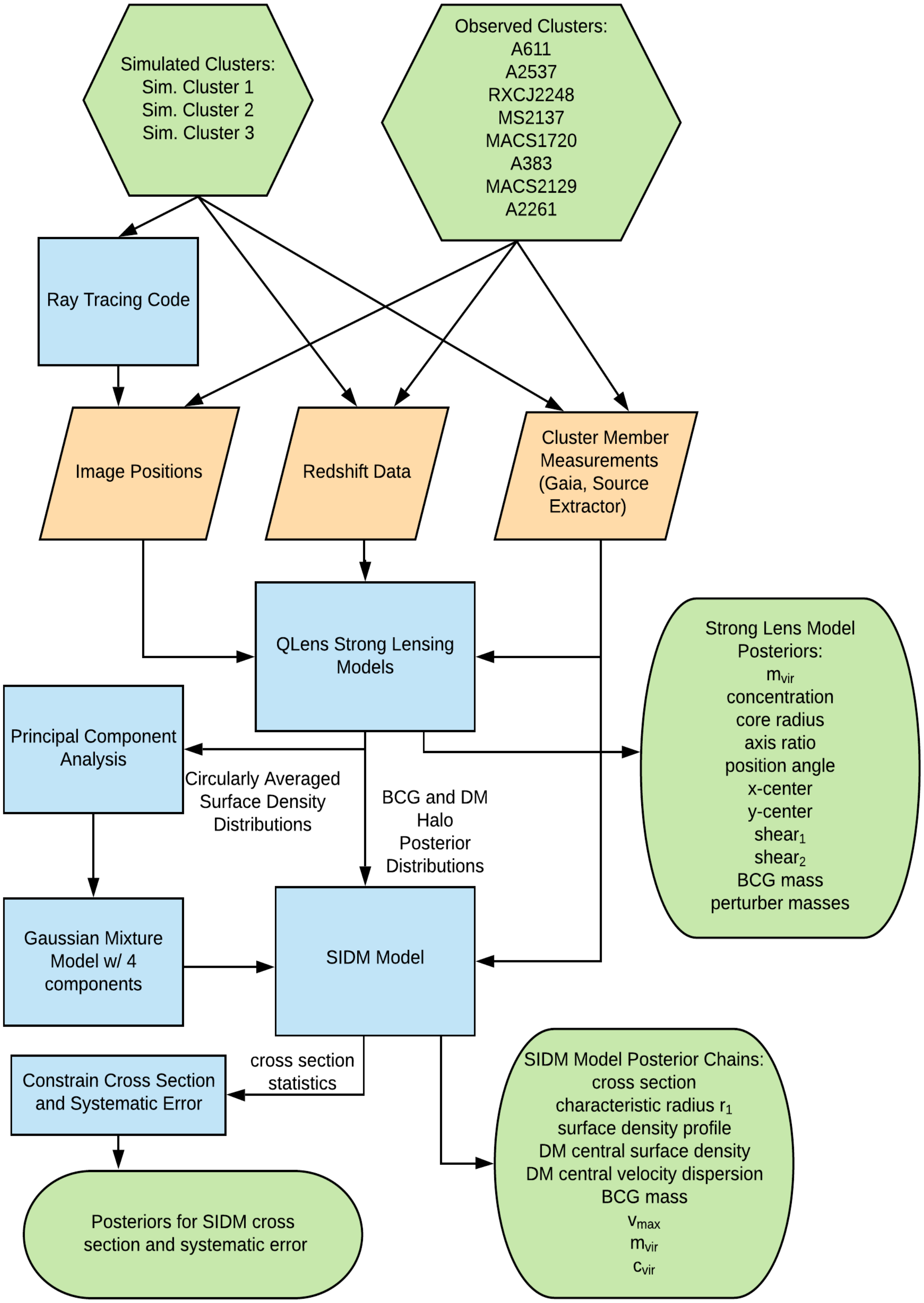
Armed with DM profiles and the aforementioned cluster member data, we employ a fitting process to match an NFW-like outer halo with an SIDM thermalized inner halo, following the process described in Kaplinghat et al. [80]. The resulting fit yields posterior chains for various parameters, including those for cross section per unit mass and average particle speed, for each cluster. In the third analysis stage, we synthesize the results of the the prior stages into an overall distribution for SIDM cross section to obtain an upper limit on self-interaction cross section.

This Chapter is organized as follows. In Section 3.2, we describe our approach to the strong lens modeling analysis. In Section 3.3 we detail the strong lens models for the simulated clusters, and do the same for the observed clusters in Section 3.4. The results of the strong lensing models for the observed clusters are presented in Section 3.5. In Section 3.6 we describe our SIDM halo matching model and its results, and also describe the model for putting constraints on the cross section, and its resulting inference of SIDM cross section and systematic error. We summarize our findings in Section 3.7.

3.2 Strong Lens Modeling Methods

Our first step is to use strong lens modeling to determine the surface density profiles for a sample of galaxy clusters. For the cluster lensing models, we use previously reduced data in the form of (1) coordinates for multiple images and (2) source redshift data. The availability of multiple image coordinate data was a requirement for all the clusters in our sample. The positional error is assumed to be $0''.5$ in each dimension (see Section 3.5.11 for a discussion of this choice). We construct mass models of the important features of each cluster: the (potentially cored) DM halo, BCG, important member galaxies ("perturbers"), and the DM subhalos associated with the perturbers. Many lensing analyses do not disentangle the DM subhalos from the luminous part, but doing so allows us to model the perturbers in a manner that is more consistent with cosmological simulations. The obvious challenge for creating accurate mass models is that the majority of the mass is in the form of invisible dark matter, the distribution of which must be inferred indirectly via lensing or by assuming some correlation with the luminous matter. We have designed a systematic approach that we first test on simulated galaxy clusters from the Illustris TNG simulation [128, 183, 144, 127, 123, 105] as described in section 3.3 of this work, allowing us to verify the accuracy of the models in reproducing cluster properties. We then apply this approach to actual galaxy clusters as described in section 3.4. For each cluster, we obtain a Markov Chain Monte Carlo (MCMC) chain for the model parameters, from which we can estimate parameter posteriors and also infer posterior distribution in other derived parameters. We use the `Multinest`

Figure 3.1: Schematic diagram of the analysis pipeline. Data sources are at the top. The final results are the parameter posteriors shown at the bottom.



sampler to create the MCMC chain [56].

In brief, we construct mass models for clusters in the following way.

- Starting with a CCD image of the cluster (or a simulated image in the case of Illustris clusters), the coordinates, shapes and fluxes of the BCG and key member galaxies are measured, using one or both of the Starlink Gaia software package [37] or a proprietary 2D fitting code. A coordinate system is created with its origin at the center of the BCG.
- The main DM halo is modeled using a cNFW profile, the parameters of which are mass, scale radius, core radius, position angle, axis ratio, and (x, y) center offset coordinates. All are allowed to vary within ranges defined by wide, uninformative priors.
- The stellar baryonic component of selected significant member galaxies (referred to as "perturbers") and the BCG are modeled using dPIE profiles. The mass parameter of each is allowed to vary (with perturbers typically tied together via a mass-follows-light scaling relation), while the scale radius, center coordinates, position angle and axis ratio are fixed by photometric measurement.
- For the perturbers, each is given an isothermal DM subhalo with a mass related to their baryonic mass via a power law. The scale radii of these DM subhalos are determined by their mass according to another power law, both described in Section 3.2.5. The center coordinates used are the same as those of the underlying baryonic model.
- To account for the effects of other masses along the line of sight that are not specifically included in the model, external shear is modeled, requiring two additional varying parameters, γ_1 and γ_2 (for a derivation of these, see Voigt and Bridle [197]).

3.2.1 Lens Profiles

Our goal in this analysis is to construct mass models of galaxy clusters that accurately infer the surface density profile in the inner ∼200 kpc, and subsequently infer SIDM properties. To model the point images and infer the halo density profile, we use the `QLens` software package, as in Andrade et al. [3]. [1]

We employ the cored NFW lens profile (cNFW) as the primary model for the cluster halos. The spherically symmetric form for the cNFW density profile is defined as:

$$\rho = \frac{\rho_s r_s^3}{(r_c + r)\,(r_s + r)^2}. \tag{3.1}$$

This profile reduces to the canonical NFW form for $r \gg r_c$. Other authors have used this profile to fit cored DM halos in both hydrodynamical simulations [143] and actual cluster lenses [133, 132, 3]. Analytic formulas for the projected density profile and deflection of the corresponding spherical model are given in Appendix B.1.

The BCG and member galaxy baryonic components, as well as galaxy member DM subhalos are modeled with a dual pseudo-isothermal ellipsoid (dPIE) profile [52], also known as a "Pseudo-Jaffe" profile. The spherical version of this profile is

$$\rho = \frac{\rho_{\mathrm{cut}} r_{\mathrm{cut}}^4}{\left(r^2 + r_{\mathrm{core}}^2\right)^2 \left(r^2 + r_{\mathrm{cut}}^2\right)^2}, \tag{3.2}$$

where ρ_{cut} is the density at the scale radius, r_{cut} is the scale (or tidal break) radius, and r_{core} is the core radius.

We model elliptical lens profiles by making the replacement $R^2 \to qx^2 + y^2/q$ in the projected density profiles. The deflection and Hessian of the lens mapping must be calculated by numerical integration (see Schramm 173, Keeton 85), which can be computationally expensive.

---

[1] For more information on `QLens` contact Quinn Minor: qminor@bmcc.cuny.edu.

However, by using Gauss-Patterson quadrature for integration [38], `QLens` is able to compute such integrals in a reasonable time. This approach avoids using the pseudo-elliptical approximation, in which the lensing *potential* is assumed to be elliptical rather than the mass profile. Using the pseudo-elliptical approximation can lead to inaccurate inferences of the density profile in cases of high ellipticity (e.g., $q < 0.5$, see Andrade et al. [3]).

3.2.2 Photometric Measurement

We use the Starlink Gaia software package [37], and the integrated version of Source Extractor [9] to measure the key photometric properties of the BCG and perterbers in each cluster: the coordinates, flux, axis ratio and position angle. We first select the HST image from the available object images that gives a target rest-frame wavelength in center of the visual band, based on the redshift of the lens. For background modeling, we use mesh-based RMS background detection. The projected half-light radii are measured using Source Extractor's "flux radius" feature by setting the flux fraction to 0.5. Fluxes are converted to magnitudes, and then corrected for galactic extinction using the tool at https://ned.ipac.caltech.edu.

3.2.3 Central DM Halo

The main DM halo is modeled using a cNFW profile, the parameters of which are mass, concentration, core radius, position angle, axis ratio, and (x, y) center offset coordinates. Flat log-priors were used on the mass and core radius parameters, while flat priors were used on position angle, axis ratio and center offset.

Merten et al. [114] studied the mass-concentration relation for cluster halos in 19 X-ray clusters from the CLASH sample. They observed an average concentration of 3.7, with a standard deviation of 0.65, and a negative correlation of concentration with halo mass. Since several of the clusters in this work are CLASH clusters, we might expect concentration to be in a similar range. However, we note that there are inherent biases in modeling concentration in strong lensing clusters, as outlined in Appendix B.3. Also, Fielder et al. [57] found that concentration will be higher for a DM halo when its associated subhalos are excluded from the calculation. Therefore we opt for a weak Gaussian prior on concentration, with a mean of 6.0 and standard deviation of 1.5. Thus the 2σ range is from 3.0 to 9.0, accommodating these anticipated biases if present.

3.2.4 BCG and Luminous Contribution of Member Galaxies

The BCG and perturbers for each cluster were modeled with a dPIE profile, with mass as a free parameter and other parameters set by observed photometry. The core radii of galaxies (r_{core}) in the observed clusters are difficult to measure with high certainty. For our purposes, we are not particularly interested in the core sizes of the member galaxies, as structure at that scale will not have a significant impact on our results. Other authors have assumed a constant core radius for cluster member galaxies, typically in the range of 100 pc to 300 pc [97, 99], or alternatively $0.1''$[98], which for our clusters is approximately 300 pc to 600 pc. We adopt a core radius of 300 pc in each member galaxy in our observed clusters. For the Illustris clusters, we fit their two dimensional shapes from the simulated image to a dPIE profile, and use those parameters in our `QLens` model. The measured core radii range from 0 kpc to 2.3 kpc.

We tie together the masses of most of the perturbers in a given cluster and employ a mass-follows-light approach for scaling them. This is a reasonable approach, given that we have separate mass models for the DM subhalos. The perturber masses scale with the observed luminosity of each perturber, resulting in one free mass parameter describing the set of perturbers. One perturber is selected as the anchor, and the others are tied to it. We do this in order to limit the proliferation of model parameters. However, there are some perturbers that are quite close to images and that strongly affect them, in some cases even causing splitting of images into multiple images. For these perturbers, their mass is allowed to vary separately from the group, at the cost of including another degree of freedom in the model.

3.2.5 DM Subhalos

In many strong lensing analyses, the baryonic and dark components of perturbers are modeled as one object, often based solely upon the photometry of the stellar component. This has potential for bias, since the DM component is likely to be strongly dominant in both mass and size. To guard against this, we model the DM subhalos of each perturbing galaxy.

We modeled the dark matter profile of cluster galaxies with a dPIE profile. We set the half-mass radius of the dPIE profile equal to that of a NFW density profile with the median concentration in Diemer and Joyce [45], as implemented in `Colossus` package [44]. Neglecting the mild redshift dependence in the concentration mass-relation for the relevant lens redshift range (about 0.2 to 0.6), we obtain the half-mass radius $= 0.0019 M_{\text{halo}}^{0.37}$. The core radii for the dPIE profiles are set to zero in the case of the observed clusters. For the simulated clusters, we find that the particle size of the Illustris simulation had an impact on the shape of the DM subhalos. In effect, this gives a core to the DM subhalos. The simulation has a DM particle size of $5.9 \times 10^7 M_\odot$, and it has been shown that two-body relaxation effects will give rise to cores at radii fewer than ~ 100 particles and at the densities typical in our model [148]. As the DM subhalos are typically in the range of 2×10^{11} to 6×10^{12} $M_\odot$, the 100-particle region has a typical radius 5 kpc to 10 kpc. In addition, we employ a Gaussian filter with a characteristic radius of 2 pixels (2.5 kpc) to help reduce noise in the ray tracing calculations, as noted in Section 3.3.2. To model these effects we allow the dPIE DM subhalos to have a core in the simulated cluster models. All subhalos in a given model have the same core, the radius of which is a varied parameter during the MCMC process.

We determine the mass of the subhalos from their fitted stellar mass, as follows. Infalling galaxies and their DM halos will be partially stripped of their DM by tidal forces due to the massive cluster. To quantify this for the simulated clusters, we looked directly at the the stellar-mass to halo-mass (SMHM) relation in Illustris TNG. We selected all subhalos within a 400 kpc radius of the center of the host halos, and with a stellar mass range similar to those of perturbers in the real cluster sample, i.e., between $1 \times 10^{10} M_{\odot}$ and $5 \times 10^{12} M_{\odot}$. These are well fit by a power relation of the form

$$M_{\text{halo}} = kM_{*}^{\alpha}, \tag{3.3}$$

with $k = 1.115$ and $\alpha = 1.024$. We used this relation to generate DM subhalo masses in the models of the simulated clusters. For the observed clusters, we turned to [134], where they examine the stellar mass-halo mass relation for cluster galaxies. We fit a power law to their data (see Figure B.1 in Appendix B.2). The resulting power law (in the same form as Equation 3.3) gives $k = 1.157$ and $\alpha = 1.1171$, which is similar to that found for Illustris. As can be seen by the error bars in the figure, this relation has significant uncertainty in the lower bound. We therefore use it as an upper bound only, and explore departures from this upper bound by varying the parameter "k". Ideally, we would fit this parameter with the other parameters through the MCMC sampling process, but unfortunately the `QLens` software does not have that capability. Instead, we manually optimize it by making runs with values of k, 0.5k, 0.25k and 0.125k. The value that yields the best model χ^2 is adopted. The value adopted is noted in the individual cluster models.

We find that in Illustris TNG, the position angles of dark matter halos of cluster galaxies are usually quite similar to that of the associated stellar mass, but the dark matter halos are rounder on average. We fix the position angle to be the same as that of the stellar spheroid and choose an axis ratio for the DM subhalos equal to $\sqrt{q_*}$, i.e., the geometric mean between the stellar axis ratio and unity. We ran tests on the effect of this assumption, comparing it to using $q = q_*$ and $q = 1$ for three representative clusters, A611, A2537 and MACS2129. The impact on the inferred BCG mass and halo mass was well within one standard deviation for all three clusters.

3.2.6 Use of Spectroscopic Redshift Data

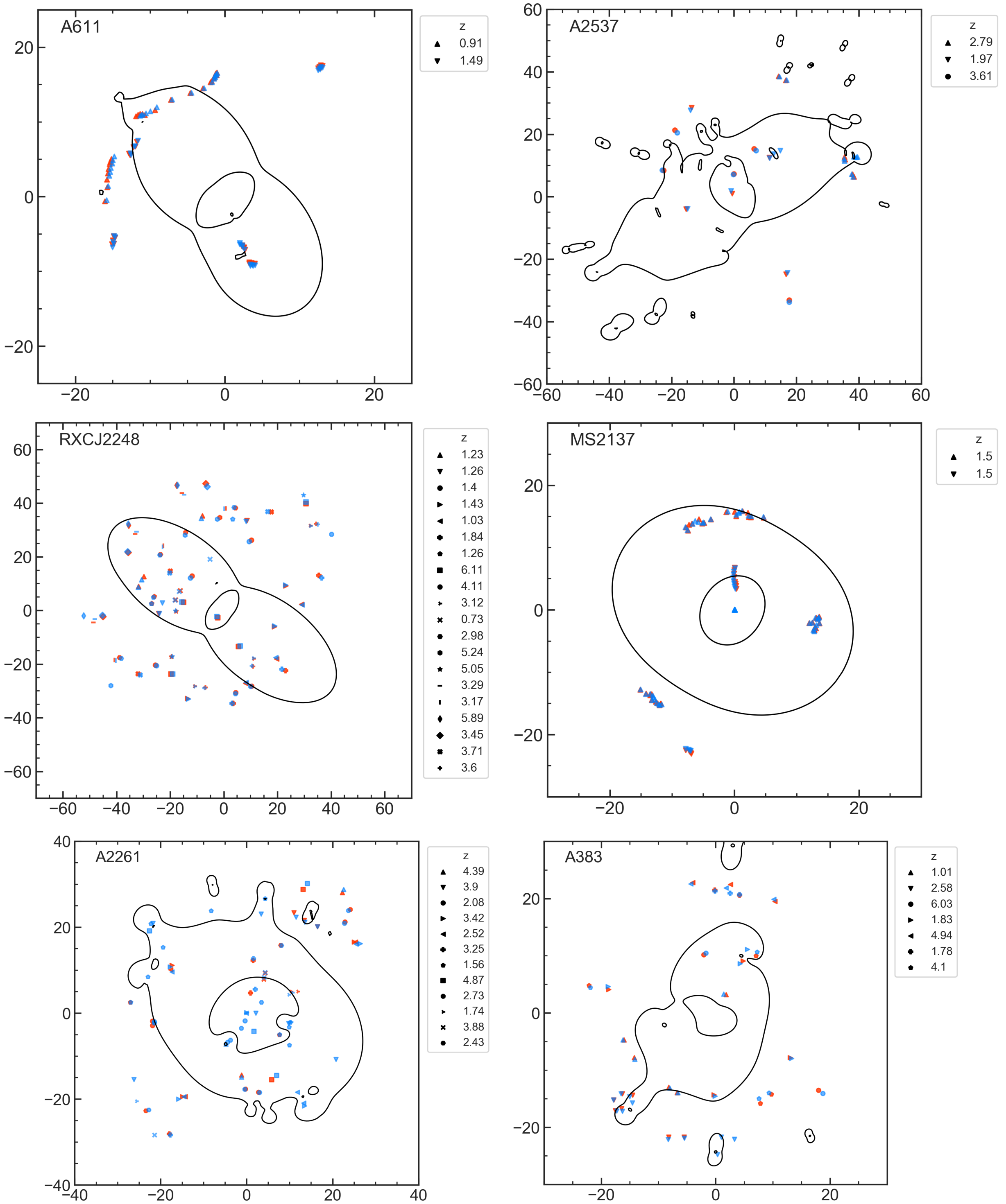
The data for each model is in the form of multiple image point locations, together with the redshift of each. Both spectroscopic and photometric redshift data are used, as noted in for each individual cluster. Where photometric data with significant confidence intervals was used, we adopted a redshift z_a for each source that minimized $\chi^2 = \sum_{i=1}^{n} \frac{(z_i - z_a)^2}{\sigma_i^2}$,

where n is is the number of images in the image group, z_i is the quoted median redshift of the image and σ_i is the 68% uncertainty interval for the measured redshift. In the case that the error in redshift is not given as a single value but in the form $z^{+\sigma \mathrm{hi}}_{-\sigma \mathrm{low}}$, we use σhi as the error when the trial value is above the median and σlow when it is below.

3.2.7 Use of Central Images

Gravitationally lensed objects can produce positive-parity "central" images that are typically demagnified and near the center of the lens, and as a result are often obscured by bright objects in the central region of the field. When they are visible, they are helpful in constraining the lens parameters. Central images occur in five of the eight clusters in our sample: A2537, RXCJ2248, MS2137, A2261 and MACS1720. The `QLens` software can be configured to look for central images and produce model central images for a given cluster, but the setting is the same for all sources in the cluster, even if only one source has a central

Figure 3.2: Image plane plots for each cluster showing the data image positions (red) and modeling image positions (cyan). The points appear purple where they overlap. Critical curves are shown for a source redshift of 2.0. Redshift values for each family of points are indicated in the box to the right of each plot. North is up, East is to the left. Axes scales are in arc seconds.



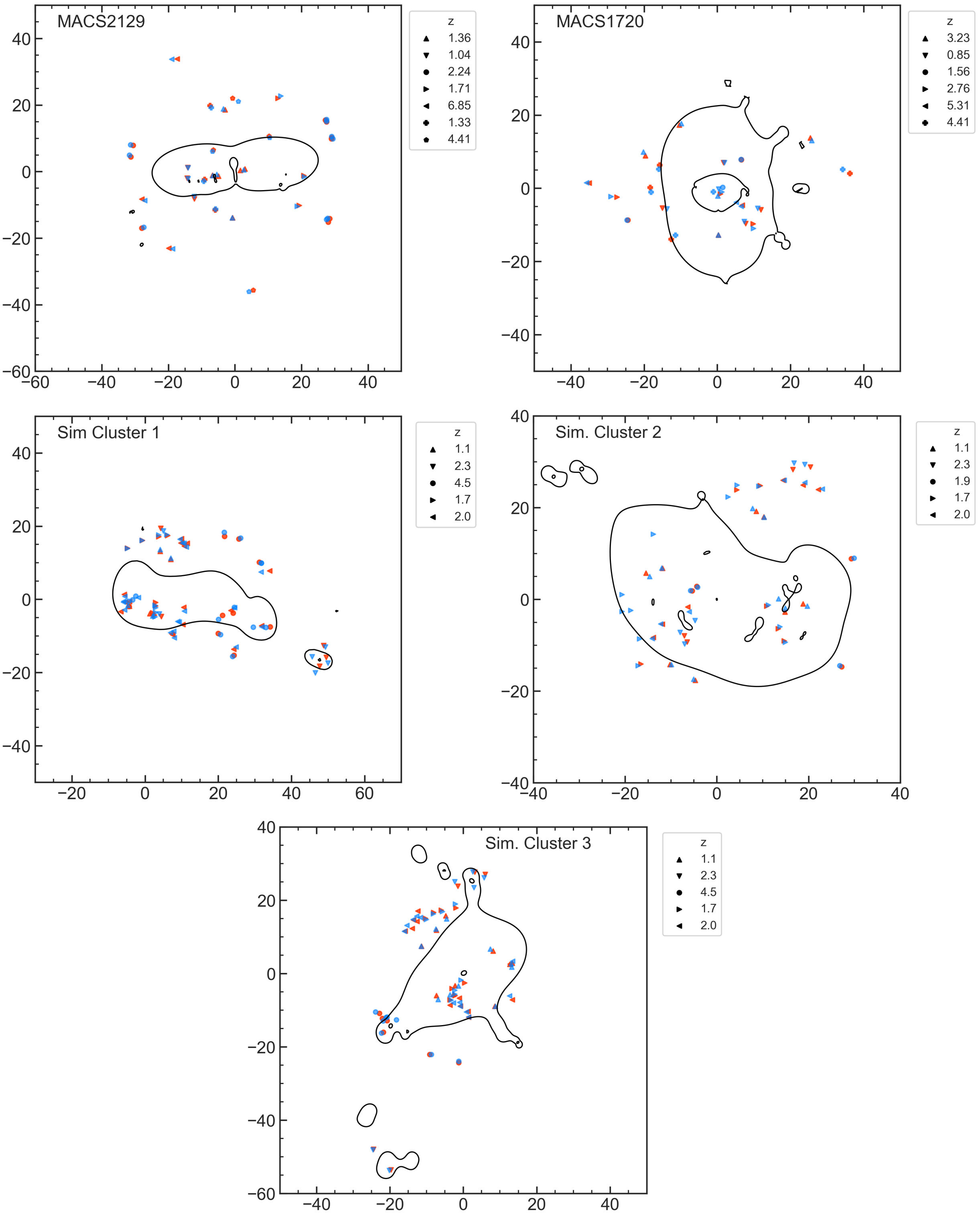


Figure 3.2: , continued.

Table 3.1: Key cluster halo parameters of simulated lensing clusters from Illustris and fitted values from strong lensing models. All simulated clusters were given a redshift of 0.3. The Illustris parameters are from 3D fits using Colossus. We were not able to obtain a meaningful strong lensng fit with Cluster 1 due to its unrelaxed nature. The strong lensing fit m_{200} and concentration values are for the main DM halo only and do not include that in subhalos. RMS position error is for the x and y coordinates, combined in quadrature.

		Illustris 3D Fit			Strong Lensing Fit				
Cluster Name	Relaxed?	ID	m_{200} ($10^{14}M_\odot$)	Concentration	m_{200} ($10^{14}M_\odot$)	Concentration	Image Points	Source Points	RMS Position Error (arc sec)
Cluster 1	no	5941	7.08	2.96	N/A	N/A	42	16	1.01
Cluster 2	yes	10359	4.79	5.87	$4.79^{+0.93}_{-0.72}$	$8.08^{+0.94}_{-0.84}$	28	12	0.69
Cluster 3	yes	19512	3.09	5.88	$4.77^{+0.83}_{-0.68}$	$7.85^{+0.89}_{-0.81}$	41	16	0.88

image. As a result, for the five clusters with central images, our models produce some images in the central region that do not have a corresponding "observed" match in the data. These predicted images might be present but difficult to observe due to their inherently low magnification and/or proximity to the bright objects usually present in the center of clusters.

3.3 Models of Simulated Clusters

To test the validity of our methods, we use data from the IllustrisTNG simulation [128, 183, 144, 127, 123, 105] to simulate strong lensing in three massive TNG galaxy clusters. Their characteristics are summarized in Table 3.1.

3.3.1 Selection of Simulated Cluster Halos

To find analogues of our observed cluster sample, we used the largest volume simulation available at a box length of 300 Mpc, in the highest resolution for that suite of simulations, TNG300-1. While the large box size comes at the expense of resolution for DM particle mass ($5.9 \times 10^7 M_\odot$), the TNG300-1 suite are the best choice among the TNG simulations for studying large (cluster) scales. We used the gravo-magnetohydrodynamical suite of simulations which includes baryons, to account for the effects of baryonic matter on the DM in these halos.

To match the general range of redshifts in our observed cluster sample, we selected three of the most massive host halos as our simulated clusters from the TNG300-1 snapshot corresponding to a redshift of z = 0.3. All three clusters have masses $10^{14} M_{\odot} < m_{200} < 10^{15} M_{\odot}$. For each, the bright galaxy at the center of the potential in the central subhalo was identified as the BCG. Of the three simulated clusters, two appear relaxed and have a clearly dominant BCG, while one (Sim. Cluster 1) appears to be merging, as evidenced by a second major galaxy nearly comparable in size to the BCG and located only ∼120 kpc away from it. We chose to keep this cluster in our analysis to examine its impact on our inferences. The remaining two hosts have only one BCG.

Cluster strong lensing is subject to selection bias dependent upon the line of sight (LOS), because mass concentrations along specific LOS can increase the surface density and lensing strength for subhalos that depart from spherical symmetry [33, 70, 136, 176, and see also Appendix B.3 for a discussion of concentration bias]. To simulate this effect and to ensure strong lensing occurs, we shot 10 random lines of sight through each halo and used the LOS with largest central surface density. We then obtained 3D fits of each simulated halo, using the `Colossus` software package [44]. The resulting fitted values for m_{200} and concentration, are listed in Table 3.1. Note that the LOS that produced the highest central surface density for a given cluster halo also yielded the highest concentration for that cluster halo, suggesting our simulated cluster analogs are affected by selection bias.

3.3.2 Simulated Image Production

After choosing the LOS, we then created a surface density map of the stars, gas, and DM, at a resolution of 1.25 kpc per pixel. We noticed that the surface density map exhibited significant

Poisson noise due to the finite particle size of the Illustris simulation. To partially ameliorate the noise, the map was smoothed with a Gaussian kernel, with a 2-pixel characteristic radius. We chose a 2-pixel smoothing radius as a way help reduce the statistical noise arising from the finite particle size but without substantially altering DM structure at scales relevant to our analysis. From the surface density map we calculate the scaled lensing potential:

$$\psi(\vec{\theta}) = \frac{1}{\pi} \int d^2\theta' \kappa(\vec{\theta'}) \ln |\vec{\theta} - \vec{\theta'}|, \tag{3.4}$$

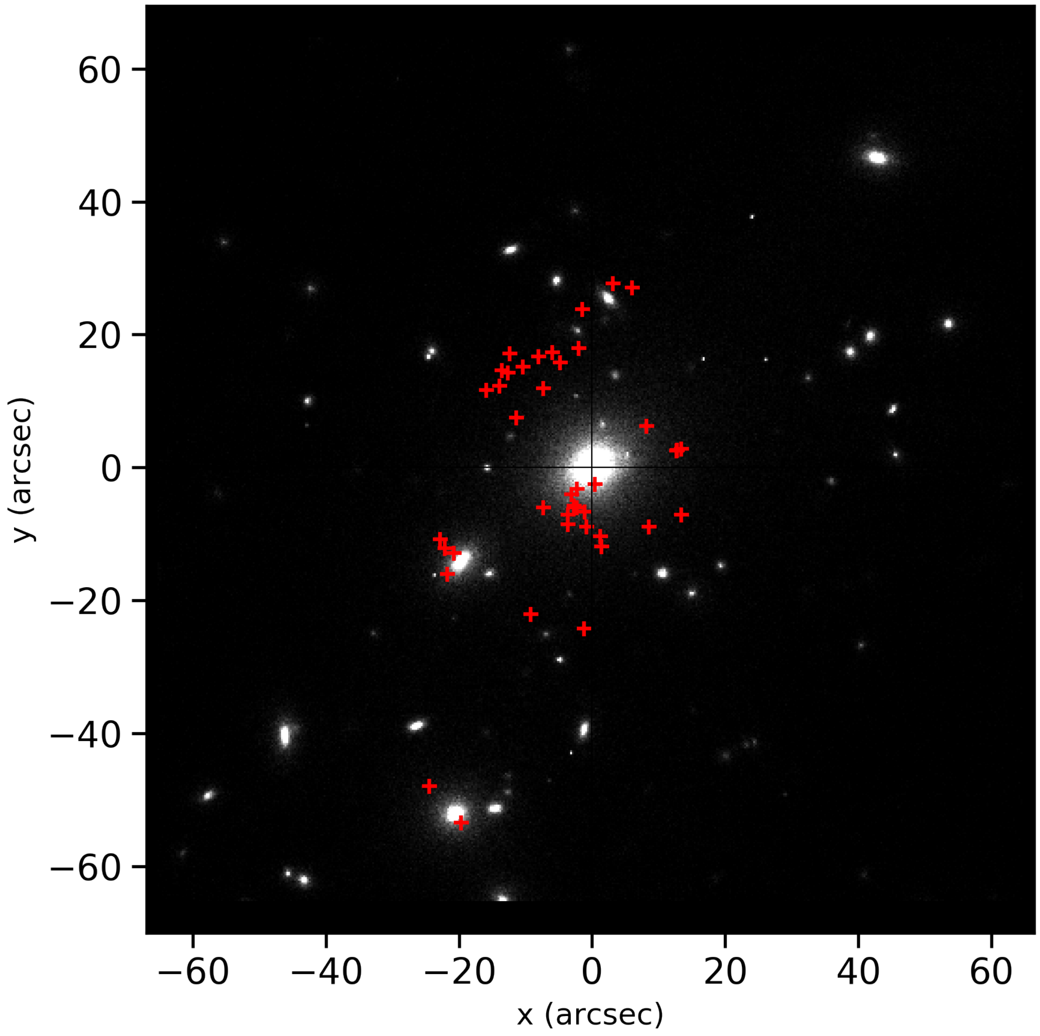
where ψ is the scaled potential, $\vec{\theta}$ is the deflection angle, and κ is the convergence. The reduced deflection angle $\vec{\alpha}$ can be found by taking the gradient of the scaled potential, i.e., $\vec{\alpha} = \vec{\nabla}_\theta \, \psi = D_L \vec{\nabla} \, \psi$, where D_L is the distance from the observer to the lens. This, together with the lens equation $\vec{\alpha} = \vec{\beta} - \vec{\theta}$, allows us to solve for image positions for any given angular position $\vec{\beta}$ of a source object.

For each cluster, from 12 to 16 point sources were created, each generating 3 or 5 images, resulting in 28 to 41 images. There were 5 unique redshift groups for each cluster, ranging from $z = 1.1$ to $z = 4.5$, which represent a typical range of source redshifts found in actual clusters. Random Gaussian errors with a standard deviation of $0\rlap{.}''5$ were added to each x and y coordinate of the images, resulting in a mean position error of $0\rlap{.}''71$. An example of a mock image for Simulated Cluster 3 is shown in Figure 3.3.

3.3.3 Model Details for Simulated Clusters

The models for each of the three simulated clusters are discussed below. Key model results are summarized in Table 3.1. Image plane plots are shown in Figure 3.2, showing data image locations, modeled image locations and critical curves.

Figure 3.3: An example image of a simulated cluster (Sim. Cluster 3). The BCG is centered, and the red markers indicate 41 simulated point images from 16 source points.



3.3.3.1 Cluster 1

Cluster 1 appears to be a merging cluster, with two large central galaxies separated by approximately 120 kpc. The unrelaxed DM profile proved problematic to model. The data set had 16 source points and 42 images, and we identified 9 perturbers in addition to the two large central galaxies. We attempted several approaches; treating the smaller of the two central galaxies as an ordinary perturber did not work well, nor did using two blended cNFW halos. While we were able to reproduce many of the simulated images, there is not enough freedom in the model to accurately reproduce the DM distribution, which is crucial for accurate inference of SIDM cross section. The modeled images had a root-mean-square (RMS) position error of $1''.01$ (x and y combined).

3.3.3.2 Cluster 2

Cluster 2 is a relaxed cluster, and the simulated data set had 12 source points and 28 images. From the created HST-like image, we determined the photometric parameters (luminosity, center coordinates, dPIE radii, position angles and axis ratios) of the BCG and perturbers. We identified 13 perturbing galaxies, choosing those ≤ 60 arcsec of the BCG and with a mass $\geq 10^{10} M_{\odot}$. All but one of the perturbers were "anchored" together so that their masses scaled as one group. We individually optimized one perturber that was close to several images. All data images were reproduced by the model, however there were 3 extra images. The modeled images had an RMS position error of $0''.69$. The surface density was reproduced with an mean precision of 0.048 dex over the radius range where images are located (i.e., from 23 kpc to 157 kpc from the BCG).

3.3.3.3 Cluster 3

Cluster 3 is a relaxed cluster with 16 source points and 41 images in the simulated data set. Using the created HST-like image, we reduced photometric measurements of perturbers and selected 10 perturbers, based on their proximity to the center and luminosity. Of these, 9 were anchored together so that their mass was varied as one, and one perturber was optimized individually. To achieve better fitting and image reproduction, we reduced the perturbers' DM halo mass normalization by 50%. All data images were reproduced by the model, with three extra images. The RMS position error of the modeled images was $0''.88$. The surface density was reproduced with an mean precision of 0.04 dex over the radius range where images are located (i.e., from 11 kpc to 254 kpc from the BCG).

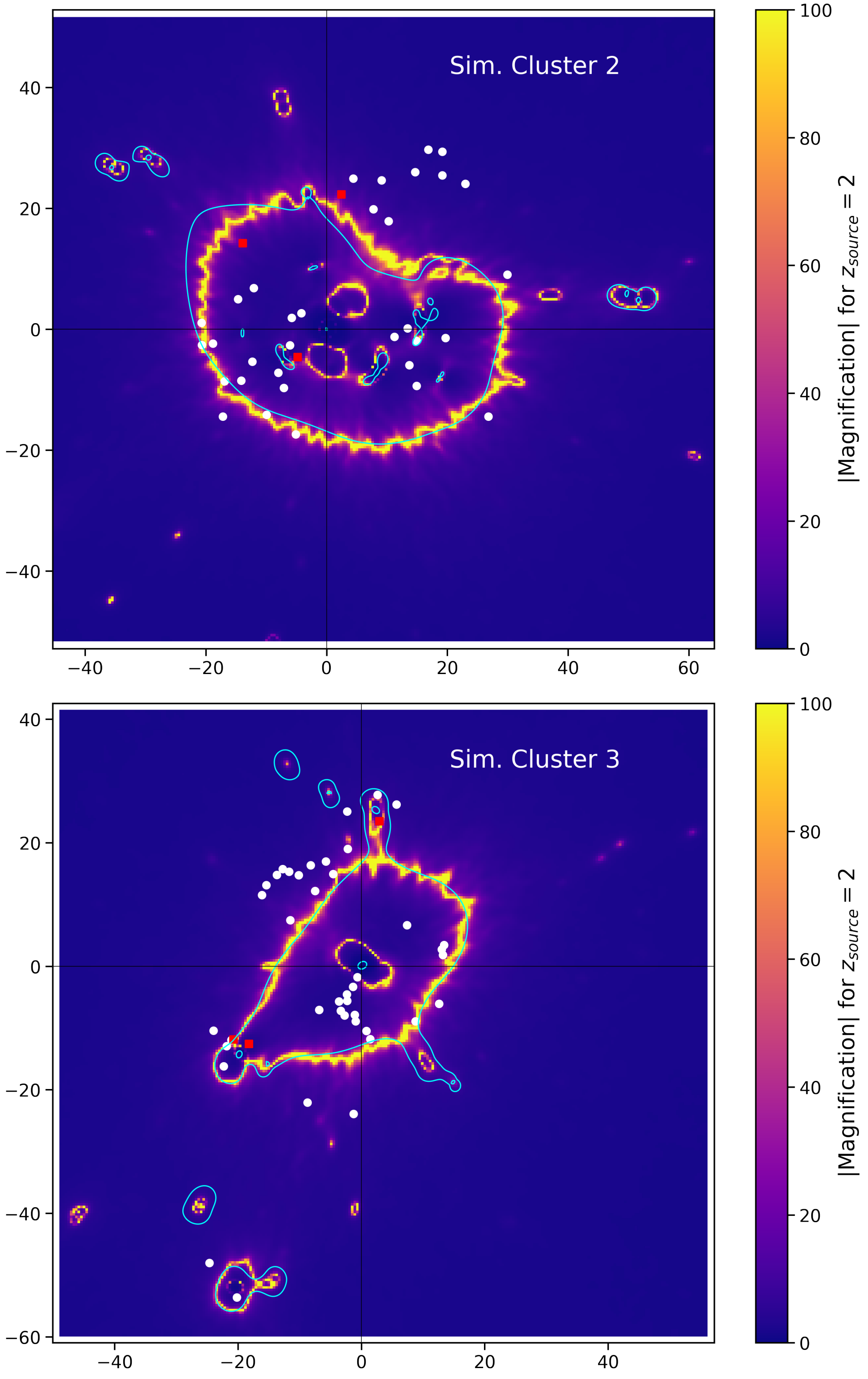
3.3.4 Discussion

For Sim. Cluster 2, the inferred halo mass from strong lensing closely matches the value obtained from measuring binned data from the Illustris surface density projection, whereas for Sim. Cluster 3, the inferred mass is 35% lower than the measured value. It should not be expected that the masses inferred from this process would closely match the actual values, because the strong lensing fits only the inner 200 kpc or so, whereas the r_{200} of these halos is more than 1.5 Mpc.

We measured the size of the BCGs with a 2-D fitting code, assuming an elliptical dPIE model (see Section 3.2.1). The measured scale radii values were quite small and compact, with scale radii of 0.5 kpc and 2 kpc respectively for Sim. Clusters 2 and 3. We note that these values are much smaller than the half-light radii quoted in the Illustris catalog, 66 kpc and 76 kpc, respectively. This difference could be due to the extended light distribution in the cluster. As a check, we reran the strong lensing models using the larger BCG radii. In each case, there were modest differences in some parameter posteriors, but in all cases the preferred halo core size was very small, i.e., consistent with no core.

The median inferred concentrations were 30% to 40% higher than those found from the 3D fit (although within 2σ). This is consistent with expectation, since the line of sight to the clusters was purposefully chosen to maximize strong lensing, making it likely that the line of sight is preferentially oriented near the major axis of the halo [33, 70, 136, 176]. The bias of concentration in such cases is explored in more detail in Appendix B.3. Another plausible contributing factor is that the concentration will be higher for a DM halo when mass from its associated subhalos are excluded, as discussed in Fielder et al. [57]. This is at least partially true in our models, as we model the subhalos of the largest perturbers individually.

Figure 3.4: Magnification maps for the simulated clusters. Regions of high magnification appear as yellow bands. The noise in the bands is due to the finite particle size of the simulation. The critical curves found from strong lensing are overlaid as cyan lines. Model image locations are shown as white circles for images matched to data, and red squares for extra images. Axes scales are in arc seconds.



Three extra modeled images appeared in each of the relaxed cluster models. Extra images can appear if the source point is close to a caustic curve, or equivalently, if images are close to critical curves in the image plane. To check on this, we compared the critical curves generated by `QLens` to the curves of high magnification as generated by the ray tracing code. This is a meaningful comparison because magnification becomes infinite at the critical curves. Figure 3.4 show those comparisons for Sim. Clusters 2 and 3. The extra images are indeed close to critical curves, and the predicted critical curves closely follow the high magnification curves. As can be seen in the figures, the high magnification curves are noisy; this is due to the finite particle size of the Illustris simulation. The ray tracing grid has a pixel size of 1.25 kpc. The critical surface density (where $\kappa = 1$, and where the tangential critical curves are located) at a lens redshift of 0.3 and assuming a source redshift of 2.0 is $\Sigma_{\text{crit}} = 2.352 \times 10^9 M_\odot$ /kpc^2. Since the mass of a DM particle in the Illustris simulation is $5.9 \times 10^7 M_\odot$, there are an average of only ~ 60 particles per pixel, the statistical variation of which can be seen in the magnification curves. By making changes to grid settings in the `QLens` software, we could reduce the number of extra images somewhat, but at the cost of higher RMS image position error in the model, which would be undesirable. We checked that the halo and BCG parameters do not change significantly under such changes. We interpret the extra images as due to either artifacts of the modeling process or unmodeled substructure, and do not believe that they significantly impact the inference of the model parameters.

3.3.5 Summary of Key Findings from Simulated Clusters

We used ray tracing to simulate strongly lensed images for three Illustris clusters, and tested our strong lensing analysis pipeline on those clusters. Our key findings for those models are summarized here.

- Although the lens models constrain only the inner ~ 200 kpc of the DM halo, they were nevertheless able to recover the halo mass of their corresponding simulated clusters to within 20% to 30%.

- Because of the finite particle size of the simulation, the surface density map exhibits significant Poisson noise. By smoothing the map with a Gaussian kernel with 2-pixel characteristic radius the noise was reduced. The simulation's finite particle size also causes the DM subhalos to suffer from two-body relaxation effects in their centers, effectively giving them cores. We therefore modeled those halos using a cored profile. Doing so moderately improved the accuracy of the image reproduction for Sim. Cluster 2, but yielded only a minor improvement for Sim. Cluster 3.

- The critical curves produced by lens models followed those of the magnification map from the underlying potential, save for noise in the magnification map due to the simulation's finite particle size.

- Image positions were recovered with good accuracy, commensurate with the simulated position error added to the mock data. The total simulated position error was $0''.71$ ($0''.5$ each in x and y), while the RMS position error of the models were $0''.69$ and $0''.88$ for Sim. Clusters 2 and 3, respectively.

- Three extra images were produced in each of the simulated cluster models, but did not appear to significantly impact the parameter inference of the models.

- The merging cluster, Sim. Cluster 1, proved difficult to model, as the mass distribution is complex and is not well characterized by a potentially cored NFW halo with additional perturbing bodies. We were not able to recover meaningful parameter values for that cluster. This highlights the importance of choosing clusters that are not major mergers for our analysis. We have been cognizant of this restriction in choosing the sample of observed clusters to model.

Table 3.2: Summary of the observed cluster sample.

Cluster Name	Adopted Redshift	kpc per arcsec	t_{age} (Gyr)	Relaxed?	Image and Galaxy Data Sources	Image Points	Source Points
Abell 611	0.288	4.33	10.2	yes	Donnarumma et al. [47], Newman et al. [131]	49	13
Abell 2537	0.294	4.39	10.1	yes*	Newman et al. [133], Cerny et al. [26]; this work	16	4
RXC J2248.7-4431	0.348	4.92	9.7	yes	Bonamigo et al. [12]; this work	55	20
MS 2137.3-2353	0.314	4.60	9.9	yes	Donnarumma et al. [46]; Newman et al. [133];this work	50	14
Abell 383	0.189	3.16	11.1	yes	Monna et al. [116]; this work	27	9
Abell 2261	0.225	3.61	10.8	yes*	Coe et al. [34]; this work	30	12
MACS 2129.4-0741	0.589	6.63	7.8	yes*	Monna et al. [117]; this work	31	9*
MACS 1720.3+3536	0.387	5.27	9.3	yes	Zitrin et al. [211]; this work	19	6

* but see discussion in the relevant paragraph of Section 3.4

3.4 Models of Observed Clusters

We began by selecting a subset of galaxy clusters, based on the following criteria:

1. a generally relaxed shape, with a dominant main halo and the absence of vigorous current merger activity, based on either X-ray or optical observation;
2. the presence of strong lensing, resulting in several multiple-image systems
3. existing multiple image data, including redshifts,
4. ideally, the availability of galaxy member data, to determine likely lensing perturbers, although in many cases we reduced our own perturber information from HST data.

We targeted a sample size of approximately 8 to 10 halos, so as to enable us to draw cogent general conclusions from the sample. The resulting sample set of 8 clusters is listed in Table 3.2.

Figure 3.2 shows plots of the image plane for the cluster models, showing the data image positions, their modeled counterparts, and representative critical curves for a redshift of 2.0.

3.4.1 Abell 611

In Andrade et al. [3] this cluster was modeled in detail, using image position and galaxy member data from Donnarumma et al. [47], but with redshift corrections indicated in Newman et al. [133], Belli et al. [8]. The redshift data is spectroscopic. In Andrade et al. [3], both cNFW and Corecusp halo models were explored, which gave similar results, with the cNFW model being the preferred model.

In this work, our approach is similar, except that we model DM halos for each of the perturbers, as described in Section 3.2.5. The BCG, the seven perturbing galaxy members and their associated DM halos were modeled with dPIE profiles. The BCG mass parameter was a varied parameter, while the other BCG parameters were fixed based on the photometry from Newman et al. [133]. The mass parameters for the perturbing members were varied in three groups, one each for perturbers 1 and 2, and one group for perturbers 3 through 7. The normalization of the stellar-halo mass relation (i.e., parameter "k" in Equation 3.3) was reduced by 75%, to allow for the lowest χ^2 while still matching all data and model images. The resulting model reproduced all data images, with no extra images.

3.4.2 Abell 2537

Abell 2537 is an efficient gravitational lens that has been studied by several others, including Newman et al. [133, 132], Cerny et al. [26]. It appears relaxed and uniform in X-ray images [172]. Newman et al. [133] describes the cluster as likely disturbed, perhaps along the line-of-sight, but we nevertheless were able to construct a satisfactory model that explains the image positions with reasonably good fidelity. The stellar-halo mass normalization "k" was reduced to 12.5% of its nominal value, to minimize the image position χ^2.

Image positions, spectroscopic redshift and BCG photometry data from Newman et al. [133]

were used in the model. Perturbing galaxy data from Cerny et al. [26] was also utilized. Only those perturbers with V-band luminosity greater than $5 \times 10^{10}\ L_\odot$ and located less than $60''$ from the BCG were included in the model, resulting in 32 perturbers. The BCG mass and one anchor galaxy mass (to which the other 31 perturbers are anchored) were varied parameters. There were three extra non-central images in the model. Those images are in the vicinity of a bright perturbers and could be present but washed out in the perturber light.

3.4.3 RXC J2248.7-4431

Also known as Abell 1063S, this cluster has been previously modeled in some detail by Caminha et al. [22] and Bonamigo et al. [12]. It appears relaxed and uniform in X-ray images [172]. We used data from Bonamigo et al. [12] for the image positions, redshifts and perturbers. We included only those perturbers with V-band luminosity greater than $2 \times 10^{10}\ L_\odot$ and a distance from the BGC less than $60''$, resulting in 13 perturbers. The mass of the BCG and one anchor galaxy (to which the other 12 perturbers were anchored) were varied. The stellar-halo mass normalization "k" was not reduced from its nominal value. The best-fit models for this cluster produced 7 extra images. Examination of the HST image [147] revealed what may be image candidates at most of the predicted locations, but without spectroscopic data we cannot confirm them.

3.4.4 MS 2137.3-2353

MS2137 appears very relaxed and uniform in X-ray images [172]. We adopt the image positions and redshift data from Donnarumma et al. [46]. The point image locations follow two great arcs at nearly identical spectroscopic redshifts. BCG photometry from Newman et al. [133] was used. There is one prominent perturber, the position and photometry data of which we reduced using HST data [147]. The mass of the BCG and perturber were allowed to vary, while the other parameters were fixed, based on photometric measurements. The stellar-halo mass normalization parameter "k" was reduced by 50% to achieve better fitting and image reproduction. All data images were reproduced, with no extras.

3.4.5 Abell 383

Abell 383 is a relaxed cluster with 9 lensed sources for a total of 27 images. Following Monna et al. [116], we also adopt 19 perturbers for this cluster. For the BCG and all other cluster members, the photometric parameters were reduced from the HST F814W filter [147] using SourceExtractor provided in the Gaia software package. As in Monna et al. [116] we use the GR galaxy as our reference galaxy and anchor all other perturber masses to the GR galaxy for optimization. Due to localized radial arcs near the G1 and G2 perturbers, we individually optimize these perturbers in addition to the BCG and GR. The stellar-halo mass normalization "k" was reduced to 25% of its nominal value, to minimize the image position χ^2. The model reproduced all data images but yielded 6 extra images. Examination of the HST image for potential new images at their predicted locations was inconclusive due to noise in the image.

3.4.6 Abell 2261

Abell 2261 is a borderline relaxed cluster with 12 strongly lensed sources for a total 30 images (see Table 3 of Coe et al. [34]). Our search of past work on this cluster did not yield data on potential perturber membership in the cluster. We therefore adopted 12 perturbers by choosing the galaxies which distort critical curves in Figure 1 of Coe et al. [34]. We measured effective radius, axis ratio, luminosity and position angle from the HST F775W filter [147] using SourceExtractor provided in the Gaia software package for the BCG and perturbers.

We individually optimized 4 perturbers which are located in close proximity to images; the others were grouped and optimized as one. The stellar-halo mass normalization "k" was reduced to 50% of its nominal value, to minimize the image position χ^2. The model was able to match all data images with an RMS position error of $0\farcs83$, however 8 extra non-central images with |magnification|>1 were produced.

3.4.7 MACS 2129.4-0741

MACS2129 is described by Mann and Ebeling [104] as a recent but well separated merger, although it appears relaxed. It has 8 lensed sources and a total of 31 images according to Monna et al. [117] (see Table 3 there). However, system 4 is much better characterized as a system of two images being lensed rather than one image being lensed, thus we use 9 sources as in Monna et al. [117], where they also make a distinction between the two images in system 4. We adopt 10 perturbing galaxies, two of which are clear strong lensing sources that impact image positions significantly. The masses of those two were optimized individually. The masses of the other perturbers were anchored to a reference galaxy, the mass of which was varied and optimized. The stellar-halo mass normalization "k" was reduced to 12.5% of its nominal value, to minimize the image position χ^2. The model reproduced 30 out of 31 data images, with no extra images. The model matched data images with an RMS position error of $0\farcs79$. The radii, position angle, and axis ratio of all cluster members were measured from the HST F814W filter [50].

3.4.8 MACS 1720.3+3536

MACS1720 also appears relaxed and uniform in X-ray images [172]. We use the image position and photometric redshift data of Zitrin et al. [212]. A single redshift value for each image group was selected by the photo-z optimization process described in Section 3.2.6.

Image set 7 from Zitrin et al. [212], a 3-image set in the far southern part of the image, was problematic to model. It contains a pair of images within close proximity to one another, but no apparent perturbing bodies nearby. In addition, the photometric redshift of this source has a high uncertainty. We therefore excluded that source point. Interestingly, we were able to accurately produce the three images in Image set 7 if we included a perturbing subhalo of mass $\sim 2 \times 10^{12} M_{\odot}$ at coordinates of (12.5, -25.5) arcsec relative to the BCG. However, inspection of the HST images does not reveal any significant luminous body at that location.

BCG photometry and perturber data were reduced from HST images [147]. We included only those perturbers with F814W luminosity greater than $5 \times 10^{9}\ L_{\odot}$ and a distance from the BGC less than $30''$, resulting in 13 perturbers. The masses of these 13 were varied as a group. The stellar-halo mass normalization "k" was not reduced from its nominal value. The resulting model had one extra image with |magnification| $>$ 1.0, a magnification 2.7 image located at (4.2, -3.5) arcsec relative to the BCG. Examination of the HST reveals a good candidate object near that location, although we cannot confirm it absent redshift measurement.

3.5 Strong Lensing Results

Using the methods described in Section 3.2, we constructed mass models for each cluster that reproduced the observed image positions. We used separate lens elements for the baryonic and dark mass components of member galaxies, which is not a common approach in cluster lensing analysis. The strong lensing models were able to reproduce the image position data with good accuracy. We assumed measurement error of $0\overset{''}{.}5$ in each of the x and y coordinates of the data image positions. The models recovered the data image positions with root-mean-square position errors (combining the x and y components) ranging from $0\overset{''}{.}32$ to $1\overset{''}{.}07$, with a median of $0\overset{''}{.}66$, as shown in Table 3.3. The reduced image plane χ^2 (i.e., χ^2 per degree

Table 3.3: Strong lensing results summary. Where confidence intervals are given, the 14th- and 86th-percentile values are indicated. The parameter r_c is the cNFW core radius. The BCG luminosity is measured using the closest HST filter to the rest frame wavelength, so as to be approximately comparable to V-band. Measurement uncertainties in luminosity are <2%. The halo offset from the BCG is the inferred projected offset in the plane of the sky.

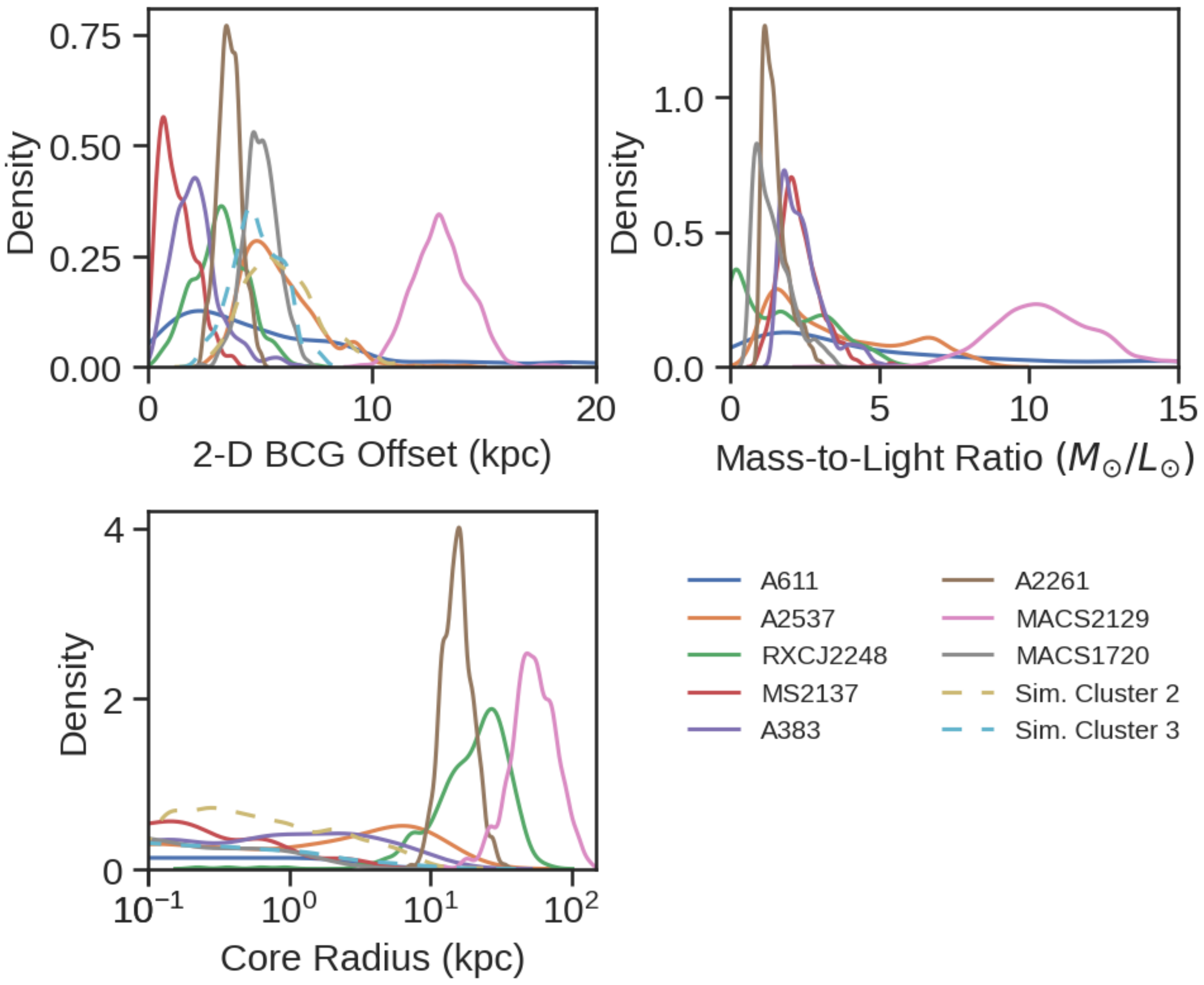
Cluster Name	Img. Plane χ^2[red.]	RMS Pos. Error (arcsec)	r_c (kpc)	Halo Mass ($10^{14} M_\odot$)	Halo Concen-tration	BCG Mass ($10^{12} M_\odot$)	BCG Lum. ($10^{11} L_\odot$)	2-D Halo BCG M/L ($M_\odot$ /$L_\odot$)	Offset from BCG (kpc)
Abell 611	20.7 [0.36]	0.32	$0.52^{+6.48}_{-0.48}$	$8.77^{+3.34}_{-2.30}$	$6.20^{+1.17}_{-1.05}$	$1.21^{+2.74}_{-0.88}$	3.25	$3.72^{+8.43}_{-2.71}$	$4.18^{+5.75}_{-2.80}$
Abell 2537	22.0 [1.57]	0.59	$1.18^{+6.76}_{-1.12}$	$9.42^{+1.77}_{-1.42}$	$6.77^{+0.90}_{-0.76}$	$0.93^{+1.09}_{-0.48}$	3.32	$2.79^{+3.28}_{-1.43}$	$5.44^{+1.18}_{-1.10}$
RXC J2248.7-4431	116.8 [1.98]	0.73	$19.83^{+13.03}_{-9.41}$	$18.39^{+1.49}_{-1.30}$	$5.39^{+0.71}_{-0.59}$	$0.82^{+0.88}_{-0.74}$	4.80	$1.70^{+1.84}_{-1.54}$	$3.11^{+1.27}_{-1.17}$
MS 2137.3-2353	39.0 [0.64]	0.44	$0.10^{+0.52}_{-0.08}$	$5.85^{+1.22}_{-0.96}$	$7.91^{+1.13}_{-0.99}$	$0.72^{+0.88}_{-0.19}$	3.26	$2.22^{+0.75}_{-0.59}$	$1.22^{+0.96}_{-0.69}$
Abell 383	24.9 [1.08]	0.48	$0.39^{+2.61}_{-0.36}$	$7.79^{+1.92}_{-1.38}$	$6.71^{+1.01}_{-0.89}$	$1.05^{+0.40}_{-0.33}$	3.22	$3.29^{+1.25}_{-1.02}$	$2.53^{+1.63}_{-1.13}$
Abell 2261	83.5 [3.34]	0.83	$15.53^{+4.43}_{-3.28}$	$8.23^{+0.98}_{0.82}$	$10.10^{+1.08}_{-0.99}$	$0.77^{+0.30}_{-0.16}$	5.64	$1.36^{+0.53}_{-0.28}$	$3.54^{+4.22}_{-1.24}$
MACS 2129.4-0741	75.6 [2.44]	0.79	$54.80^{+22.98}_{-17.57}$	$10.70^{+0.74}_{-0.64}$	$7.23^{+1.13}_{-0.99}$	$2.13^{+0.38}_{-0.36}$	2.03	$10.49^{+1.89}_{-1.75}$	$12.99^{+1.30}_{-1.18}$
MACS 1720.3+3536	86.8 [5.79]	1.07	$0.04^{+0.24}_{-0.04}$	$4.28^{+0.31}_{0.27}$	$10.83^{+0.65}_{-0.66}$	$0.38^{+0.25}_{-0.14}$	3.01	$1.27^{+0.84}_{-0.46}$	$5.06^{+0.82}_{-0.83}$

of freedom) for the fits ranged from 0.36 to 5.79. The degrees of freedom are counted as follows: two for each image point (one each for the x and y components), less two for each source point as they are, in essence, free parameters, less one for each varied parameter in the MCMC model. For the 8 observed clusters, 5 have reduced chi square statistics < 2, indicating that the extent of the match between data and observation for those models is generally in accordance with the error variances. For the other 3 (Abell 2261, MACS 2129 and MACS 1720), this indicates that the assumed errors may have been underestimated. Systematic error is the likely dominant component, which is discussed in Section 3.5.11.

All data images were reproduced except one. Occasionally, extra images were produced by the models, as noted in the relevant paragraphs of Section 3.4. Posterior distributions for all the cluster models are shown in the supplemental online material.

Strong lensing parameter inferences are summarized in Table 3.3. Figure 3.5 shows histograms for the 2-dimensional BCG offset, mass-to-light ratio, and Core Radius. Below we discuss the parameter inferences for each cluster for these key parameters: DM halo mass (i.e., m_{200}), BCG mass, and core radius (r_c).

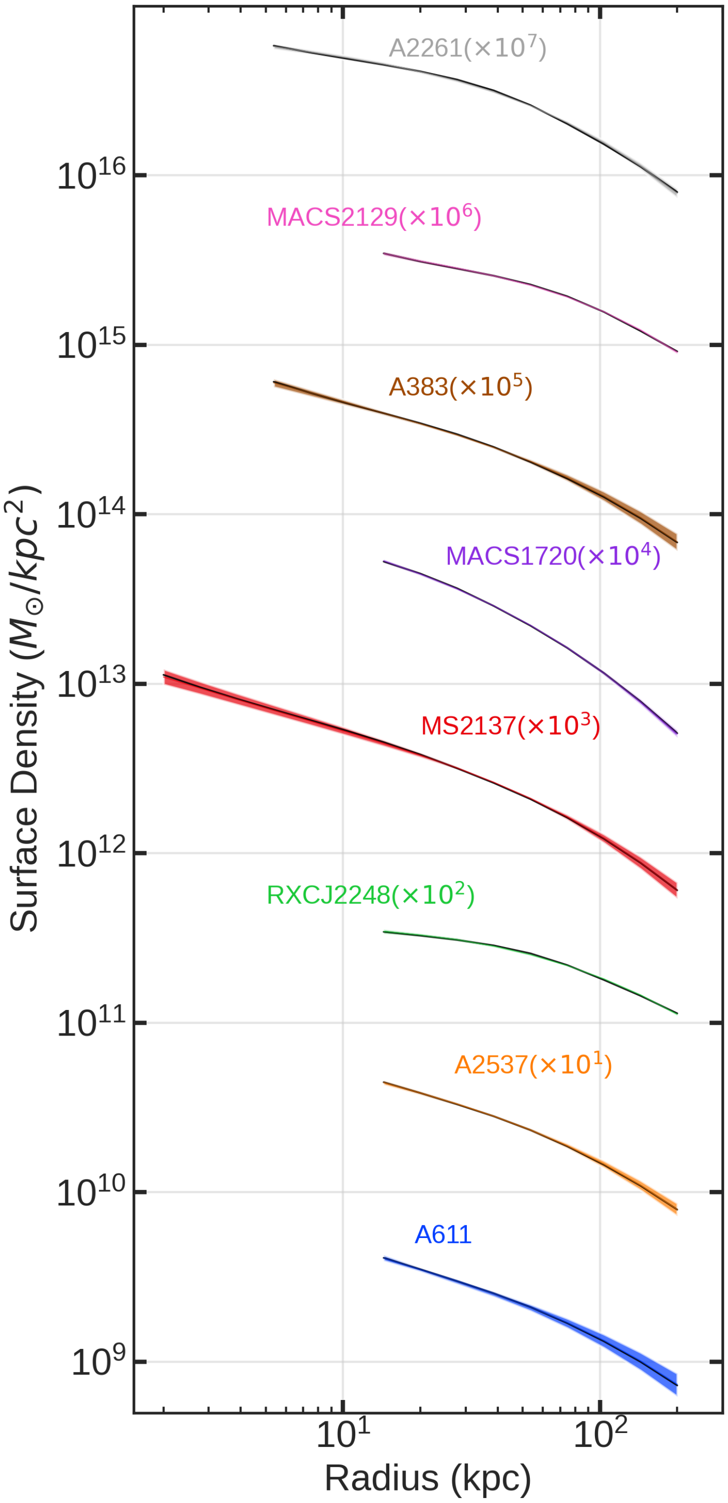
Figure 3.5: Histograms for 2-dimensional BCG offset, BCG mass-to-light ratio and DM halo core radius posteriors. The histograms are normalized to unity area, and have been smoothed slightly with a Gaussian kernel for display purposes.



3.5.1 Abell 611

Our inference of the mass of the central DM halo is $8.77^{+3.34}_{-2.30} \times 10^{14}\ M_\odot$. In Newman et al. [130] they infer a value of $6.2^{+0.7}_{-0.5} \times 10^{14}\ M_\odot$ using a combination of weak lensing, strong lensing and kinematic data and employing a generalized NFW profile. In a subsequent work, Newman et al. [133] found m_{200} of $8.31^{+1.46}_{-1.23} \times 10^{14}\ M_\odot$ using weak and strong lensing. In Donnarumma et al. [47] they use an NFW profile for the DM halo and infer strong lensing m_{200} values of $4.68 \pm 0.31 \times 10^{14}\ M_\odot$ using a dPIE profile for the BCG, and $6.32^{+0.51}_{-0.23} \times$

Figure 3.6: Surface density inferred from strong lensing ("data", in color) and the corresponding median SIDM modeled surface density ("model", in black) versus radius for each real cluster. The area shown for each is the 68% confidence interval. The clusters are offset by one decade each for display purposes.



$10^{14}\ M_\odot$ using a Sersic profile for the BCG. They arrive at an m_{200} of $9.32 \pm 1.39 \times 10^{14}\ M_\odot$ using X-ray gas temperature techniques. Our inference is consistent with all of these.

Our inferred BCG mass is $1.21^{+2.74}_{-0.88} \times 10^{12}\ M_\odot$. In Newman et al. [130] they infer a BCG mass of $1.01^{+0.24}_{-0.29} \times 10^{12}\ M_\odot$, while in Donnarumma et al. [46] they infer $6.17^{+1.32}_{-1.79} \times 10^{12}\ M_\odot$ for their case 6, which is the most comparable to our configuration. Our inference falls in the middle, and is consistent with the Newman et al. [130] result but in mild tension with the Donnarumma et al. [46] result.

We infer a cNFW core radius of $0.52^{+6.48}_{-0.48}$ kpc, which is consistent with the findings of Newman et al. [132], in which they infer a core size of $2.95^{+4.29}_{-2.01}$ kpc.

3.5.2 Abell 2537

For this cluster we infer a central DM halo mass of $9.42^{+1.77}_{-1.42} \times 10^{14}\ M_\odot$. In Newman et al. [133] they infer a somewhat higher value of $13.1^{+1.3}_{-1.2} \times 10^{14}\ M_\odot$ using a purely NFW profile and combining strong a weak lensing. In Cerny et al. [26], they employ a model with two dPIE DM halos with different center coordinates, and find a mass within 400 kpc of the BCG of $2.6 \pm 0.5 \times 10^{14}\ M_\odot$. Our inferred value for mass within that radius is nearly the same, with a median value of $2.5 \times 10^{14}\ M_\odot$.

Our inference for the BCG mass is $0.93^{+1.09}_{-0.48} \times 10^{12}\ M_\odot$. Other authors do not specify their inferred BCG masses, but we note that in Newman et al. [133] they measure a luminosity of $5.86 \times 10^{11}\ M_\odot$, and in Cerny et al. [26] they measure a luminosity of $5.27 \times 10^{11}\ M_\odot$, each of which when combined with our inferred mass results in a reasonable mass-to-light ratio of approximately 1.6 to 1.8.

For Abell 2537, we infer a modest core radius of $1.18^{+6.76}_{-1.12}$ kpc. This is in tension with the inference of Newman et al. [132] of $46.8^{+34.5}_{-19.2}$ kpc. We are unsure of the reason for the discrepancy, but note that the measured half-light radius of the BCG is 15.7 kpc, which goes beyond the DM core radius, increasing the possibility for degenerate solutions between the two. Examination of the posterior plot (see supplemental online material) shows a mild bimodality in the BCG mass, possibly also admitting a larger core size. Note also that this cluster has the fewest images in the sample (16), and the fewest source points (4). The image position data is the same as that used in Newman et al. [133]. A small data set could potentially limit the ability of the model to accurately constrain the mass profile.

3.5.3 RXC J2248.7-4431

For this cluster we infer a central DM halo mass of $18.39^{+1.49}_{-1.30} \times 10^{14}\ M_{\odot}$. In Caminha et al. [22], in their reference case 2 they find a mass of $2.90 \pm 0.02 \times 10^{14}\ M_{\odot}$ within an aperture of 250 kpc. The corresponding median value for our inference would be very similar; $2.88 \times 10^{14}\ M_{\odot}$. In Bonamigo et al. [12], they infer an m_{200} of $20.3 \pm 6.7 \times 10^{14}\ M_{\odot}$, in good agreement with our inference. We infer a BCG mass for RXCJ2248 of $0.82^{+0.88}_{-0.74} \times 10^{12}\ M_{\odot}$. We were unable to find comparable inferences from other authors. We infer a core radius of $19.83^{+13.03}_{-9.41}$ kpc, and we were not able to find comparable cNFW core size inferences from other authors.

3.5.4 MS 2137.3-2353

We infer a central DM halo mass of $5.85^{+1.22}_{-0.96} \times 10^{14}\ M_{\odot}$ for this cluster. In Donnarumma et al. [46], the found an m_{200} of $4.4^{+0.6}_{-0.5} \times 10^{14}\ M_{\odot}$ using a strong lensing model, and Newman et al. [133] found an m_{200} of $3.63^{+1.26}_{-0.81} \times 10^{14}\ M_{\odot}$ using a combination of strong and weak lensing. Both of these figures are consistent with our inference.

While we infer a BCG mass of $0.72^{+0.88}_{-0.19} \times 10^{12}\ M_{\odot}$, other authors do not report their

BCG masses. However, we note that our meaured BCG luminosity of $3.26 \times 10^{11}\ L_{\odot}$ and that measured in Newman et al. [133] of $3.20 \times 10^{11}\ L_{\odot}$ are similar, and result reasonable mass-to-light ratios of approximately 2.2.

We infer a very small core radius of $0.10^{+0.52}_{-0.08}$ kpc, as does Newman et al. [132], at $2.82^{+3.01}_{-2.39}$ kpc.

3.5.5 Abell 383

For this cluster we infer a central DM halo mass of $7.79^{+1.92}_{-1.38} \times 10^{14}\ M_{\odot}$. In Zitrin et al. [210] they find a similar mass of $7.67 \times 10^{14}\ M_{\odot}$ from a generalized NFW profile. In addition, we find the 2D projected mass on the halo within 50 kpc, the distance to the large tangential arc of systems 1 and 2, is $2 \times 10^{13}\ M_{\odot}$. This value is consistent with $1.7 \times 10^{13}\ M_{\odot}$, $2.2 \times 10^{13}\ M_{\odot}$, and $2 \times 10^{13}\ M_{\odot}$ from Monna et al. [116], Newman et al. [131], and Zitrin et al. [210] respectively.

We infer a core radius of $6.66^{+6.89}_{-4.20}$ kpc for the DM halo. We were not able to find comparable cNFW core size inferences from other authors. Our inferred BCG mass is $1.05^{+0.40}_{-0.33}\times 10^{12}\ M_{\odot}$ which is lower than the results Monna et al. [116] of $6.05 \times 10^{12}\ M_{\odot}$ or $6.13 \times 10^{12}\ M_{\odot}$, depending if they included or did not include velocity dispersion measurements. We note that our measured luminosity of $3.22 \times 10^{11} L_{\odot}$ when coupled with our inferred BCG mass results in a reasonable mass-to-light ratio of ~ 3.3.

3.5.6 Abell 2261

For cluster Abell 2261, we infer a central DM halo mass of $8.23^{+0.98}_{-0.82} \times 10^{14}\ M_{\odot}$. In Coe et al. [34], they found m_{200} of $9.54^{+0.84}_{-0.84} \times 10^{14}\ M_{\odot}$ by using a combination of weak lensing and strong lensing when assuming a spherical NFW profile halo, which is consistent with our inference.

We inferred a BCG mass of $0.77^{+0.30}_{-0.16} \times 10^{12}\ M_{\odot}$, and a core radius of $15.53^{+4.43}_{-3.28}$ kpc. We were unable to find inferences from other authors for these parameters.

3.5.7 MACS 2129.4-0741

For this cluster we infer a central DM halo mass of $10.70^{+0.74}_{-0.64} \times 10^{14}\ M_{\odot}$. In Monna et al. [117], they specify the mass enclosed within the Einstein Parameter, Θ_E, given as $29 \pm 4"$. Within this radius, they find a mass of $8.6 \pm 0.6 \times 10^{13}\ M_{\odot}$. Using our model, we find a corresponding median value of $17.5 \times 10^{13}\ M_{\odot}$ within that same radius.

Similarly, our model infers a median BCG mass within $\Theta_E = 29"$ of $2.03 \times 10^{12}\ M_{\odot}$, whereas Monna et al. [117] finds $8.4 \pm 2 \times 10^{12}\ M_{\odot}$.

This cluster has the largest inferred core size in our sample, at $54.8^{+23.0}_{-17.6}$ kpc. In Monna et al. [117] they find an even larger core radius of 101^{+13}_{-11} kpc. However, as noted in Monna et al. [117], the DM halo core radius correlates with the BCG mass profile, thus they speculate that the core radius may be large because of an overestimate of the BCG mass.

3.5.8 MACS 1720.3+3536

For this cluster we infer a central DM halo mass of $4.28^{+0.31}_{0.27} \times 10^{14}\ M_{\odot}$. In Zitrin et al. [212] they use strong lensing to specify the mass enclosed within 136" of the center. They find a value of $3.35 \times 10^{14}\ M_{\odot}$ for their NFW model. The corresponding median value for our model is a similar $3.57 \times 10^{14}\ M_{\odot}$. Note that our model is very close to NFW, as the inferred core radius is quite small at $0.04^{+0.24}_{-0.04}$ kpc, the smallest in our sample. We can find no other works that present estimates of the core radius or BCG mass for this cluster.

3.5.9 Halo and BCG masses

The primary results of the strong lensing models are the surface density profiles, which are shown as the colored bands in Figure 3.6. These profiles are used as data for the SIDM profile matching described in Section 3.6. The black lines in Figure 3.6 are the surface density profiles as modeled by the SIDM halo matching. The match is better than 0.1 dex for all clusters.

The posterior distributions of halo core radius, halo mass and BCG mass are shown in Table 3.3. (Note that these inferences are not used for the SIDM halo analysis.) Our DM halo core inferences range from essentially zero (i.e., MS2137, MACS1720) to more than 50 kpc (MACS2129). Core size is an important characteristic in this analysis, as a small core rules out strong self-interactions.

We infer DM halo masses in the range of $4 \times 10^{14} M_\odot$ to $1.8 \times 10^{15} M_\odot$. These are broadly consistent with those found by other authors. We infer BCG masses in the range of $7 \times 10^{11} M_\odot$ to $2.1 \times 10^{12} M_\odot$. These are are generally consistent with those found for these objects by other authors, with the exceptions of A383 (where our inference is lower than that of Monna et al. [116]) and MACS2129 (where our inference is lower than that of Monna et al. [117], however, those authors speculate that their BCG mass may be an overestimate).

A potential degeneracy can occur in lens models between BCG mass and halo mass, especially if the BCG is closely co-centered with the DM halo, since the mass is a free parameter in both profiles and both deflect image positions similarly. In a similar way, degeneracy can occur between BCG mass and core size. As a sanity check for the inferred BCG masses, we calculate the mass-to-light ratio for each, using the closest HST filter to V-band in the rest frame. The luminosities were measured in Gaia/Source Extractor and corrected for galactic extinction.

Newman et al. [133] used a stellar population synthesis model and assumed a Chabrier initial mass function (IMF) [28] to arrive at a mass-to-light ratio $\Upsilon_V = M_*/L_V$ for 7 giant elliptical BCGs, finding a range of 1.80 to 2.32 in V-band, with low scatter. The assumption of a Salpeter IMF [166] increases the ratio by a factor of 1.78, resulting in an upper value of 4.13. In Andrade et al. [3], they note the possibility of a super-Salpeter IMF in Abell 611, which would further increase the upper limit. Our results are consistent with those expected values, with the exception of MACS2129, for which we find $\Upsilon_V = 10.49^{+1.89}_{-1.75}$. We do not know the reason for this outlier, but we speculate that it may be caused by difficulty in separating the BCG light from three other bright objects within the half-light radius of the BCG (approximately 17 kpc). As a test, we made a separate strong lens model for MACS2129 with Υ_V constrained to a value of 4. The resulting fit was inferior to the original, with χ^2 increasing from 77 to 187, RMS position error increasing from $0\farcs79$ to $1\farcs30$, and only 28 out of 31 images matched. The inferred core size from that run was smaller than the original run, decreasing to a median posterior value of 39.4 kpc from the original 54.8 kpc, which would cause a decrease in the inferred SIDM cross section for that cluster.

3.5.10 Halo Concentrations

We infer halo concentrations with median posterior values ranging from $5.39^{+0.71}_{-0.59}$ (RXCJ2248) to $10.83^{+0.65}_{-0.66}$ (MACS1720), with a median value of 7.0 (see Table 3.3). For comparison, the concentration-mass relation of Diemer and Joyce [45] predicts median concentrations $\simeq 4.0$ for halos at redshifts of 0.2 to 0.6 and in the mass range of $10^{14} M_\odot$ to $10^{15} M_\odot$. Also, Merten et al. [114] observed concentrations of 3.7 ± 0.65 for 19 X-ray clusters in the CLASH sample. As we discuss in Appendix B.3, this discrepancy can potentially be explained if the lines of sight to lensing clusters are preferentially oriented along the major axis of the clusters. The concentrations for prolate halos can be biased upward by up to 60%. As described in Section 3.6, we vary the cluster's LOS axis ratio in our SIDM model to account for this effect. (The concentration parameter *per se* is not used in the subsequent SIDM analysis; rather, the entire radial halo profile is used, as described in Section 3.6.) In addition, Fielder et al. [57] showed that concentration will be significantly higher for a DM halo when mass from its associated subhalos are excluded from the calculation. Our models do account for some subhalo mass separately (for the larger perturbers), so we would expect our primary halo to be more concentrated than that of the cluster as a whole.

3.5.11 Strong Lensing Systematic Errors

Statistical errors in image positions in strong lensing studies are often quite small, on the order of 1 to 2 detector pixel widths. In our case, most images were from the Hubble Space Telescope ACS instrument, with a pixel with of 0.″05. Much more significant are systematic effects, which can include misidentified images, inaccurate image redshifts (especially for photometric redshift data; see Cerny et al. [26]), unmodeled substructure and correlated mass along the line of sight. As in many other strong lensing studies, we account for the latter item via two external shear components in the model, which adds two degrees of freedom to the model. To account for other systematic errors, most studies increase the assumed position error well beyond that of the statistical uncertainty. Some assumptions for position error from other authors for the same clusters in our sample are as follows: 0.″2 for Abell 611 [47], 0.″5 for MS2137, A383, A611 and A2537 [133]; 1.″4 for A2261 [34]; 0.″5 to 1.″0 for MACS 2129 [119]; 1.″0 for A383 [116]; 0.″5 for RXCJ 2248 [22]. Note that some of those studies involved multiple data sets, combining strong lensing with weak lensing, X-ray analysis and/or stellar kinematic analysis. In that case, the position error assumption is more important because it scales the χ^2 of the strong lensing component only, thus affecting the weighting of the strong lensing relative to the other data. In our case, we have only one component, so the relative weighting between data sources is not a concern. This was part of the motivation for using strong lensing alone in our study. We have therefore adopted a position error assumption of $0\farcs5$, which seems broadly consistent with the residuals for all of the clusters in our our analysis. Note that this applies separately to the x and y components, so that the total uncertainty is $0\farcs71$.

We note that the cluster MACS1720 has the highest reduced χ^2 in our sample (5.79) and the highest RMS postion error ($1\farcs07$). In order verify that the assumed data position error of $0\farcs5$ did not adversely impact the results, we made an additional run for that cluster with an assumed data position error of $1\farcs0$. The resulting reduced χ^2 fell to 1.63. The inferred surface density profile declined slightly and its uncertainty approximately doubled; for example, at 20 kpc the inferred surface density changed from $4.42 \pm 0.09 \times 10^9 M_\odot$ to $4.22 \pm 0.20 \times 10^9 M_\odot$. The concentration fell from from 10.83 ± 0.66 to 9.71 ± 0.84. The resulting inference for $\log_{10}(\sigma/m)$ didn't change much, declining from -1.35 ± 0.18 to -1.40 ± 0.20.

As a check on the position error assumption, we discuss and plot the error residuals from strong lensing in Appendix B.4 and Figure B.3. The residuals appear generally Gaussian in shape, and have standard deviation of approximately $0\farcs4$ to $0\farcs8$, roughly consistent with the assumed position uncertainty of $0\farcs5$.

3.5.12 BCG Offsets from DM Halo Centers

An observable consequence of cored SIDM halos would be oscillations of BCG about the center of the halo after mergers, which would persist for several Gyr [86]. In Harvey et al. [68], they used simulations to examine the effect of DM self-interaction on BCG offsets, and concluded that the distribution of such offsets from an ensemble of clusters would exhibit a median value of 3.8 ± 0.7 kpc for a CDM scenario (i.e., $\sigma/m = 0$ cm^2/g), and 8.6 ± 0.7 kpc in a scenario where $\sigma/m = 1$ cm^2/g. The rightmost column of Table 3.3 and the first histogram in Figure 3.5 show our inferred 2-dimensional offset of the DM halo centers for each cluster, which range from approximately 1 kpc to 10 kpc, with a median of approximately

4 kpc, except for MACS2129, which is an outlier at 10 kpc to 15 kpc. Since these are 2-dimensional offsets, we can estimate that the corresponding 3-dimensional offsets would be larger by a factor of approximately $\sqrt{3/2} = 1.23$, although the precise value would depend on the distribution of the offsets. Using this factor results in a corresponding median value of approximately 5 kpc for our data, which would imply $\sigma/m \ll 1\ \mathrm{cm}^2/\mathrm{g}$ in light of the findings of Harvey et al. [68]. However we also note that Harvey et al. [67] found that the error estimates from the posteriors of MCMC lens modeling could be understated by as much as a factor of 10. Applying this to the predicted offsets in Table 3 would result in uncertainties larger than the offsets themselves, making an inference discerning CMD from $1\ \mathrm{cm}^2/\mathrm{g}$ SIDM much weaker.

3.6 SIDM Halo Analysis

It has been argued from observations of groups and clusters of galaxies, that the cross section must be velocity-dependent to have a significant effect for less massive galaxies [80]. This velocity dependence can occur in several ways: resonant self-interaction of dark matter [31], light mediator models with either elastic or inelastic collisions [55, 100, 188, 187, 2, 11, 196], bound states [16], and strongly interacting massive particles [73, 29]. See Chu et al. [32] for a review of these models in addition to a model-independent approach for approximating velocity dependence using effective range theory. As the range of particle velocities in our cluster sample is relatively narrow, we opt for a constant velocity analysis here.

3.6.1 SIDM Halo Model

For the SIDM halo model, we following the procedure in Kaplinghat et al. [80] and write the full profile as an inner isothermal profile and an outer NFW profile, with the two profiles matched in mass and density at a radius r_1, which is determined by the cross section. The idea here is to get a smooth density profile that interpolates between an isothermal core and a collisionless outer envelope well enough to capture the halo profiles inferred from SIDM simulations. The characteristic radius r_1 dividing the inner and outer regions can be approximated by setting the average scattering rate per particle times the age of the halo to unity:

$$\text{rate x time} \approx \frac{<\sigma v>}{m} \rho(r_1)\, t_{\text{age}} = 1, \tag{3.5}$$

where σ is the scattering transfer cross section, v is the relative velocity between DM particles, $\rho(r_1)$ is the density of DM at the characteristic radius, t_{age} is the age of the halo, and $<...>$ denotes averaging over the velocity distribution. This is a simplification of the time-dependent process of halo assembly, but it compares well to numerical simulations [80, 161] because of the approach to equilibrium.

For the age of each halo, we adopt a value equal to the time at the redshift of the halo, as shown in Table 3.2. The average relative velocity of SIDM particles, assuming a Maxwellian velocity distribution, can be shown to be $\frac{4}{\sqrt{\pi}}\sigma_0$, where σ_0 is the central velocity dispersion. Assuming a constant cross section σ over the range of velocities accessible in the cluster, we have $\langle \sigma v \rangle = \sigma \langle v \rangle = \frac{4}{\sqrt{\pi}} \sigma\, \sigma_0$. For a model with a sharp velocity dependence, σ should be interpreted as the velocity averaged transfer cross section [189, 159, 77, 11].

Considering the two regions, one which is thermalized by self-interactions and one which remains largely non-interacting, the complete profile can be written as

$$\rho(r) = \begin{cases} \rho_{\text{iso}}(r), & r < r_1 \\ \rho_{\text{NFW}}(r), & r \geq r_1 \end{cases} \tag{3.6}$$

where $\rho_{\text{iso}}(r)$ is a cored isothermal profile [79], and $\rho_{\text{NFW}}(r)$ is the NFW profile (see Equation 3.1, but with $r_c \to 1$). For the region interior to r_1, interactions are common and we consider the DM particles to behave as an ideal gas, characterized by a pressure (p), density (ρ) and (one-dimensional) velocity dispersion (σ_0) that obey an equation of state $p = \rho\sigma_0^2$. Assuming that the DM particles in the central region achieve hydrostatic equilibrium, we have

$$\nabla p = -\rho \nabla \Phi_{\text{tot}} \tag{3.7}$$

where Φ_{tot} is the total gravitational potential from both DM and baryons. In this analysis, we use the fitted BCG from strong lensing with a dPIE profile (Equation 3.2) for the baryonic mass. The gravitational potential must satisfy the Poisson equation

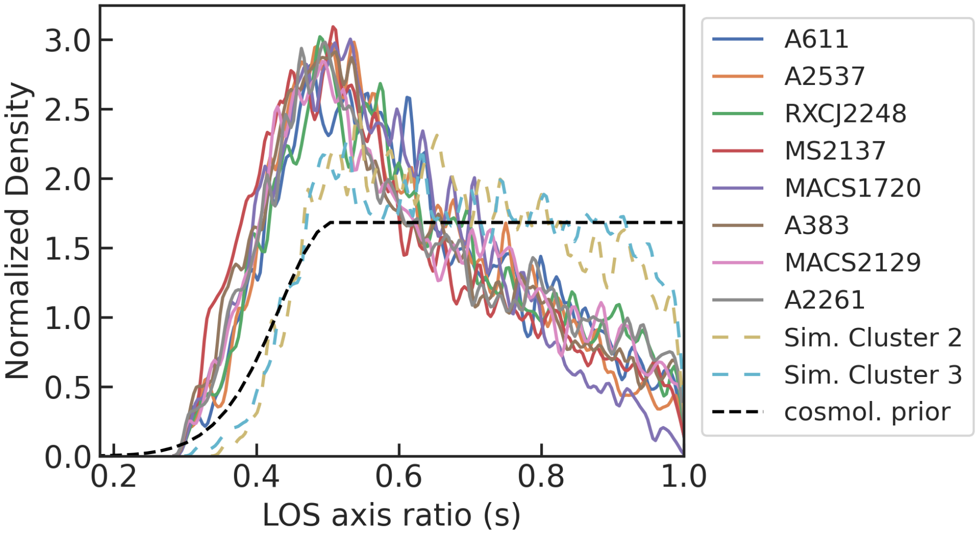
$$\nabla^2 \phi_{\text{tot}} = 4\pi G(\rho_{\text{DM}} + \rho_{\text{BCG}}) \tag{3.8}$$

where G is Newton's gravitational constant, ρ_{DM} is the mass density of DM and ρ_{BCG} is the mass density of the BCG. Thus the SIDM density profile explicitly depends on the stellar distribution [79]. The solution to two equations above gives $\rho_{\text{iso}}(r)$ for the interior region.

We assume that the isothermal solution is a cored profile at small radii with central density ρ_0 and central dispersion σ_0. For concreteness, we impose this boundary condition at 1% of the stellar core radius (r_{core} in the dPIE profile). We then evolve the isothermal Jeans equation outwards stopping at the radius r_1 when the SIDM density $\rho(r_1)$ satisfies Eq. 3.5. By matching the mass and density of ρ_{iso} to a NFW profile, we can find the NFW parameters ρ_s and r_s from ρ_0 and σ_0. Note that we do not search for a second solution at larger r_1 [80, 161], which could be just as good a fit. This second solution with the same ρ_0 is reminiscent of the core collapse phase [5, 51, 53, 135], but it is not clear if this connection holds up in detail.

As discussed in 3.5.10 and Appendix B.3, the apparent concentrations of halos that are elongated along the LOS are higher than the actual concentration. To account for this we

Figure 3.7: Histogram plot for the line of sight axis ratio "s". The prior is shown as a dashed black line and is based on Vega-Ferrero et al. [193] for s<0.5 and is flat for s>0.5.



transform our spherical halo profile $\rho(r)$ in Eq. 3.6 into an ellipsoidal profile $\rho_e(r_e) = \rho(r_e)$ with $r_e^2 = R^2 + z^2 s^2$, and orient the z-direction along the line-of-sight. This implies that the ellipsoidal surface density $\kappa_e(R) = \kappa(R)/s$. Choosing the LOS axis ratio $s < 1$ implies that for the same mass, concentration and cross section, the surface density is higher. We allow s to vary during the MCMC fitting procedure, which would allow for smaller ρ_s and hence larger cross sections. Note that the central density does not vary linearly with ρ_s as the cross section is changed.

To arrive at a prior distribution for s, we use the results in Vega-Ferrero et al. [193] and Bonamigo et al. [14] for relaxed clusters. We utilize the PDF from Vega-Ferrero et al. [193] for values of s < 0.5. The PDF reaches a maximum at $s \simeq 0.5$, and declines as s approaches unity. For $s > 0.5$, since our clusters may not always be fully oriented along the LOS, so we opt for a flat prior for $s > 0.5$. The prior for s is shown in Figure 3.7, along with the inferred posteriors for each cluster from the SIDM model. Our data prefers LOS axis ratios near 0.5 for all clusters.

There are two pitfalls that impede seamless parameter space exploration using the halo profile in Eq. 3.6. First is that when we scan over ρ_0 and σ_0, there is no guarantee that r_1 can be found. To insure against this failure, we change our variable from ρ_0 to $R_0 = \rho_0(\sigma/m)(4/\sqrt{\pi})\sigma_0 t_{\rm age}$ and put a prior on R_0 that is larger than unity [155, 161]. Since the density falls with increasing radius, we are guaranteed to get a solution for r_1. For the present study we have adopted the following priors on the five parameters that are varied to find the SIDM solutions: $5 < R_0 < 500$, $400 < \sigma_0 < 1400$, $0.01 < \sigma/m < 1.0$, $10^{11}M_\odot < M_{\rm BCG} < 10^{13}M_\odot$, $0.3 < s < 1$. With the exception of the lower limit of σ/m, our prior boundaries do not impact the inferred posteriors.

The second issue is that for a given r_1 and $\rho_{\rm iso}(r)$, it may not be possible to find a matching NFW profile (see also, Robertson et al. [161]). The procedure to find the matching NFW profile proceeds by first matching $\gamma_M(r) = M(r)/(4\pi r^3\rho(r))$ where $M(r)$ is the dark matter halo mass enclosed within radius r. Since for the NFW profile $\gamma_M(r)$ is only a function of r/r_s, we get r_1/r_s and therefore r_s. Then we can match the isothermal and NFW $M(r)$ at r_1 to infer ρ_s. However, the first step can fail if $\gamma_M(r_1) < 0.5$ for the isothermal profile, because $\gamma_M(r) \geq 0.5$ for the NFW profile. However, these cases are rare and they are not physically interesting, as it requires a core unlike that seen in simulations with $\rho(r_1) \simeq \rho_0$. With these checks in place, it is possible to find unique NFW matches for all $\rho_{\rm iso}(r)$ profiles generated by the likelihood sampler. Once we have the full density profile, then we can compute the 2D density profiles to compare to the data.

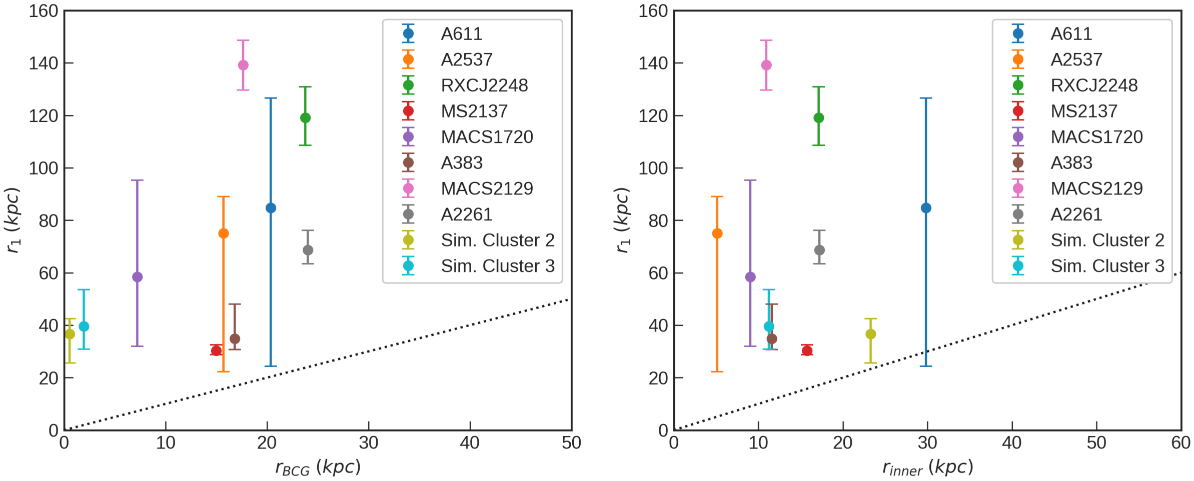
Before we discuss the results from the SIDM halo matching process discussed above, it is worth noting this analytic model (described in Kaplinghat et al. [80]) has been remarkably successful in capturing the density profiles in simulations from dwarf galaxies to clusters of galaxies [80, 158, 179, 162, 155, 161]. The study in Sokolenko et al. [179] specifically focused on clusters and investigated different ways of matching the isothermal and NFW solutions. They found that the model of Kaplinghat et al. [80] can reproduce the core density of the halos, as shown in the left panel of Figure 22 of their paper. The inferred core is recovered at the 10-20% level for cross sections below 1 cm^2/g, which are relevant for our study. Remarkably, the analytic model provides an unbiased description over almost two orders of magnitude in cross section from 0.1 to 10 cm^2/g. There seems to be a bias creeping in at $\sigma/m = 0.1\ \mathrm{cm}^2/\mathrm{g}$; if this trend continues to lower cross sections, this would imply that our inferred cross sections are biased somewhat high.

The right panel of the same figure from Sokolenko et al. [179] shows that the kinetic energy inferred from the central dispersion is systematically lower than the average kinetic energy within r_1. This bias is important for us since we need both the cross section and the average relative velocity of DM particles to extract constraints on SIDM models. The results from Sokolenko et al. [179] indicate that the average relative velocity can be biased low by about 20% if it is inferred from the central dispersion σ_0 using the halo profile in Eq. 3.6. While this seems like a small effect, it can be an important systematic if we are constraining a cross section model with a sharp velocity dependence. Since most models predict cross sections that fall with increasing relative velocity at these high velocities, a conservative way of constraining SIDM models is to use the largest possible average relative velocity for each cluster. We can estimate this maximum average relative velocity as $1.4v_{\mathrm{max}}$ using the fact that the maximum RMS velocity of DM particles is very close to v_{max} [163]. We provide both σ_0 and V_{max} posteriors for the clusters.

3.6.1.1 The Importance of the Characteristic Radius r_1

The impacts of DM self-interaction on galaxy cluster properties was examined in Robertson et al. [160] using the Bahamas simulations. In reviewing the SIDM halo-matching model, those authors conclude “we find the model to provide a good description of simulated SIDM density profiles, and (importantly) find that the isothermal Jeans model can be used to infer the cross-section from a simulated halo‘s density profile...." In the upper right panel of Figure 2 from that work, they show that the median density of the simulated cluster of

Figure 3.8: Comparison of the characteristic radii r_1 to the BCG radii and the radii of the innermost strong lensing data points. The dotted lines indicate 1:1 equality. Note that $r_1 > r_{\mathrm{BCG}}$ and $r_1 > r_{\mathrm{inner}}$ in all cases except A611, which allows a solution mode with low cross section and an r_1 value extending down to ~ 25 kpc at the 1σ level.



mass $\sim 10^{15} M_{\odot}$ at a radius of 10 kpc is approximately 70% higher in CDM as compared to SIDM with σ/m of 0.1 cm^2/g. The divergence in the median densities persists out to radii $\gtrsim 40$ kpc. The scatter about the mean in the inner parts is not random – it is mostly driven by the scatter in the outer profiles due to the concentration–mass relation and the impact of the baryons on the inner profile. Both of these effects are captured by the isothermal halo profile used in Robertson et al. [160] and our work.

The model parameter r_1 indicates the matching point between the isothermal and NFW halo profiles. If r_1 occurs well outside the BCG, then we can be reasonably confident that adiabatic contraction does not impact the outer NFW profile significantly. If the value of r_1 is within the range of strong lensing data, then we can be reasonably confident that the isothermal profile is being constrained by the data. Figure 3.8 shows r_1 for the cluster sample in relation to the BCG radius (left panel) and the innermost strong lensing data point (right panel). The r_1 radii are larger than both, with the exception of A611 cluster. A611 has a bi-modal solution that extends down to low cross sections and values of r_1 that are comparable to those of the BCG radius and innermost data points.

Writing equation 3.5 in terms of the density at r_1, we see that $\rho(r_1) = 1/(\frac{<\sigma v>}{m}\ t_{\rm age})$. Since $t_{\rm age}$ varies only modestly over our data set, constraints on $\rho(r_1)$ are the drivers of the constraints on cross section. For example, taking a cluster at redshift 0.3, $< v >$=1500 km/s, and $\sigma/m \geq$0.1 cm^2/g, we arrive at $\rho(r_1) \leq 3.12 \times 10^6 M_\odot$ kpc^{-3}. Halos that are more dense than this at their characteristic radius would disfavor SIDM cross sections at or above 0.1 cm^2/g. Note that in Robertson et al. [160], Figure 2, the densities of simulated CDM and SIDM halos (with SIDM cross section of 0.1 cm^2/g) do indeed begin to depart from one another at approximately the density calculated here.

3.6.2 SIDM Halo Matching Inference Results

We employ an MCMC fitting code to implement the second stage of our analysis, using as input the strong lensing posterior for the DM surface density profile for each cluster. By using the entire surface density profile, the model can appropriately match the halo density profile, and accommodate multiple possible modes in the strong lensing solution set. We used the Dynesty sampler [180, 72] to generate the MCMC posterior chain for this stage. The key posterior parameters are: cross section, r_1 (the characteristic radius), average particle speed, and BCG mass, as noted in Figure 3.1. From this we infer the central density (ρ_0) and the matched NFW halo parameters ($M_{\rm vir}$ and $c_{\rm vir}$). We fit cross sections at constant velocity, noting that the median maximum circular velocities (a rough proxy for relative particle speed) for the clusters in the sample are in the range of 1,000 km/s to 1,800 km/s, with most around 1,200 km/s.

As shown schematically in Figure 3.1, surface density posteriors at radii ranging from 2 kpc to 200 kpc are generated by the strong lensing model and were first decomposed into principal components using the Principal Component Analysis (PCA) in the Scikit-Learn software package [140], which typically yielded 3 or 4 components for any given cluster. We found that supplying 10 logarithmically spaced radial data points ranging from 10 kpc to 200 kpc enabled good surface density profile reproduction for most clusters, but three clusters (A383, A2261 and MS2137) required that an additional 5 bins extend farther inward, to 2 kpc, to ensure that the model made a good match to the shape of the surface density profile inferred from strong lensing.

A Gaussian Mixture Model (GMM) was used to model the likelihood using the principal components of the surface density profiles. We calculated the Bayesian Information Criteria (BIC) for each cluster while varying the number of GMM components, and found that having more than four components in the GMM did not improve the BIC significantly. We therefore used four components in the GMM.

In addition to the surface density profiles, the SIDM model also requires the parameters of the BCG density profile as described above. We use the BCG scale radius as measured from photometry for this purpose. However, we do not use the BCG mass posterior from strong lensing. By varying the BCG mass independently, we are able to explore the degeneracy between the BCG mass and cross section; see Figure 3.12 for a comparison of the BCG mass inferred from the SIDM and cNFW fits, and the associated discussion in Section 3.6.2. Inference of the BCG mass is sometimes multimodal, with larger BCG masses being associated with larger cross sections (lower DM core densities). Posterior distributions for the SIDM halo model are shown in the supplemental online material.

The surface density models produced by the SIDM halo matching model reproduced the data to within 0.1 dex or better. Plots of the matches for each cluster are shown in the Appendix in Figure B.4. Note that the model relies on surface density profile rather than direct inferences of parameters such as BCG mass, halo mass, concentration, etc. This allows multimodal solutions in those parameters to be accurately incorporated into the model.

Figure 3.11 plots the inferred $r\,\rho(r)$ of the central DM halo versus radius for each cluster

Figure 3.9: Posterior histograms for $\log_{10}(\sigma/m)$. The simulated clusters are shown as dashed lines.

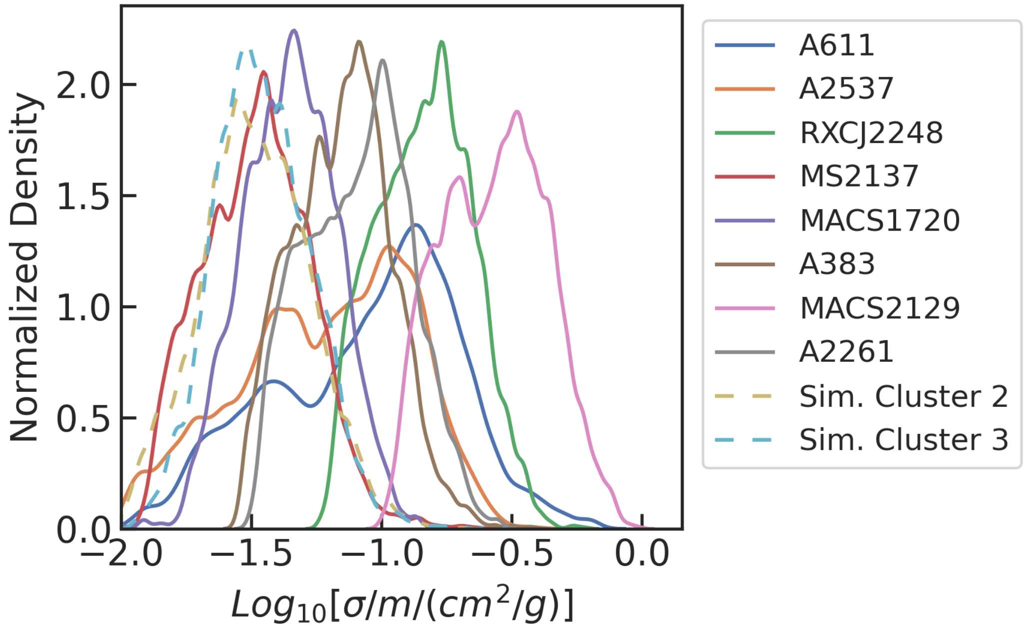


Table 3.4: SIDM parameter posterior summary. The columns are: (1) cluster name, (2) $\log_{10}$ cross section per unit mass, median and 68% C.L., (3) $\log_{10}$ cross section per unit mass, mean ± standard deviation, (4) characteristic radius, (5) central density, (6) central velocity dispersion, (7) halo maximum circular velocity, (8) BCG mass, (9) halo virial mass, and (10) concentration. The indicated ranges are the 68% confidence intervals.

	$\log_{10}\left(\frac{\sigma/m}{cm^2g^{-1}}\right)$	$\log_{10}\left(\frac{\sigma/m}{cm^2g^{-1}}\right)$	r_1	ρ_0	σ_0	$v_{\rm max}$	$m_{\rm BCG}$	$m_{\rm vir}$	$c_{\rm vir}$
Cluster Name	(median)	(mean)	(kpc)	$(10^6\,M_\odot\ kpc^{-3})$	$(km\ s^{-1})$	$(km\ s^{-1})$	$(10^{12}\,M_\odot\ )$	$(10^{14}\,M_\odot\ )$	
A611	$-1.00^{+0.27}_{-0.47}$	-1.06 ± 0.35	$94.5^{+15.4}_{-64.8}$	$165.1^{+97.8}_{-60.5}$	649^{+138}_{-168}	1205^{+206}_{-169}	$2.32^{+0.90}_{-1.72}$	$5.05^{+2.97}_{-1.79}$	$5.86^{+1.22}_{-1.11}$
A2537	$-1.18^{+0.30}_{-0.40}$	-1.21 ± 0.33	$83.6^{+18.3}_{-56.3}$	$205.4^{+58.5}_{-108.1}$	629^{+151}_{-179}	1235^{+209}_{-150}	$1.88^{+0.62}_{-1.57}$	$5.35^{+2.65}_{-1.65}$	$6.20^{+0.94}_{-0.82}$
RXCJ2248	$-0.83^{+0.18}_{-0.21}$	-0.83 ± 0.18	$117.3^{+16.9}_{-13.1}$	$17.5^{+7.2}_{-5.2}$	706^{+121}_{-92}	1554^{+277}_{-194}	$0.67^{+0.72}_{-0.28}$	$12.24^{+6.94}_{-3.78}$	$3.78^{+0.60}_{-0.45}$
MS2137	$-1.47^{+0.20}_{-0.23}$	-1.48 ± 0.21	$32.8^{+4.6}_{-2.5}$	$256.6^{+46.3}_{-43.9}$	560^{+70}_{-55}	1085^{+205}_{-136}	$1.04^{+0.18}_{-0.16}$	$3.25^{+1.74}_{-1.04}$	$7.46^{+1.34}_{-1.14}$
MACS1720	$-1.35^{+0.18}_{-0.19}$	-1.35 ± 0.18	$47.2^{+17.6}_{-6.2}$	$453.7^{+930.9}_{-337.4}$	628^{+99}_{-87}	1072^{+159}_{-129}	$0.78^{+0.66}_{-0.56}$	$2.41^{+1.03}_{-0.68}$	$10.10^{+1.08}_{-1.15}$
A383	$-1.14^{+0.18}_{-0.22}$	-1.15 ± 0.19	$61.0^{+9.3}_{-3.1}$	$142.6^{+32.8}_{-33.0}$	606^{+89}_{-63}	1146^{+215}_{-152}	$1.33^{+0.20}_{-0.18}$	$4.56^{+2.60}_{-1.54}$	$6.47^{+1.21}_{-1.05}$
MACS2129	$-0.56^{+0.20}_{-0.24}$	-0.57 ± 0.20	$153.0^{+12.7}_{-11.1}$	$54.0^{+11.9}_{-9.5}$	741^{+117}_{-84}	1349^{+245}_{-177}	$1.90^{+0.30}_{-0.28}$	$6.63^{+3.81}_{-2.14}$	$3.71^{+0.60}_{-0.47}$
A2261	$-1.07^{+0.19}_{-0.24}$	-1.09 ± 0.20	$89.6^{+7.4}_{-5.4}$	$52.2^{+8.3}_{-7.0}$	696^{+117}_{-77}	1253^{+242}_{-159}	$1.20^{+0.19}_{-0.13}$	$5.27^{+3.00}_{-1.59}$	$7.59^{+1.15}_{-0.99}$
Sim. Cluster 2	$-1.49^{+0.22}_{-0.20}$	-1.48 ± 0.21	$35.3^{+6.9}_{-10.7}$	$1414.5^{+2223.6}_{-910.6}$	509^{+78}_{-62}	1115^{+156}_{-143}	$0.16^{+0.07}_{-0.04}$	$3.28^{+1.39}_{-0.99}$	$8.24^{+1.22}_{-1.10}$
Sim. Cluster 3	$-1.46^{+0.20}_{-0.17}$	-1.45 ± 0.19	$37.3^{+10.1}_{-7.6}$	$222.9^{+787.1}_{-123.1}$	549^{+138}_{-90}	1124^{+153}_{-162}	$0.35^{+0.77}_{-0.20}$	$3.37^{+1.37}_{-1.11}$	$8.17^{+1.19}_{-1.11}$

Figure 3.10: Posteriors for relative velocity $V_{\rm rel} = 4\sigma_0/\sqrt{\pi}$ (horizontal axes, in km/s) versus SIDM cross section (vertical axes, in cm^2/g) from the SIDM halo matching model. The smaller boxes are for the clusters individually, and the larger box at the bottom is the composite of all 10 clusters. The 2σ regions are shaded, and the 1σ regions are shown in a darker shade.

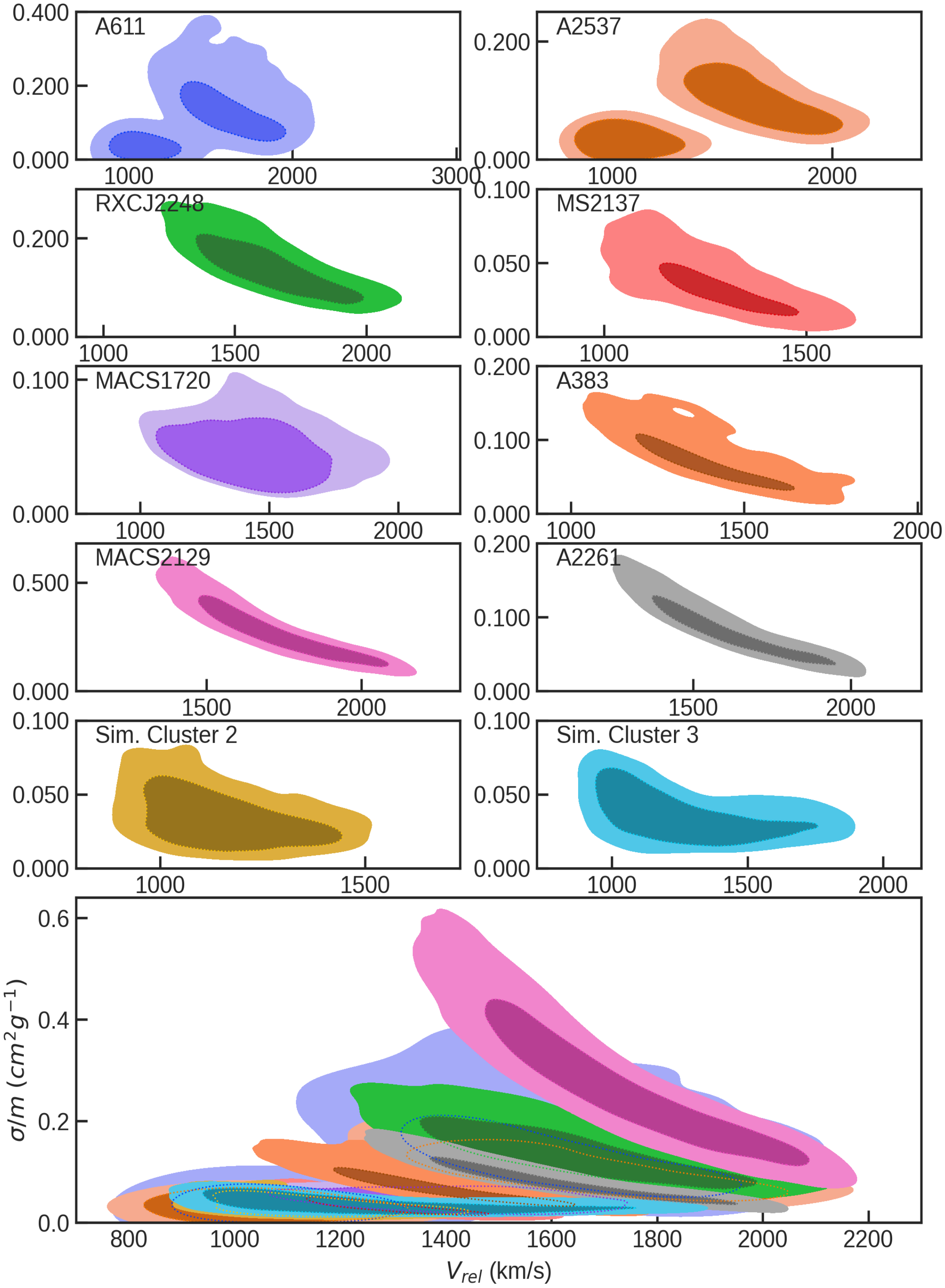
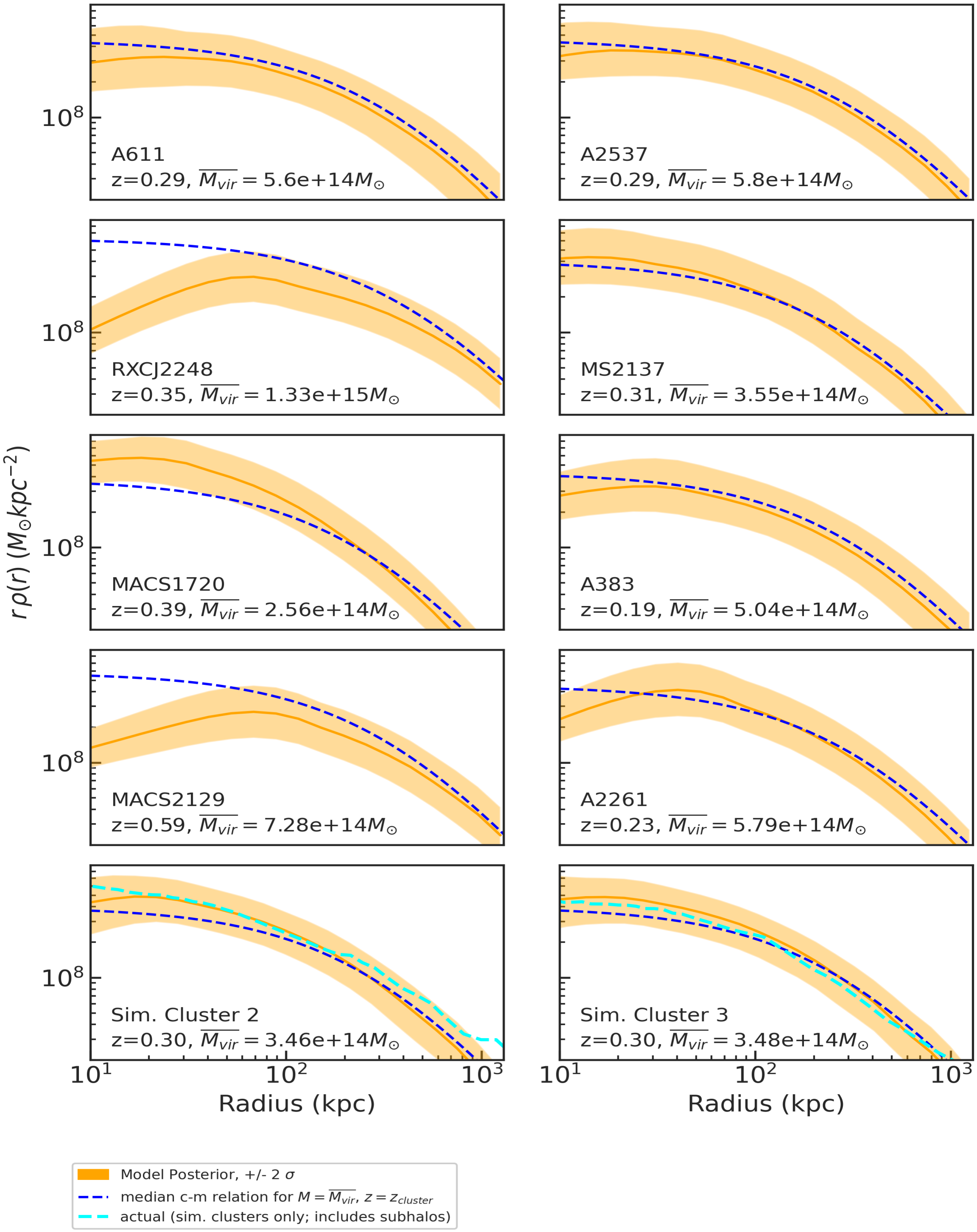


Figure 3.11: Posteriors from the SIDM model for $r\,\rho(r)$ (vertical axes, in $M_\odot$ kpc^{-2}) versus radius (horizontal axes, in kpc). The median posterior and 2σ band are shown in orange. The dashed blue line shows the median concentration-mass relation found from Diemer and Joyce [45] using the redshift and the average virial mass of the specific cluster. For the simulated clusters, the actual $r\,\rho(r)$ is shown, but note that this includes subhalo density and is not exactly comparable to the other lines.



(orange line and bands), and allows comparison to several other curves. We plot $r\,\rho(r)$ rather than density to allow better data visualization. The dashed blue line represents the median concentration-mass relation from Diemer and Joyce [45], accounting for the redshift and using the median virial mass of each specific cluster. Sim. clusters 2 and 3 are somewhat more dense than that of the median concentration-mass relation of Diemer and Joyce [45]. The cyan dashed line shows the actual $r\,\rho(r)$ for the simulated clusters, but note that this data includes subhalo DM, and so is not exacly comparable to the other lines, especially in the outer parts of the halo. Our results show notable cores (i.e., a positive slope to the $r\,\rho(r)$ curve in the inner radii) in clusters RXCJ2248, MACS2129, and A2261.

Histograms for the inferred SIDM cross section per unit mass for each cluster are shown in Figure 3.9, and the tabulated results are summarized in Table 3.4. Cross sections for the simulated clusters are inferred to be 0.037 ± 0.019 cm^2/g and 0.039 ± 0.020 cm^2/g for Sim. Clusters 2 and 3, respectively, at the 68% confidence level.

The median inferred cross section for the observed clusters ranges from 0.04 cm^2/g (MACS1720) to 0.28 cm^2/g (MACS2129). The inferred values show $\sigma/m < 0.23$ cm^2/g (68% C.L.) for all clusters except MACS2129, which shows $\sigma/m < 0.44$ cm^2/g at the 68% C.L.

Our inferred cross sections per unit mass versus mean relative particle speed are shown in Figure 3.10. We calculate the central relative velocity between particles as $V_{\rm rel} = 4\sigma_0/\sqrt{\pi}$, assuming a Maxwellian velocity distribution. Using the same method as that described above in Section 3.6.3 to account for systematic error in the sample, but in this case for $V_{\rm rel}$ rather than cross section, we find the relative velocities in the sample are 1458^{+80}_{-81} km/s. However we note that this could be biased low by up to 20%, as discussed in Section 3.6.1. There is an anti-correlation between inferred cross section and $V_{\rm rel}$, as can be seen in Figure 3.10 and Table 3.4. Note that the trend when the results from all the clusters are put together is that higher cross section inferences go with higher $V_{\rm rel}$ and higher σ_0. This is likely because a higher cross-section results in a larger interaction rate and, in turn, a larger r_1 (characteristic

Figure 3.12: Comparison of the BCG mass inferred (top panel) and $c_{\rm vir}$ (bottom panel) from strong lensing with that from the SIDM model. The error bars indicate the 68% confidence interval. The diagonal lines represent equality.

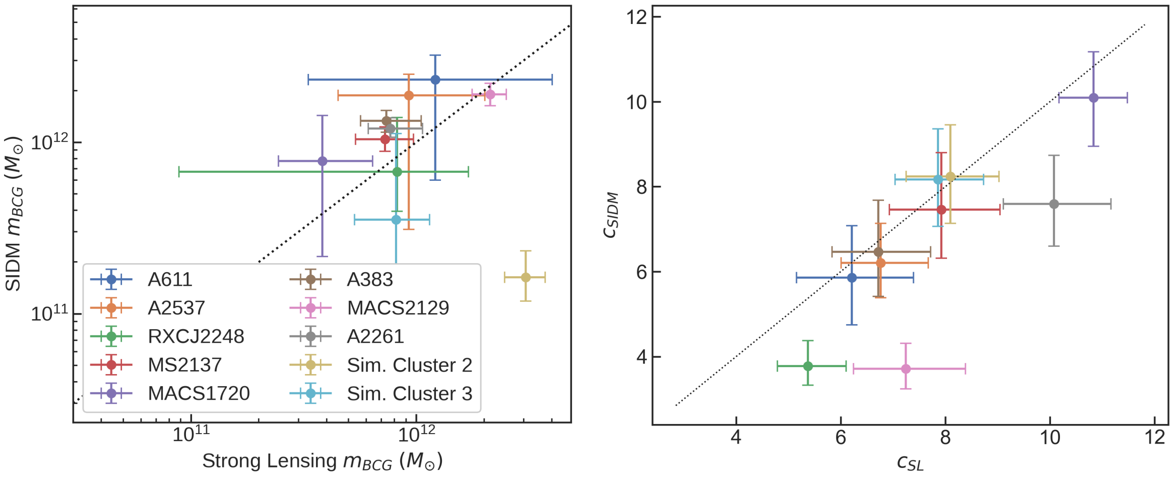
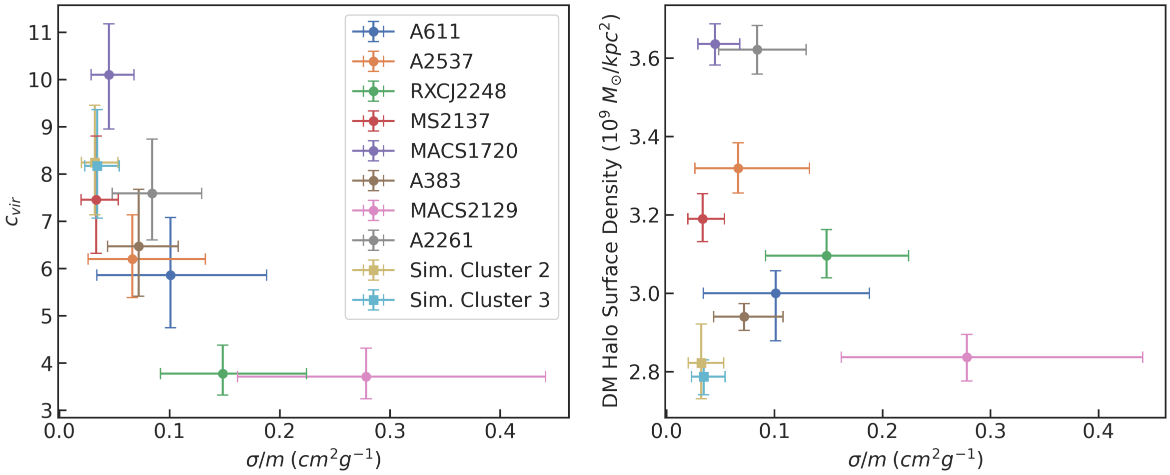


Figure 3.13: *Top:* Concentration versus inferred cross section. The error bars indicate the 68% confidence interval. The square markers are the simulated clusters. *Bottom:* DM halo surface density at a radius of 28 kpc versus inferred cross section.



radius) and greater average particle speed within r_1.

The cross section posteriors for the observed clusters are shown in Fig. 3.9. It seems clear by eye that there is some support for a common cross section around 0.1cm^2/g. We will not simply average the results, but allow for a systematic error in inferring the underlying cross section in Section 3.6.3.

The top panel of Figure 3.12 shows the BCG masses for each cluster inferred from strong lensing compared to the inference results from the SIDM model. The correlation between the the two models is good, with the notable exception of Sim. Cluster 2. We reran the SIDM analysis for that cluster with a strong Gaussian prior on BCG mass: $\log{(m_{\rm BCG}/M_\odot)} = 12.54 \pm 0.1$ (i.e., the value inferred from the strong lensing analysis of that cluster). Our results were robust under that test, showing a only very small increase in cross section.

The bottom panel of Figure 3.12 shows the concentration for each cluster inferred from strong lensing compared to the inference results from the SIDM model. For the strong lensing cNFW fits, we defined $c_{\rm vir} = r_s/r_{\rm vir}$, which should be comparable to the NFW $c_{\rm vir}$ when the cNFW core radius $r_c \ll r_s$. The SIDM model allows for elongation along the LOS by allowing the axis ratio s to vary, resulting in generally lower concentrations than that inferred in the SL analysis. This in turn results in lower inferred central densities and higher SIDM cross sections.

In Figure 3.13 we show the concentration inferred in the SIDM model for each cluster versus the SIDM cross section in the top panel. There is a large scatter in the concentrations from about 3 to 10, while the cross sections are scattered around 0.1 cm^2/g. Note that MACS2129 has a high median cross section and a low concentraion. This suggests that carefully combining strong and weak lensing data with a tailored analysis such as what we have done could be fruitful. The bottom panel shows DM halo surface density at 28 kpc versus inferred SIDM cross section. The observed clusters are systematically higher in surface density than the simulated clusters. Among the observed clusters there does not seem to be a correlation between surface density at a particular radius and cross section, although cluster MACS2129 does appear to have the highest inferred cross section while its surface density is among the lowest in the group.

Besides having the highest inferred SIDM cross section, Cluster MACS2129 is unusual in several regards. It has the highest redshift in our sample (0.589), and is the only cluster for which our lens model was unable to reproduce one of the data images. MACS2129 also has the largest DM halo core inferred from strong lensing, at $54.8^{+23.0}_{-17.6}$ kpc, which is directly relevant for setting the cross section. In addition, this cluster's mass-to-light ratio is an outlier, at $\Upsilon_V = 10.49^{+1.89}_{-1.75}$, while the inference for the other 7 clusters is for $1 < \Upsilon_V < 4$. As discussed in Section 3.5, we investigated the impact of constraining Υ_V to a value of 4 for MACS2129, and it would indeed result in a smaller core size (a median of 39.4 kpc compared to the original 54.8 kpc) but the fit of the model was significantly inferior to the original. A smaller core would lead to a lower inferred SIDM cross section for that cluster. It is possible that a better strong lensing model is required for MACS2129 but it is not clear what form that would take.

We also compared the inferences for the virial mass of the DM halo, $m_{\rm vir}$, from the strong lensing fits and the SIDM halo matching results. Our results quoted in Tables 3.3 and 3.4 show that the estimates are largely consistent with one another at the 10% level. The exceptions include the two highest mass systems, RXCJ2248 and MACS2129, that had virial masses inferred from the SIDM halo matching procedure that were about 30% higher than those from the cNFW fit to the strong lensing data. There is a systematic trend for the SIDM halo matching procedure to prefer slightly higher halo virial masses, but mostly within individual error bars.

3.6.3 Constraints on the Cross Section

After having obtained posterior chains that include cross section per unit mass ($sm = \sigma/m$) for the simulated and observed clusters, we turn to estimating the true underlying cross section. One option is to average the cross sections obtained for the observed clusters. However, this would not be appropriate without checking if the inferred cross sections are statistically consistent with each other. Therefore, we estimate the average cross section from the sample by allowing for a systematic error.

This systematic error can also allow us to (partially) account for effects not included in our analysis, including variations of the radial profile and shape away from our cored elliptical NFW profile model, and variations of subhalo mass profile away from the scaling relations used to model them. We also recall that the work of Sokolenko et al. [179] showed a spread of about 10-20% in the inferred core sizes for small cross sections ($\sigma/m < 1$ cm^2/g). We expect that this will translate into roughly constant error in $\log(\sigma/m)$. We also explored one systematic effect related to the SIDM halo profile – the elongation of halos along the line of sight by assuming an ellipsoidal halo profile that is oriented along the LOS. A more sophisticated non-spherical SIDM halo profile may be able to bring the cross section inferences for the different clusters into better agreement with each other.

For the observed clusters, we denote the means and standard deviations for the posteriors of $\log_{10}[(\sigma/m)/(\mathrm{cm}^2/\mathrm{g})]$ from the SIDM analysis as μ_i and ϵ_i, with $i = 1, ..., 8$. These are listed in Table 3.4. We used $\log_{10}(\sigma/m)$ because its posterior is more Gaussian for all the clusters compared to the posteriors for σ/m. We then add a common systematic error $\epsilon_{\mathrm{model}}$ in quadrature to the inferred errors of all the clusters. The combined likelihood can be written as,

$$L = \prod_i N\left(\mu_i | \mu_{\mathrm{true}}, \sqrt{\epsilon_i^2 + \epsilon_{\mathrm{model}}^2}\right) \tag{3.9}$$

where $N(\mathrm{x} \mid \text{mean, standard deviation})$ denotes the PDF of a normal distribution with specified mean and standard deviation. This allows us to estimate the true underlying value of $\log_{10}[(\sigma/m)/(\mathrm{cm}^2/\mathrm{g})]$ denoted by μ_{true} simultaneously with $\epsilon_{\mathrm{model}}$. We adopt uniform priors between -2 and 0 for μ_{true} (which means the true cross section is between 0.01 and 1 cm^2/g) and a uniform prior between 0 and 2 for $\epsilon_{\mathrm{model}}$.

We inferred $\mu_{\mathrm{true}} = -1.086^{+0.125}_{-0.125}$, which implies that the average SIDM cross section (=

$10^{\mu_{\rm true}}$) is $0.082^{+0.027}_{-0.021}$ cm^2/g (68% C.L.), and it is less than 0.13 cm^2/g at the 95% confidence level. We inferred a model (systematic) error ($\epsilon_{\rm model}$) of 0.27 dex. This common error is larger than the individual errors from the fits in Table 3.4. Thus, we can approximate each cluster as having an error of about 0.27 dex. Within this approximation, the error on the mean is $0.27/\sqrt{8} \simeq 0.1$ dex, and hence the 2σ upper bound on σ/m should be close to $0.082 \times 10^{0.2}$cm^2/g $= 0.13$cm^2/g, which it is. We repeated our analysis using means and standard deviations for σ/m rather than $\log_{10}(\sigma/m)$, and an analysis with a normal distribution for σ/m truncated below zero— the results of the three methods were consistent with one another.

In comparison, a constraint on SIDM σ/m of $0.10^{+0.03}_{-0.02}$ cm^2/g for galaxy clusters was reported in Kaplinghat et al. [80]. This was based on data from 6 clusters from Newman et al. [133, 132], which employed strong lensing, weak lensing and stellar kinematic approaches to determine the cluster profiles. In Sagunski et al. [165], they use a Jeans analysis technique and find $\sigma/m < 0.35$ cm^2/g (95% C.L.) for galaxy clusters. Cluster mergers can also put constraints on SIDM cross section, as collisionless DM would not be slowed in a cluster merger. In Randall et al. [150] they analyze the collision in the Bullet Cluster (galaxy cluster 1E 0657-56) and find $\sigma/m < 1.25$ cm^2/g at the 68% confidence level. In Robertson et al. [158], they use simulations on the same cluster and find a more relaxed constraint of 2 cm^2/g in their fiducial model. Harvey et al. [66] analyzed an ensemble of 72 clusters to find $\sigma/m < 0.47$ cm^2/g at the 95% confidence level, but see Wittman et al. [202]. Analyses of other cluster collisions that put an upper constraint on SIDM cross section include $\sigma/m < 4$ cm^2/g in MACS J0025.4-1222 [17], $\sigma/m < 3$ cm^2/g in Abell 2744 [113] and $\sigma/m < 7$ cm^2/g in cluster DLSCL J0916.2+2951 [39]. Harvey et al. [68] found $\sigma/m < 0.39$ cm^2/g at the 95% confidence level based on BCG oscillations in otherwise relaxed galaxy clusters.

The limits we have obtained are very stringent, as the comparison above shows. Unlike previous work, our dedicated SIDM analysis used the full radial surface density profile, allowing for stronger constraints. The fact that we selected clusters with masses close to $10^{15} M_\odot$ and with images over a wide range of radii also likely played a part in this. We can test this hypothesis with more heterogeneous cluster data sets in the future. One may worry that our reliance on the analytic model down to such low cross sections is leading to overly stringent constraints. As we discussed previously, the model has only been validated for cross sections of 0.1 cm^2/g and higher by Sokolenko et al. [179]. However, the cNFW fits to the strong lensing data recovered uniformly high surface densities at 30 kpc for the clusters, close to or larger than that of the simulated clusters. These high surface densities directly lead to the strong constraints on the cross section. A related issue is that the analytic model may not be capturing the effects of BCG stellar distribution on the SIDM density profile well enough at low cross sections because the inner halo is not close to isothermal. However, it is not clear why there would be a bias to lower cross sections in this case. Note that the inferred r_1, despite the small cross sections, is in the range of 30 to 140 kpc, where the cluster is dark matter dominated. This is again related to the high dark matter densities required to model the strong lensing images. It is also useful to keep in mind that the impact of the baryons on the SIDM density profile of these $\sim 10^{15} M_\odot$ clusters is much smaller than lower mass clusters (which would have almost similar stellar masses). Nevertheless, we cannot rule out that systematic errors of order 0.1 cm^2/g are introduced by the method, and this needs to be investigated using simulations.

3.7 Conclusions

With the aim of constraining the self-interaction cross section of DM particles, we have constructed strong lensing models of 8 observed clusters (see Figure 3.2). We included DM subhalos for the perturbing galaxies to more accurately model the perturber mass distribution. The inferred surface density from strong lensing for the main halo was fit with a

SIDM profile to infer the self-interaction cross section. For the SIDM profile, we used a well-tested analytic model, in which the outer region follows a NFW profile and the inner region is isothermal due to DM self-interactions, with the transition radius being set by the self-interaction cross section. We allowed the SIDM halo profile to be elongated along the line-of-sight with a cosmological prior to regulate axis ratios larger than 2. We used the posteriors for the cross sections inferred from all the clusters to infer the true underlying SIDM cross section, allowing for a common systematic error due to mismodeling. We have tested our inference pipeline on two mock data sets obtained from the Illustris-TNG simulation. Our key findings are summarized below.

Using strong lensing alone we were able to reproduce the image positions in our cluster sample with RMS image position errors ranging from $0''.32$ to $1''.07$. Our methodology includes separate lens model components for the member galaxy's baryonic and DM components, allowing for more flexible characterization of each galaxy's matter distribution. The models reproduced all data images, with the exception of one image in the outlier MACS2129.

We find that the strong lensing inference for concentration was biased high for the two simulated relaxed clusters (see Table 3.1) compared to the cosmological expectation for spherically-averaged halos. For the observed clusters we find a median concentration of 7.0 from strong lensing, which is also higher than would be expected from the concentration-mass relation, and higher than that found in X-ray observations of CLASH clusters. As discussed in Appendix B.3, this may be expected in situations where the line of sight is preferentially oriented along the major axis of halos, which boosts strong lensing probability. In the common case of prolate halos with their major axis oriented along the line of sight, halo concentration can be biased by up to $\sim 60\%$. Previous works [e.g., 133] have also noted that observed concentrations for a sample of clusters similar to ours are indeed higher than what is predicted from mass-concentration relations in CDM simulations. In our SIDM halo model, we allow the halos to be elongated along the line of sight, which reduces the inferred concentrations when fitting the SIDM halo model to the inferred surface density profiles.

We found that the BCGs centers in our models are either coincident with the DM halo centers or offset by few to 10 kpc (Figure 3.5). The median offset is about 4 kpc, as measured in the plane of the sky, which would correspond to approximately 5 kpc in 3D distance assuming isotropy. Harvey et al. [68] used simulations to conclude that the distribution of offsets from an ensemble of clusters would exhibit a median value of 3.8±0.7 kpc for a CDM scenario (i.e., $\sigma/m = 0$), and 8.6±0.7 kpc in a scenario where $\sigma/m = 1\ \mathrm{cm}^2/\mathrm{g}$. Comparing to Harvey et al. [68], the median value we obtained is consistent with $\sigma/m \ll 1\ \mathrm{cm}^2/\mathrm{g}$. However we also note that Harvey et al. [67] found that MCMC lensing code inferences of offset are susceptible to substantially understating the error in position offset, thus tempering the ability to make strong inferences from the offsets.

Using the SIDM halo model, the cross section per unit mass for all observed clusters had median values in the range of 0.034 cm^2/g to 0.15 cm^2/g (Figure 3.9 and Table 3.4), with the exception of outlier MACS2129 at 0.28 cm^2/g. We combined these individual measurements allowing for an unknown common modeling error, and inferred an SIDM cross section of $0.082^{+0.027}_{-0.021}\ \mathrm{cm}^2/\mathrm{g}$ (68% C.L.), with an upper limit of 0.13 cm^2/g at the 95% C.L. (see Sections 3.6.3). We infer a systematic (modeling) error of 0.27 dex. In comparison, a constraint on SIDM σ/m of $0.10^{+0.03}_{-0.02}\ \mathrm{cm}^2/\mathrm{g}$ for galaxy clusters was reported in Kaplinghat et al. [80] using a simpler analysis, and Sagunski et al. [165] found $\sigma/m < 0.35\ \mathrm{cm}^2/\mathrm{g}$ (95% C.L.) for galaxy clusters using Jeans analysis. The mean relative velocity of dark matter particles for the eight clusters we have analyzed is 1458^{+80}_{-81} km/s. Since relative particle speeds are much lower in galaxies, the cross sections for dark matter interactions can be larger in galaxies, as many concrete particle physics models predict [187].

Acknowledgements

We gratefully acknowledge a grant of computer time from XSEDE allocation TG-AST130007. QM was supported by NSF grant AST-1615306 and MK was supported by NSF grant PHY-1620638.

Chapter 4

Core Densities of Classical Milky Way Dwarf Galaxies: a Distribution Function Approach

4.1 Introduction

The prevailing theory of the evolution of the universe, ΛCDM, is quite successful in predicting the large scale structure we observe. However, at scales below ~ 1 Mpc, discrepancies between prediction and observation begin to emerge [21]. At that scale, we observe galaxies whose central dark matter (“DM") density profiles are shallower (or less “cuspy") than that predicted by ΛCDM. We also see fewer dwarf galaxies than we would expect, unless an abnormal fraction of the DM subhalos have no visible baryonic component. Further, the mass function of the satellites that we do observe is much lower than the prediction of ΛCDM. These puzzles are known as the "cusp/core" [120, 58, 82], "missing satellite" [83, 87] and "too big to fail" [15, 81] problems, respectively. Dwarf galaxy satellites of the Milky Way

("MW") can be helpful in resolving these small-scale problems. They are typically dominated by DM, making the analysis easier, and they are close enough to permit observation of some elements of the trajectories of individual stars. These trajectories can be used to characterize the distribution of the non-visible DM halo that hosts them.

While ΛCDM predicts a cuspy "NFW" profile, with $\rho \propto r^{-1}$, in the inner region [125], observations of nearby dwarf galaxies have been inconclusive with regard to the shape of their inner DM profile. The data have largely been shown to be consistent with either cusped or cored DM profiles.

Wheeler et al. [201] found strong evidence that most MW dwarf spheroidal galaxies ("dSphs") are dispersion supported rather than rotationally supported. The spherically symmetric Jeans equation [10] has been used extensively to study dispersion supported systems such as these. Unfortunately, it suffers from a well-known degeneracy between mass and velocity anisotropy. There have been many efforts to overcome this [10, 101, 43, 157]. Read et al. [152] used a Jeans equation solver and fourth-order velocity moments to examine the inner densities of MW classical dwarf galaxies. Kaplinghat et al. [81] used Jeans analysis coupled with fourth-order velocity moments to predict the inner densities of bright MW dSphs. Walker and Penarrubia [198] used distinct chemodynamic populations in Fornax and Sculptor to constrain the inner slope of the DM profile, and Pace et al. [138] did the same for the Ursa Minor dwarf spheroidal. Read et al. [153] used N-body simulations to study core formation in dwarf galaxies, concluding that supernova feedback causes cores to form if star formation proceeds long enough. Schwarzchild [175] described a numerical modeling approach for triaxial stellar systems, which allow flexible configurations and do not require assumptions about orbital anisotropy. These models have been used by others to model MW dSphs [18, 91, 75, 76, as examples].

An alternative approach is to use a phase space distribution function ("DF") in six dimensions as a tool for characterizing the orbits of the stars in their DM potential. This approach allows for flexible forms for the stellar distribution, and can also allow consideration of velocity moments above second order, potentially mitigate the mass-anisotropy degeneracy. As examples, Wu and Tremaine [204] used distribution functions to derive the mass distribution of Messier 87 using its globular clusters as tracers. Strigari et al. [184] used an approximate DF model to examine the DM profile Sculptor dwarf galaxy. In that work, they did not consider higher-order velocity moments but did consider two different chemodynamic populations.

We use a DF approach in this work to determine the inner DM characteristics of dSphs. First, we examine 24 data sets of various configurations from the Gaia Challenge project [154]. We use those data sets to validate our modeling method. We then examine the bright dSphs of the MW: Draco, Fornax, Sculptor, Carina, Sextans, Leo I, Leo II, Ursa Minor and Canes Venatici I. Our aim is to constrain the DM characteristics in the inner regions. We intend to follow this study with an enhanced model that considers two distinct chemodynamic populations for the observed targets where justified by metallicity data. We we believe this may allow tighter constraints on the resulting inferences.

4.2 Distribution Function Models

In this section we describe our approach in modeling the dwarf galaxy stellar and DM distributions. It is possible to describe the position and velocity of stars (or other objects) in a galaxy using a phase space distribution function in six dimensions, three for position and three for velocity. Our intent is to use DFs to analyze dSphs of the MW, using the stars as tracers to determine the DM distribution.

We define a Cartesian coordinate system, centered on the galaxy center, with the z-axis along the line of sight to the galaxy. The projected radius of a star as seen from the observer is then $R = \sqrt{x^2 + y^2}$. An individual star will have have a position coordinate $\mathbf{x}$, given by coordinates (x, y, z). The star will have a velocity vector $\mathbf{v}$, with components (v_x, v_y, v_z). We define η as the angle between the position and velocity vectors.

We can define the distribution function f such that $f(x, v, t)d^3x\ d^3v$ is the probability of finding a star in the infinitesimal volume element $d^3x\ d^3v$. For our purposes we will assume that the DF is constant in time, and becomes $f(x, v)d^3x\ d^3v$. We require that the DF is normalized to unity probability over all phase space, i.e.,

$$\int f(x, v)d^3x\ d^3v = 1. \tag{4.1}$$

Motions of particles like stars in a stationary potential can be determined by the collisionless Boltzman equation. The Strong Jeans Theorem then tells us that solutions to the collisionless Boltzman equation depend only upon three integrals of motion, which may be taken to be the actions H, L and L_z [10].

We consider a potential $\Phi(r)$ that is spherically symmetric. In that case, the DF will depend on the energy E and the angular momentum L. The energy of a star per unit mass is $E(r, v) = \Phi(r) + \frac{1}{2}v^2$, and the angular momentum per unit mass is $L(r, v) = rv\ sin(\eta)$. Several useful quantities can be derived from the DF, including the density profile, the radial velocity dispersion profile and the tangential velocity dispersion profile [10, 184]:

$$\nu(r) = \int d^3\mathbf{v}\ f(x, v) = 2\pi \int_0^\pi d\eta\ sin\eta \int_0^{v_{\rm lim}} dv\ v^2\ f(E, L) \tag{4.2}$$

$$\sigma_r^2(r) = \frac{2\pi}{\nu} \int_0^\pi d\eta\ sin\eta\ cos^2\eta \int_0^{v_{\rm lim}} dv\ f(E, L)v^4 \tag{4.3}$$

$$\sigma_t^2(r) = \frac{\pi}{\nu}\int_0^\pi d\eta\ sin^3\eta\ \int_0^{v_{\rm lim}} dv f(E,L) v^4 \tag{4.4}$$

We define $\nu(r)$ as the probability per unit volume of finding a star at radius r. The number density of stars at radius r is then

$$n(r) = w\ \nu(r) \tag{4.5}$$

where w is the total number of stars in the population. We define the velocity above which stars become unbound as $v_{\rm lim} = \sqrt{2(\Phi_{\rm lim} - \Phi(r))}$, where ϕ and $\Phi_{\rm lim}$ are defined in Section 4.2.1. The total velocity dispersion can be found by combining the radial and tangential components:

$$\sigma_{tot}^2(r) = \sigma_r^2(r) + 2\sigma_t^2(r). \tag{4.6}$$

The projected stellar density Σ_* at a radius R can be found by integrating over the line of sight (LOS):

$$\Sigma_*(R) = 2\int_0^\infty dz\ n(r), \tag{4.7}$$

where $r = z^2 + R^2$. The LOS velocity dispersion can be found from

$$\begin{aligned}\Sigma_*(R)\ \sigma_{\rm LOS}^2(R) &= 2\int_0^\infty dz\ n(r)\frac{z^2\sigma_r^2 + R^2\sigma_t^2}{z^2+R^2} \\ &= 2\pi w\int_0^\pi d\eta \sin\eta \int_0^{v_{\rm lim}} dv v^4 \\ * \int_0^\infty dz \frac{(2z^2\cos^2\eta + R^2\sin^2\eta)}{z^2+R^2} f(E,L)\end{aligned} \tag{4.8}$$

Higher order moments of velocity can also be predicted by this method. We will use a virial shape parameter ("vsp") that is the fourth moment of velocity in our analysis. For our purposes we opt to compute the global vsp rather than one that varies with radius, which helps to minimize noise in the calculation. The derivation of the vsp is presented in Appendix C.1.

4.2.1 DM Potentials

We assume that the mass of the stars are negligible in comparison to that of the dark matter, and so the stars are tracers of the dark matter potential but do not influence it. We will consider three potential/density profiles: “NFW", “cored" and “cNFW". The NFW and cored profiles can be completely described by two parameters, while the cNFW profile has one additional parameter, the core parameter “c". The cNFW profile core parameter $c \equiv r_c/r_s$, where $\log_{10}[r_c/\text{kpc}]$ is the parameter used in the model (which we distinguish from the core radius r_{core}, defined below). We also use the scale radius r_s and scale velocity v_s as specifying parameters for all three profiles. The scale density ρ_s and the scale potential Φ_s are determined via the relation $\Phi_s = v_s^2 = 4\pi G r_s^2 \rho_s$, where G represents Newton’s gravitational constant.

Let $x \equiv r/r_s$. The NFW profile density and potential pair is

$$\rho(r) = \frac{\rho_s}{x(x+1)^2}, \text{ and} \tag{4.9}$$

$$\Phi(r) = \Phi_s \left(1 - \frac{\log(x+1)}{x}\right) \tag{4.10}$$

Note that Φ has been defined so that it is non-negative everywhere, with a value of zero at

$r = 0$, and goes to Φ_s as $r \to \infty$.

Define the peak circular velocity in a potential as $V_{\max}$, and the radius at which the peak occurs as $r_{\max}$. For the NFW profile, it can be shown that $r_{\max} = 2.163\ r_s$, and $V_{\max} = 0.465\ v_s$.

The "cored" profile is a generalized Hernquist profile [71, 208] of the form

$$\rho(r) = \frac{\rho_s}{(x+1)^3}, \text{ and} \tag{4.11}$$

$$\Phi(r) = \Phi_s \frac{x\,(x+2) - 2\,(x+1)\log\,(x+1)}{2x\,(x+1)}$$

The potential in the cored case has a zero value at $r = 0$, and goes to $\frac{1}{2}\Phi_s$ as $r \to \infty$. For the cored case, $r_{\max} = 4.4247\ r_s$, and $V_{\max} = 0.3502\ v_s$.

The cNFW profile is

$$\rho = \frac{\rho_s}{(x+c)\,(x+1)^2}. \tag{4.12}$$

The potential is

$$\Phi(r) = \frac{\Phi_s}{(c-1)^2}\left(\frac{x\,(c-1)}{(x+1)} + (1-2c)\log(x+1) + c^2\log\left(\frac{c+x}{c}\right)\right)$$

This profile reduces to the canonical NFW form for $c \to 0$, and reduces to the cored form when $c \to 1$. The relation for conversion between r_s and $r_{\max}$ (and similarly between v_s and $V_{\max}$) become nonlinear but can be solved numerically.

For all profiles, wee define the core radius r_{core} as the radius at which the DM density falls to one-half of its central value. For the NFW and cored profiles, r_{core} is 0 and 0.26

r_s, respectively. For the cNFW profile, the core radius is a nonlinear function of c and is computed numerically. We define Φ_∞ as the value of the potential as $r \to \infty$.

4.2.2 Stellar DF Form

We take the form of the stellar distribution function to be the product of an energy function and an angular momentum function, which are

$$h(E) = \begin{cases} E^a(E_c^q + E^q)^{d/q}(\Phi_{\text{lim}} - E)^e, & E < \Phi_{\text{lim}} \\ 0, & E \geq \Phi_{\text{lim}} \end{cases} \tag{4.13}$$

$$g(L) = \left(\frac{1}{2}\Big(\big(\frac{L}{L_\beta}\big)^{\frac{b_{\text{in}}}{\alpha}} + \big(\frac{L}{L_\beta}\big)^{\frac{b_{\text{out}}}{\alpha}}\Big)\right)^{\alpha}, \tag{4.14}$$

with α nonnegative for $b_{\text{in}} <= b_{\text{out}}$, and α negative for $b_{\text{in}} > b_{\text{out}}$. The total DF is their normalized product:

$$f(E, L) = n_f \; h(E) \; g(L),$$

and so the total phase space density distribution of w stars is

$$w \; f(E, L) = w \; n_f \; h(E) \; g(L),$$

In these equations, n_f is a normalizing factor that ensures the DF integrates to unity over all phase space, i.e.,

$$n_f = 1/\int f(x, v) d^3x \; d^3v$$

.

These equations are similar to those in [184], except we have inserted a factor of $\frac{1}{2}$ in the angular momentum function to ensure that the function transitions smoothly as α changes sign, which improves the behavior of MCMC computations. The parameter Φ_{lim} is a limiting potential beyond which no stars exist, analogous to a tidal cutoff potential, and we define r_{lim} as the radius at which this cutoff occurs for a particle with zero velocity. The e parameter controls the shape of the tidal cutoff. The parameters a and d control the log-slope of the energy response. E_c is a cutoff energy, below which the log-slope is approximately a, and above which the log-slope is approximately $a + d$. We restrict d such that $d < 0$.

The parameter L_β characterizes the angular momentum scale, and the parameters b_{in} and b_{out} control the inner and outer log-slopes of the angular momentum function, respectively. At angular momenta above L_β, the log-slope is approximately b_{out}, and below L_β the slope is approximately b_{in}. As a result, the parameters b_{in} and b_{out} determine the anisotropy of the system. The anisotropy parameter β is given by

$$\beta(r) = 1 - \sigma_t^2(r)/\sigma_r^2(r). \tag{4.15}$$

If $b_{\text{out}} \approx 0$, then $\beta \approx -\frac{1}{2}b_0$ for $L \ll L_\beta$. Similarly, if $b_{\text{in}} \approx 0$, $\beta \approx -\frac{1}{2}b_1$ for $L \gg L_\beta$.

4.2.3 Approximate Likelihood Function

The full likelihood function that considers the positions and velocity of all stars is discussed in Appendix C.3. A significant problem with the full likelihood function is its intensive computation requirement. For each star, we are required to perform a multi-dimensional integration of our DF. For data sets with hundreds or even thousands of stars, the time to compute the normalized likelihood can be too long to make MCMC analysis appealing, even with supercomputers. To make the model faster to calculate, we employ an approximation method as described below.

Using the equations described in Section 4.2 and Appendix C.1, the DF can be used to make predictions of the radial profiles of surface density and velocity dispersion, and a prediction of the global vsp. We can compare these predictions to observed values from photometry (in the case of surface brightness) or from spectroscopy (in the cases of velocity dispersion and vsp). The surface density and dispersion observations use binned data, with bins at 8 to 25 radial locations, typically. The χ^2 for each characteristic is calculated by comparison of the predicted points with the observed values, relative to the uncertainty in the observation:

$$\chi^2 = \frac{(\mathrm{data} - \mathrm{prediction})^2}{\mathrm{uncertainty}^2}. \tag{4.16}$$

The χ^2 components for each of the three predictions become terms in the model likelihood as described below. The total χ^2 is then

$$\chi^2_{\mathrm{tot}} = \chi^2_{\mathrm{SD}} + \chi^2_{\mathrm{disp}} + \chi^2_{\mathrm{vsp}}, \tag{4.17}$$

where the subscripts refer to surface density, dispersion and virial shape parameter, respectively. We form the log likelihood according to $\log \mathcal{L} = -\chi^2_{tot}/2$. We use a Monte Carlo-based model to derive parameter posteriors. The model employs sampling via the Emcee sampler [59]. Table 4.1 shows the upper and lower parameter limits for the model.

4.2.4 Derived Parameters

Once the parameters specifying the DM potential and the DF are inferred we can calculate distributions of surface density and velocity dispersion at a range of radii, and we can derive other parameters of interest such as the half-light radius $r_{1/2}$, the anisotropy paramter β, the density at 150 pc ρ_{150} and the halo mass M_{200}.

Since the DF model makes a smooth prediction for surface density, calculation of the half-

Table 4.1: Parameter limits in the MCMC model. Units are kpc for r_s and r_c, and km/s for v_s. The units for w are the number of stars in the population. The other parameters are dimensionless. The parameters $\tilde{E_c}$ and $\tilde{\Phi}_{\rm lim}$ are made dimensionless by dividing by Φ_∞ for the distribution being used, and $\tilde{L_\beta}$ is made dimensionless by dividing by $r_s\sqrt{\Phi_\infty}$.

	$\log_{10}(\rm r_s/kpc)$	$\log_{10}(\rm v_s/[kms/s])$	$\log_{10}(\rm r_c/kpc)$	a	q	$\tilde{E_c}$	d	$\tilde{\Phi}_{\rm lim}$	e	$\tilde{L_\beta}$	$b_{\rm in}$	$b_{\rm out}$	α	$\log_{10}(\rm w/stars)$
lower limit	-2.5	0	-2	-4	0.1	0.01	-12	0.01	0.1	0.01	-10	-10	0.1	1
upper limit	1	2.5	1	5	25	1	0	1	10	1	10	10	10	7

light radius $r_{1/2}$ is relatively straightforward. The 2D half light radius $R_{1/2}$ satisfies the equation

$$\frac{\int_0^{R_{1/2}} \Sigma_*(R) R dR}{\int_0^{R_{\rm max}} \Sigma_*(R) R dR} = \frac{1}{2}, \tag{4.18}$$

where $R_{\rm max}$ is the radius of the outermost surface density data point. We verified that using $R_{\rm max}$ rather than an infinite limit did not have a significant effect on the result. This equation can be solved numerically using spline interpolation and iteration for each point in the posterior chain. We then multiply the result by 1.33 to derive the 3D half-light radius. [203] found that the ratio of 1.33 is valid for a variety of stellar profile shapes, and we confirmed this to be a very good approximation for our own mock data sets. We also verified that for the mock data sets, the value obtained by this method was very close to the median radius of the stars in the data set. We use the photometry integration method to calculate the half-light radius posteriors directly from the density predicted by the distribution function (see Equation 4.2). We also calculate $M(< r_{1/2})$, the mass enclosed within the half-light radius.

4.3 Mock Data Modeling

Testing the model with mock data allows us to validate our approach and provides an indication of what we can reliably infer from the model. We use mock data from the Gaia Challenge spherical data sets [154].

4.3.1 Mock Data Characteristics

The stellar density profile in the mock data is given by a generalized Hernquist profile [71, 208]:

$$\nu_*(r) = \nu_0 \left(\frac{r}{r_*}\right)^{-\gamma_*} \left(1 + \left(\frac{r}{r_*}\right)^2\right)^{\frac{(\gamma_* - 5)}{2}} \tag{4.19}$$

The parameter γ_* is set to 0.1 for the cored stellar profile and 1.0 for the cuspy stellar profile. The parameter r_* determines how embedded the star population is placed in the DM potential, and was varied among three values: 0.1 kpc, 0.5 kpc and 1.0 kpc.

The DM potential in the mock data is either "cored" or "NFW", as described in Section 4.2.1. The DM central density ρ_0 is also determined by this choice, with $\rho_0 = 400 \times 10^6 M_\odot kpc^{-3}$ for the cored case and $\rho_0 = 64 \times 10^6 M_\odot kpc^{-3}$ for the NFW case. All of the mock data sets have scale radius $r_s = 1$ kpc. The scale velocity v_s is 147.1 km/s in the cored case and 58.8 km/s in the NFW case.

The stellar velocity anisotropy profile is also varied among two cases. The anisotropy parameter β is defined as $\beta(r) = 1 - \sigma_{\text{tan}}(r)/\sigma_{\text{rad}}(r)$, where σ_{tan} and σ_{rad} are the tangential and radial velocity dispersion components, respectively. β is varied according to an Osipkov-Merrit form [10]: $\beta(r) = r^2/(r^2 + r_a^2)$, where r_a is the anisotropy radius. The parameter r_a takes the values of either 1 kpc or 10,000 kpc. A value of 1 kpc creates a profile in which β rises from 0 in the center to 1 in the outer parts, reaching 0.5 at a radius of 1 kpc. A value of $r_a = 10,000$ kpc creates an essentially isotropic profile with $\beta = 0$ everywhere. The mock data sets therefore have $2 \cdot 3 \cdot 2 \cdot 2 = 24$ possible unique configurations.

The Gaia Challenge data sets provide good model validation cases for our model, since certain key parameters are known: r_s, v_s and w. The data sets contain multiple populations. We selected stars from only one population in each set, and did not include non-member foreground stars. The stars were binned into bins with equal number of stars. We found that the data sets typically had a small fraction of stars with very large orbital radii, which made the outer bins very wide and presented computational challenges. To address this, we opted to exclude the outermost stars from the data sets. Stars farther than 5 half-light radii from the center were excluded. Less than 10% of the stars from any data set were excluded in this fashion, typically about 5%. To simulate measurement error in the line-of-sight velocities, Gaussian error was added with a standard deviation of 2 km/s. The data set characteristics are summarized in Appendix Table C.1.

4.3.2 Mock Data Modeling Results

The approximate DF model was applied to the 24 mock data sets, the results of which are presented below. Since we wish to simulate that we do not have a priori knowledge of the DM profile, we used the cNFW profile in the model, in which the core size is a varying parameter. The model found very good fits to the surface density curves, disperion curves and vsp values in all cases, with χ^2 per degree of freedom ≤ 1.3 for all data sets. A typical fit is shown in Figure 4.1.

Figure 4.2 shows the posterior inferences in $r_{\max}$ versus $V_{\max}$ for the 24 mock data sets, with the true value shown as an "x" near the centers. We used GetDist [96] for two-dimensional plots. The models have a wide diversity of shapes in the $r_{\max}$ - $V_{\max}$ plane, depending on the various profiles for DM density, stellar density, anisotropy and "embeddedness" (i.e., the depth of the stars in the DM potential). The bottom row of Figure 4.2 shows how the embeddedness impacts the shape of the posteriors, the degeneracy characteristics between the two parameters, and the inference capability. We found that the highly embedded data sets

Figure 4.1: Typical fits to surface density (top), velocity dispersion (bottom) and vsp (bottom, inset), in this case for mock data set 10 (ID bcbO_2349). The best fit DF solution is shown in green.

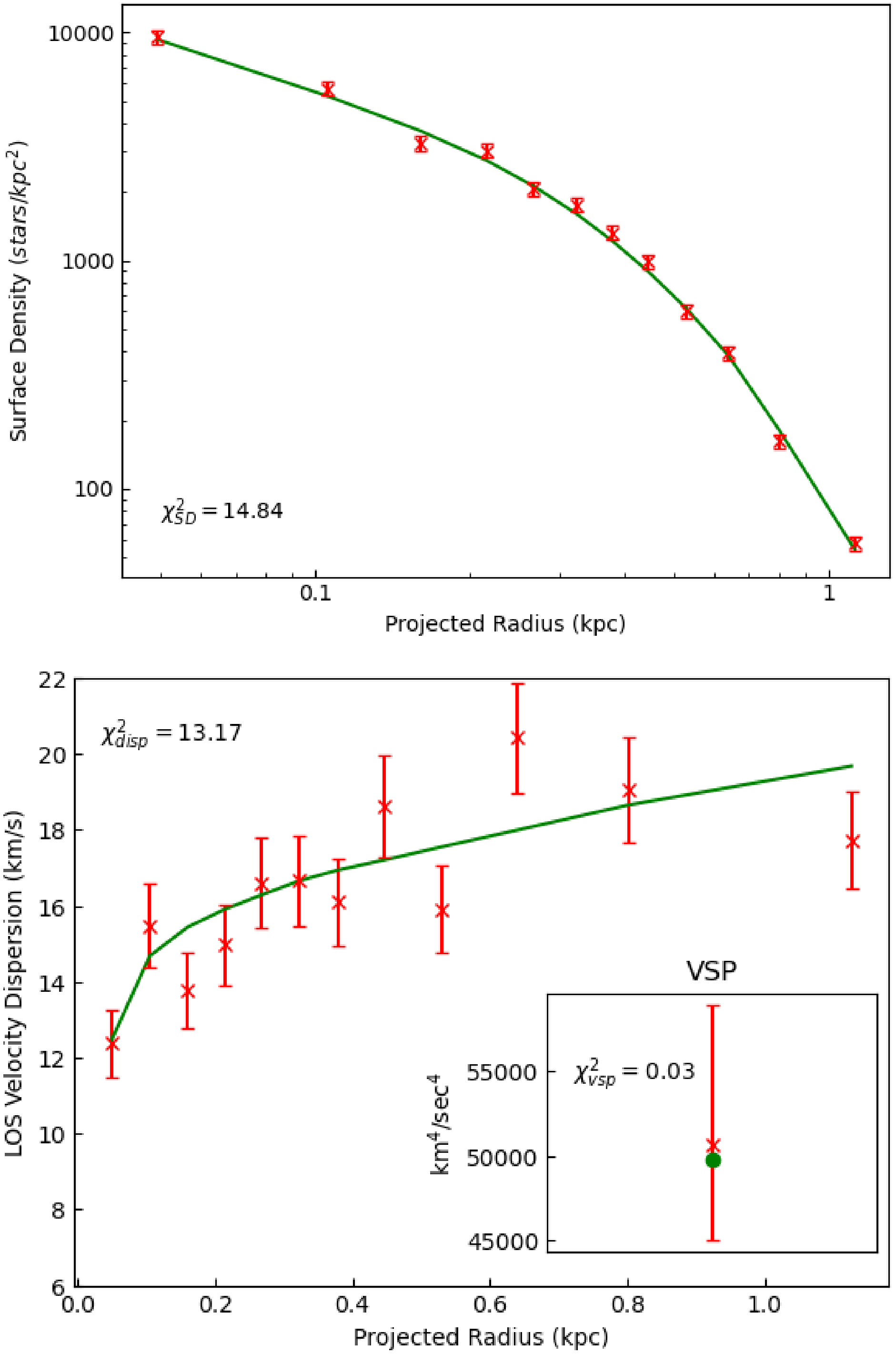


Figure 4.2: Posteriors for $r_{\max}$ vs. $V_{\max}$ for the mock data sets. The 68% and 95% levels are shown, with the 68% level in a darker color. The black "x" indicates the true value. Top row: color-coded by data set. Bottom row: The same data, color-coded by their embeddedness in the DM halo. Right column: Cored profile. Left Column: NFW profile.

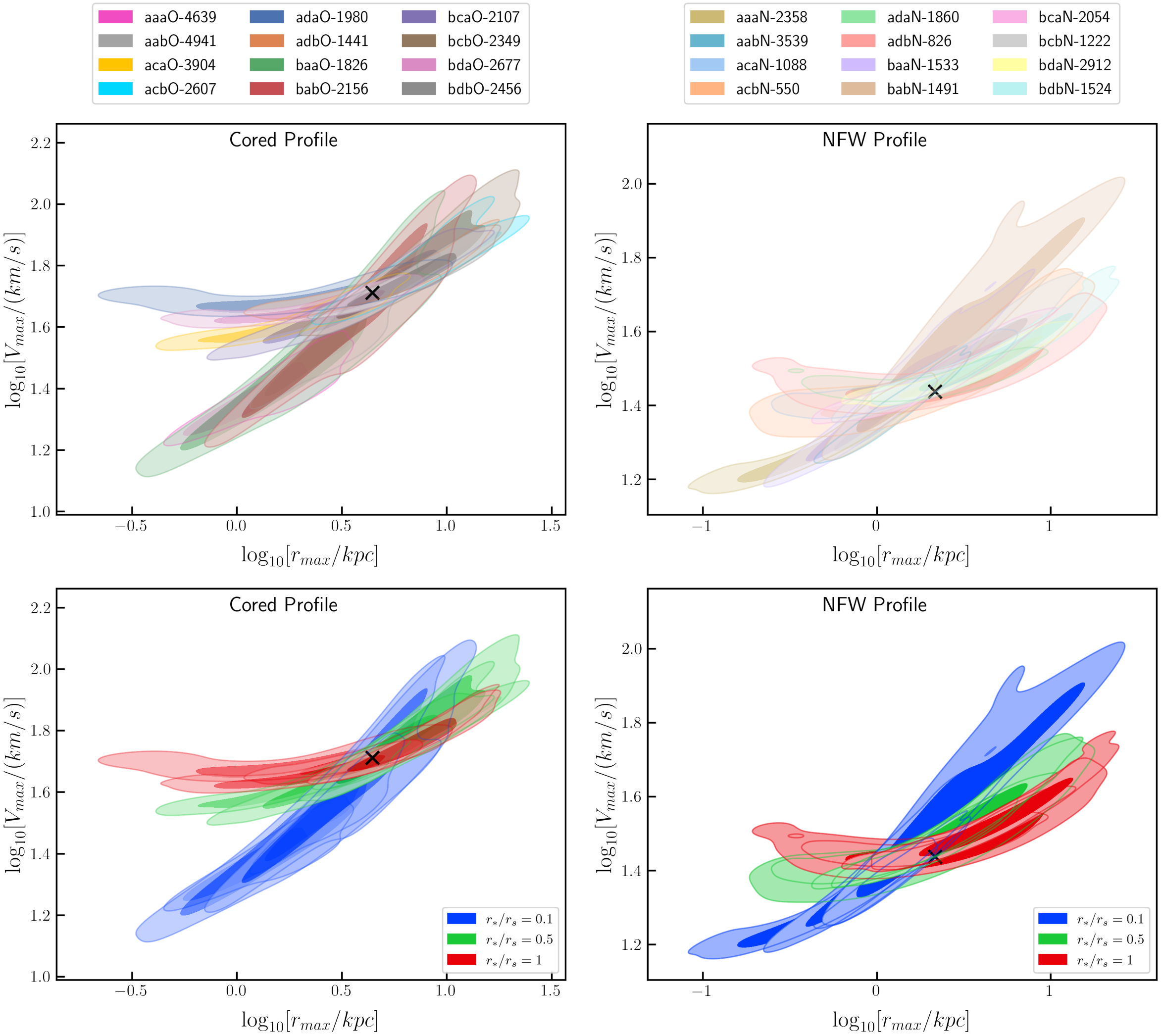
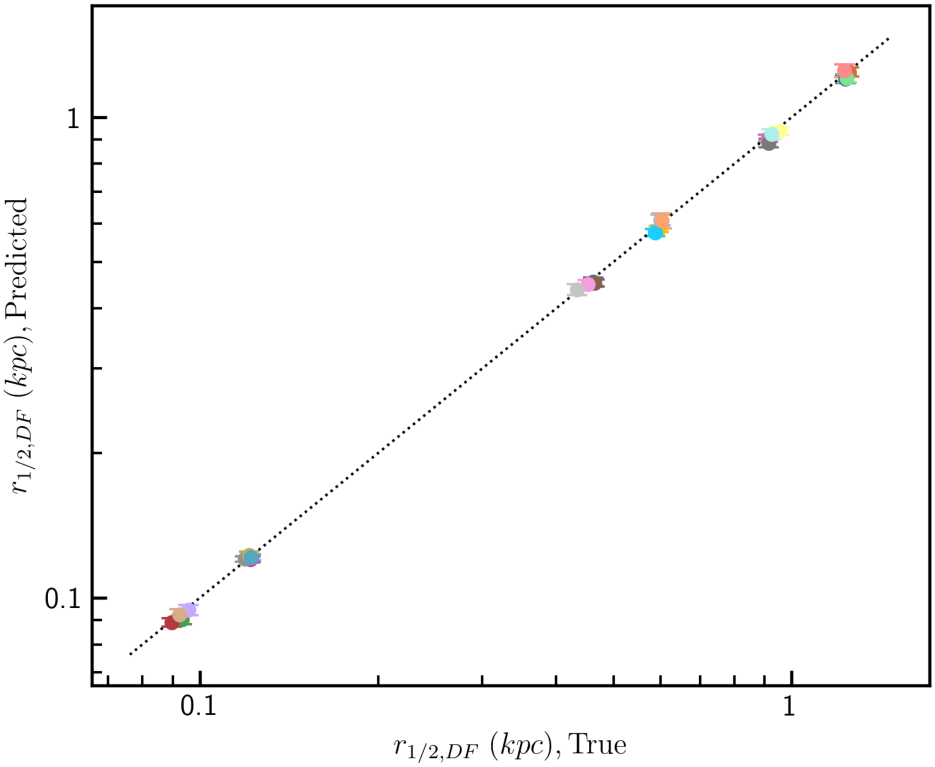


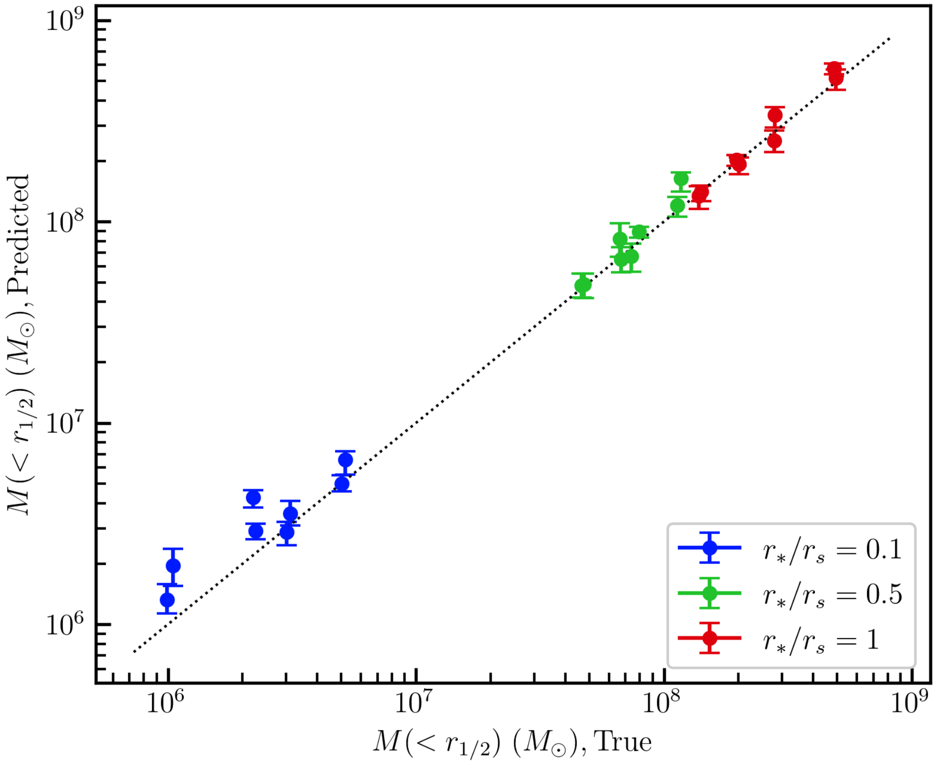
Figure 4.3: Predicted and true values for the half-light radius for the 24 mock data sets, using the same colors as Figure 4.2. The predictions are determined from the DF (see Equation 4.2) by calculating and integrating the surface density curve, finding the radius that yields have the total value (see Equation 4.18). The error bars indicate the 68% confidence interval. The true value is taken to be the median radius of the stars in the given data set.



$(r_*/r_s = 0.1)$ were the least accurate in their inferences of $r_{\max}$ and $V_{\max}$, and that tendency carried over into inferences of other parameters. The model made reliable inferences for the data sets with $r_*/r_s >= 0.5$. The reason for the difference is that the highly embedded data sets do not trace the potential near the scale radius r_s, and so have limited accuracy in that region.

Figure 4.4 compares the posterior for the calculated half-light radius to the true value, which is taken to be the median radius of the stars in the data set. The accuracy is very good, with a difference of less than 2% between the median prediction and the true value for all data sets.

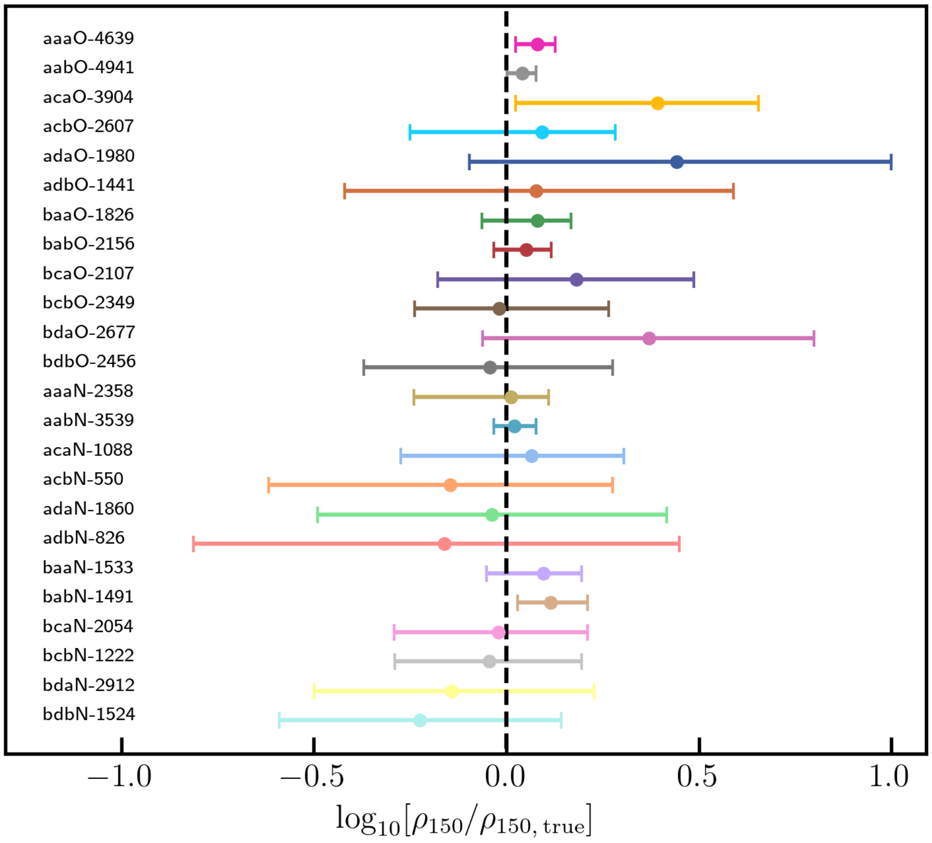
Figure 4.4: Predicted and true values for the mass within the half-light radius for the 24 mock data sets, color-coded by their embeddedness in the DM halo. The error bars indicate the 68% confidence interval. The diagonal line indicates equality.



The mass within the inferred half-light radius can be determined for the cNFW profile by using the posterior values for r_s, v_s and r_c. Figure 4.4 shows the true and predicted values for the mass within the half-light radius for the mock data sets. The predictions are fairly accurate for the data sets with $r_*/r_s \geq 0.5$, i.e., those not deeply embedded in the DM potential. For the data sets with the lowest mass enclosed (and correspondingly very deeply embedded in the DM halo), the model tends to systematically overestimate the mass enclosed.

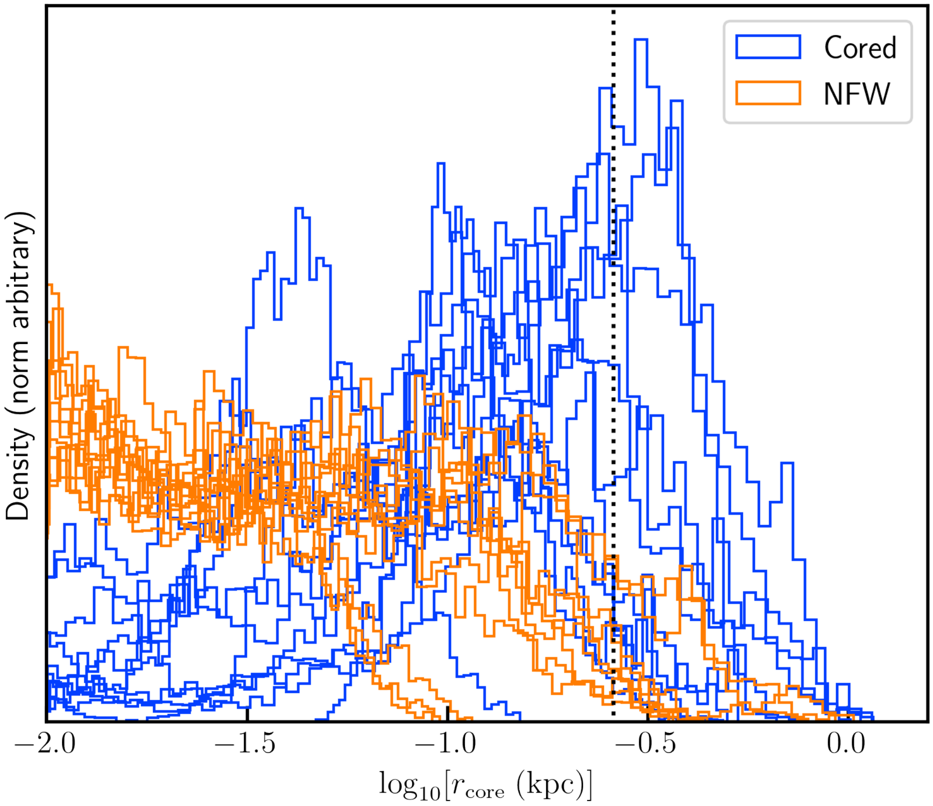
Predictions for the density at 150 pc as compared to their true values are shown in Figure 4.5. The median predictions are generally within 0.3 dex of the true value, with one case near 0.5 dex. In four cases the true values were outside the 95% confidence level of the posterior, all of which were over-estimations of the density.

Figure 4.5: Predicted and true values for DM density at 150 pc for the 24 mock data sets, using the same colors as in Figure 4.2 . The error bars indicate the 95% confidence interval.



The parameter $\log_{10}[r_c/\mathrm{kpc}]$ is allowed to vary in the model to explore the best fitting value, with prior limits $-2 < \log_{10}[r_c/\mathrm{kpc}] < 1$. As described in Section 4.2.1, the core radius r_{core} is calculated as the radius at which the density falls to one-half its central value. The true core radii of the mock data sets are either 0 kpc (NFW) or 0.26 kpc (cored), corresponding to $\log_{10}$ values of $-\infty$ and -0.585, respectively, although we use -2 as a practical lower limit, corresponding to $r_{\mathrm{core}} = 0.01$ kpc. Figure 4.6 shows a composite plot of all 24 posteriors, with blue for cored profiles and orange for NFW profiles. The model shows some ability to distinguish between the two profiles, with an uncertainty of approximately 0.5 to 1.0 dex, although there is bias towards lower values for the cored profiles. This is further illustrated in Figure 4.7 which shows the true and predicted core radii color-coded by anisotropy profile. The points in the right portion of the plot represent the cored profiles, and it can be seen

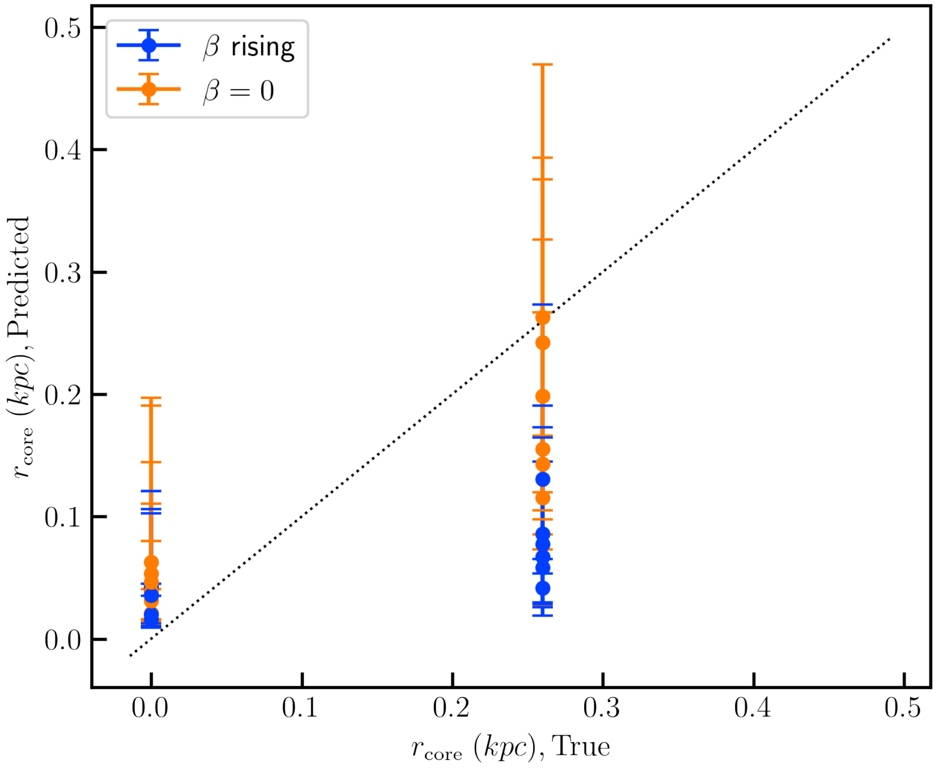
Figure 4.6: Posterior histograms of $\log_{10}[r_{\mathrm{core}}/\mathrm{kpc}]$ for the 24 mock data sets. The 12 cored sets are shown in blue; their true value is -0.585 (corresponding to $r_{\mathrm{core}} = 1$ kpc) and is indicated by the black dotted line. The 12 NFW sets are shown orange; their true value is $-\infty$ (corresponding to $r_{\mathrm{core}} = 0$ kpc), although we limit the parameter to -2 in log space (corresponding to $r_{\mathrm{core}} = 0.01$ kpc.)



that the profiles with positive, rising anisotropy parameter β (colored blue) suffer the most downward bias in the prediction of core size, evidencing a similar mass-anisotropy degeneracy as that seen in other analysis methods.

Figure 4.8 shows the true and predicted posteriors for the anisotropy parameter β at the half-light radius. Half of the mock data sets are anisotropic over their entire range, while the other half have rising β profiles, with a true value between 0.4 and 0.6 at the half-light radius. For the isotropic data sets, the model predictions have median values centered near zero and with a range of -0.2 to 0.2. For the anisotropic data sets, the model tends to systematically underestimates β, especially for the highly embedded data sets denoted in blue.

Figure 4.7: True and predicted values of the core radius $r_{\rm core}$, in kpc, and color-coded by anisotropy profile. The group on the left is have NFW profiles, while the group on the right have cored profiles. The mock data sets with positive, rising β (blue) show a bias towards smaller cores.



4.3.3 Summary of Model Performance with Mock Data

The approximate DF model makes accurate predictions in the $r_{\rm max}$ - $V_{\rm max}$ plane and for half-light radius of the data sets. The mass within the half-light radius is predicted well for those data sets not too deeply embedded in the DM potential, but for those highly embedded there is a tendency to overestimate the mass. The density at 150 pc (ρ_{150}) is accurate to within 0.5 dex in all cases, and within 0.3 dex in most cases. Although the model shows some ability to distinguish between NFW and cored profiles, the accuracy is low, with uncertainty of approximately 0.5 to 1.0 dex, and there is evident mass-anisotropy degeneracy. The predictions of $\beta(r_{1/2})$ are of limited accuracy, especially for the cases that are deeply embedded and that have rising β profiles.

Figure 4.8: True and predicted values for $\beta(r_{1/2})$, color-coded by embeddedness.

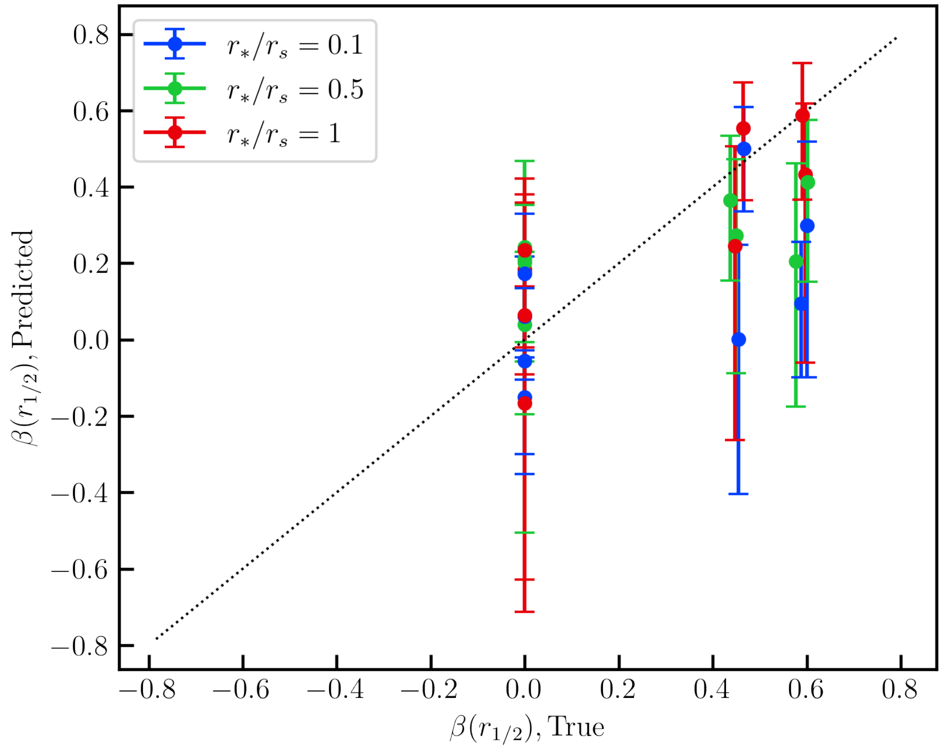
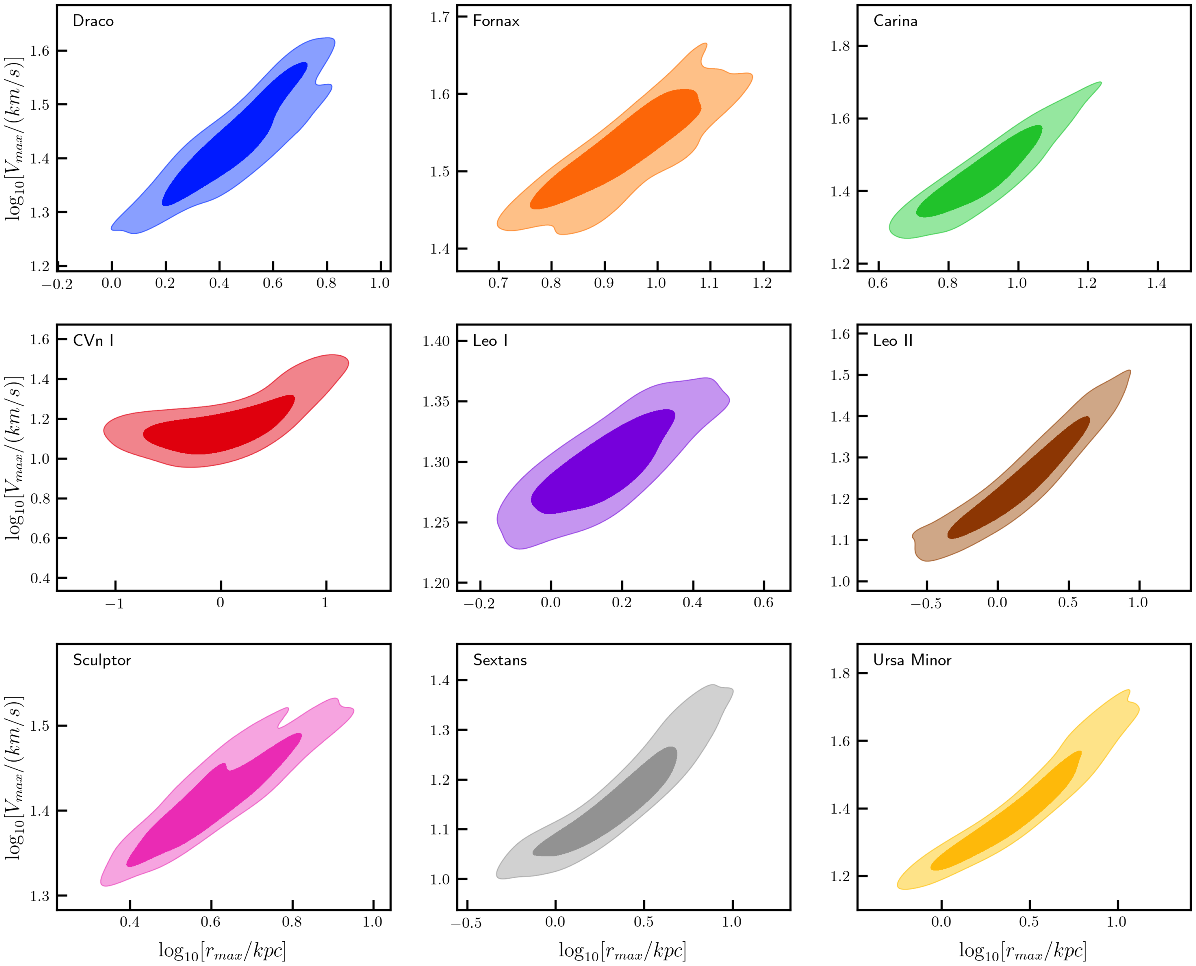


Table 4.2: The Dwarf Galaxy Sample. The 2D half-light radius is quoted in this table. References: (1) Simon [177], (2) McConnachie [109] (3) The NASA/IPAC Extragalactic Database (NED), (4) MW 1 model from Patel et al. [139] (5) Light MW model from Battaglia et al. [6].

Name	Adopted Distance (kpc)	Center RA (deg.)	Center DEC (deg.)	$R_{1/2}$ (pc)	ϵ	Pericenter (kpc)	References
Draco	82.0 ± 6.0	260.051625	57.915361	237 ± 17	0.31	84.94 ± 19.16	1, 3, 3, 1, 2, 4
Fornax	139.0 ± 3.0	39.997200	-34.449187	792 ± 18	0.30	108.89 ± 25.51	1, 3, 3, 1, 2, 4
Carina	106.0 ± 5.0	100.402888	-50.966196	311 ± 15	0.33	80.14 ± 18.25	1, 3, 3, 1, 2, 4
CnV I	211.0 ± 6.0	202.014583	33.555833	437 ± 18	0.39	$68.09^{+71.49}_{-42.17}$	1, 3, 3, 1, 2, 5
Leo I	254.0 ± 15.5	152.117083	12.306389	270 ± 17	0.21	$46.53^{+30.50}_{-26.54}$	1, 3, 3, 1, 2, 5
Leo II	233.0 ± 14.0	168.370000	22.151667	171 ± 10	0.13	$115.55^{+88.35}_{-58.87}$	1, 3, 3, 1, 2, 5
Sculptor	86.0 ± 5.0	15.038984	-33.709029	279 ± 16	0.32	57.21 ± 6.24	1, 3, 3, 1, 2, 4
Sextans	95.0 ± 3.0	153.262319	-1.614602	456 ± 15	0.35	$74.45^{+4.38}_{-5.68}$	1, 3, 3, 1, 2, 5
Ursa Minor	76.0 ± 4.0	227.285379	67.222605	405 ± 21	0.56	77.05 ± 16.17	1, 3, 3, 1, 2, 4

Figure 4.9: Posterior inferences for $r_{\max}$ vs. $V_{\max}$ for the observed sample. The 68% and 95% levels are shown, with the 68% level in a darker color.



4.4 Bright Dwarf Spheroidal Models

We selected as our sample the eight classical dSphs of the MW, plus Canes Venatici I, as shown in Table 4.2. The results from applying the DF model are described here. We use cNFW as the DM profile, as it is the most general of our profiles. We use surface density data from Muñoz et al. [122]. Dispersion data is from Walker et al. [199, 200], Spencer et al. [181], Mateo et al. [108] and M. Walker, private communication, except Canes Venatici I from Simon et al. [178]. VSP data is from Kaplinghat et al. [81]. The results from the analysis are discussed here.

Table 4.3: Comparison of findings for $r_{\max}$ vs. $V_{\max}$ with the Jeans analysis of Kaplinghat et al. [81] for the observed sample. The Jeans analysis considered NFW and cored isothermal profiles as separate cases.

	$\log_{10}[r_{\max}/(\text{kpc})]$			$\log_{10}[V_{\max}/(\text{km/s})]$		
dSph Name	This Work	Jeans NFW	Jeans cISO	This Work	Jeans NFW	Jeans cISO
Draco	$0.45^{+0.18}_{-0.16}$	$0.50^{+0.23}_{-0.24}$	$0.03^{+0.12}_{-0.13}$	$1.43^{+0.10}_{-0.07}$	$1.26^{+0.07}_{-0.03}$	$1.24^{+0.03}_{-0.02}$
Fornax	$0.95^{+0.09}_{-0.11}$	$0.80^{+0.12}_{-0.19}$	$0.45^{+0.17}_{-0.14}$	$1.53^{+0.05}_{-0.06}$	$1.36^{+0.04}_{-0.05}$	$1.33^{+0.06}_{-0.04}$
Carina	$0.91^{+0.11}_{-0.13}$	$0.57^{+0.16}_{-0.26}$	$-0.09^{+0.18}_{-0.32}$	$1.44^{+0.09}_{-0.08}$	$1.28^{+0.06}_{-0.08}$	$1.14^{+0.07}_{-0.05}$
CVn I	$0.05^{+0.50}_{-0.49}$	$0.33^{+0.44}_{-0.60}$	$-0.09^{+0.38}_{-0.41}$	$1.15^{+0.10}_{-0.07}$	$1.24^{+0.15}_{-0.10}$	$1.19^{+0.15}_{-0.06}$
Leo I	$0.16^{+0.12}_{-0.14}$	$0.70^{+0.35}_{-0.38}$	$0.12^{+0.17}_{-0.19}$	$1.30^{+0.03}_{-0.03}$	$1.42^{+0.13}_{-0.12}$	$1.33^{+0.07}_{-0.06}$
Leo II	$0.17^{+0.34}_{-0.36}$	$-0.58^{+0.47}_{-0.49}$	$-1.30^{+0.68}_{-0.46}$	$1.23^{+0.12}_{-0.09}$	$1.16^{+0.06}_{-0.03}$	$1.17^{+0.04}_{-0.03}$
Sculptor	$0.60^{+0.14}_{-0.13}$	$0.62^{+0.29}_{-0.25}$	$0.06^{+0.11}_{-0.08}$	$1.41^{+0.06}_{-0.05}$	$1.43^{+0.10}_{-0.08}$	$1.35^{+0.04}_{-0.03}$
Sextans	$0.36^{+0.25}_{-0.31}$	$0.31^{+0.27}_{-0.34}$	$0.05^{+0.21}_{-0.28}$	$1.15^{+0.09}_{-0.07}$	$1.19^{+0.07}_{-0.06}$	$1.35^{+0.04}_{-0.03}$
Ursa Minor	$0.40^{+0.34}_{-0.28}$	$-0.19^{+0.46}_{-0.46}$	$-0.75^{+0.41}_{-0.77}$	$1.36^{+0.16}_{-0.09}$	$1.26^{+0.07}_{-0.03}$	$1.24^{+0.03}_{-0.02}$

Predictions for $r_{\max}$ and $V_{\max}$ for the observed sample are presented in Figure 4.9. The two parameters show strong positive correlation. Because the halo scale density $\rho_s \propto r_s^2/v_s^2$, this type of degeneracy is approximately along lines of constant density, so that the density is relatively well constrained, as discussed below. In Table 4.3, we compare the $r_{\max}$ and $V_{\max}$ inferences to those of Kaplinghat et al. [81], who used Jeans analysis for their inference and also utilized the vsp. They analyzed two cases, one for an NFW profile and a second for a cored isothermal profile. The results are similar, with inferences from the DF and Jeans methods overlapping at their 1σ boundaries except for $V_{\max}$ of Fornax and both $r_{\max}$ and $V_{\max}$ of Carina. In those, the DF predictions are larger than those from either profile in the Jeans analysis. Possible reasons for the differences could include (1) different prior assumptions between the two methods, and (2) for Fornax, that Kaplinghat et al. [81] accounted for the stellar mass in the potential, in contrast to this work where we have assumed that the stars are massless tracers of the DM potential.

The half-light radius posteriors for the observed sample is shown in Table 4.4. Comparison

Figure 4.10: Posteriors for the half-light radius and mass within that radius for the observed sample. The 68% and 95% levels are shown, with the 68% level in a darker color. Isodensity contours are shown as dotted lines, with the density value indicated, in units of $10^7 M_\odot$ kpc^{-3}.

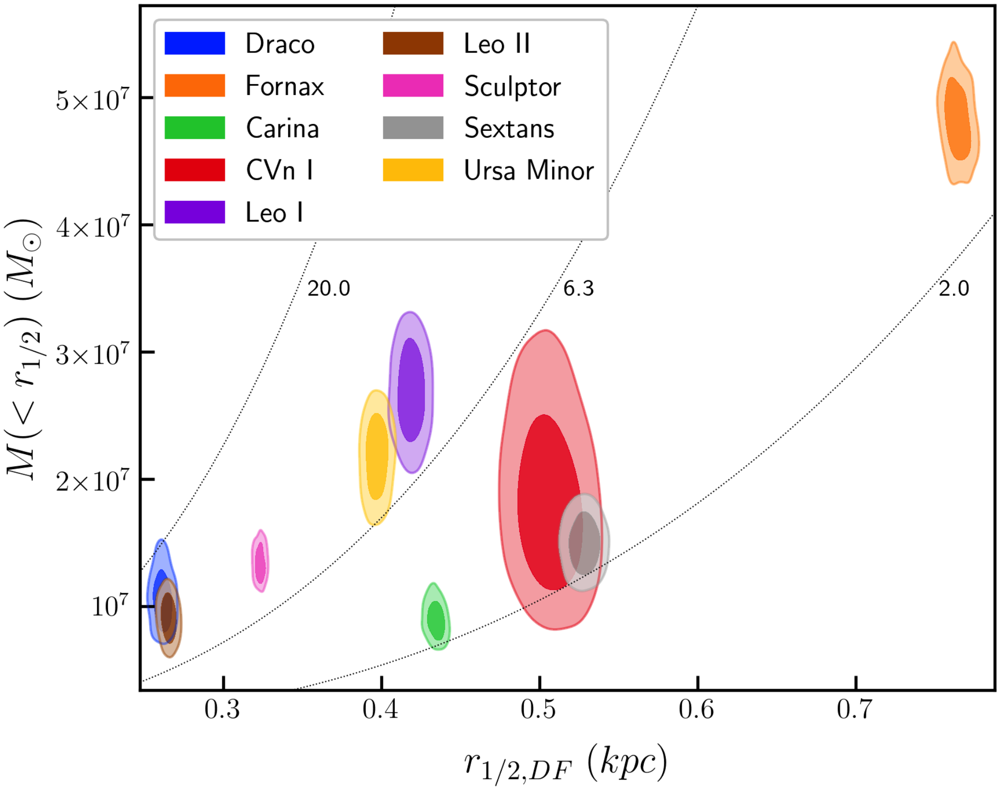


Table 4.4: Comparison of 2D projected half-light radius ($R_{1/2}$), in kpc. The "DF" column is the $r_{1/2}$ posterior result from the DF, converted to 2D projected $R_{1/2}$ by dividing by 1.33. The "Plummer Fit" is the result from the best fitting 2-parameter Plummer profile, as applied to the (one-dimensional) sphericalized surface density data. The rightmost column is the half-light radius reported by Muñoz et al. [122] for a Sersic profile fit to 2D surface density maps, sphericalized as described in the text.

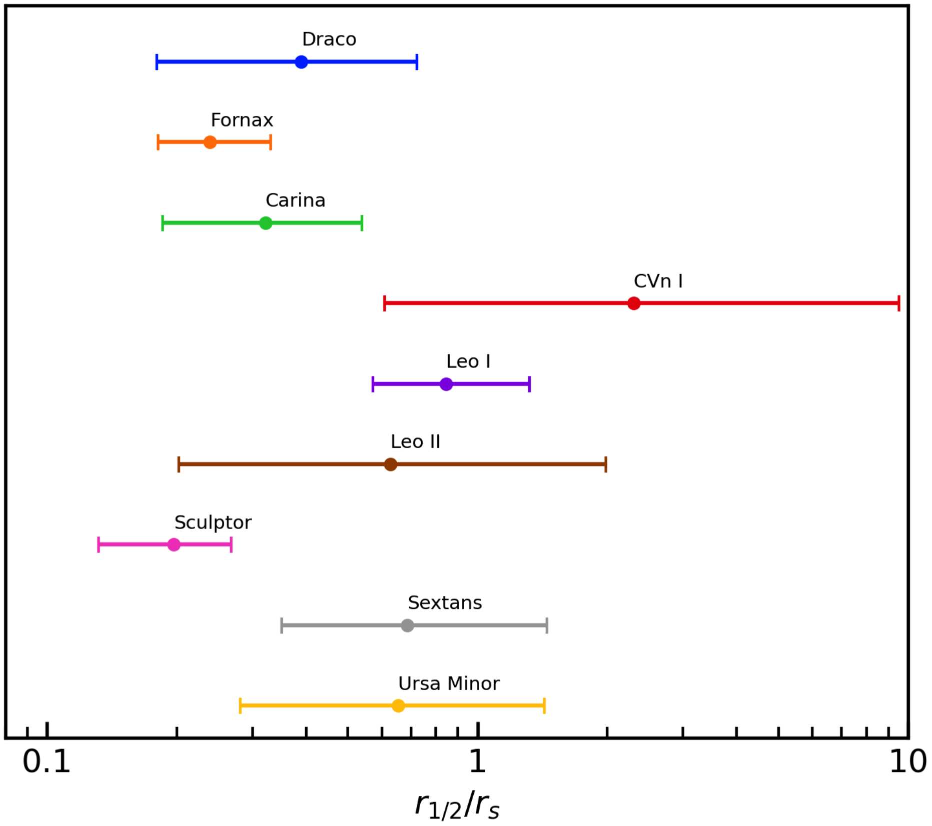
dSph Name	DF	Plummer Fit	Muñoz et al. [122]
Draco	0.197 ± 0.003	0.235	0.183
Fornax	0.574 ± 0.004	0.688	0.668
Carina	0.327 ± 0.003	0.344	0.277
CVn I	0.381 ± 0.010	0.445	0.357
Leo I	0.315 ± 0.004	0.308	0.204
Leo II	0.200 ± 0.002	0.206	0.162
Sculptor	0.243 ± 0.002	0.276	0.244
Sextans	0.397 ± 0.005	0.470	0.370
Ursa Minor	0.299 ± 0.004	0.325	0.257

with previous authors is not straightforward, because of fundamentally different approaches in computation. We compare to Muñoz et al. [122], who fit a Sersic profile to 2-dimensional data maps of the dwarfs. We multiply the Muñoz et al. [122] result by the axis ratio of its elliptical profile, $\sqrt{1-\epsilon}$, to convert from elliptical radius to a spherical one. The data used for our models is a 1-dimensional equivalent of their data, also adjusted for ellipticity. We also show the $R_{1/2}$ resulting from a 2-parameter Plummer profile [145] fit of our input data. The DF approach does not rely on any profile shape; we simply find the radius that encloses half of the stars. As can be seen in the table, there can be substantial differences between the various methods. One notable difference is in Leo I, for which the DF predicts a median value of 0.315 kpc while the Plummer fit to the same data yields 0.308 kpc, and Muñoz et al. [122] find 0.204 kpc. Possible reasons for the difference are (a) the surface density map for Leo I is quite boxy, with ellipticity that appears to change with position angle, and (b) the surface density plateaus considerably at larger radii, making it a poor fit for most parameterized profiles. We note that Read et al. [152] used Jeans analysis combined with virial shape parameters to examine these objects and found 2D half-light radii of 0.298 kpc and 0.194 kpc for Leo I and Leo II, respectively, consistent with our findings.

Figure 4.10 shows 2D posteriors for the half-light radius of the observed sample versus the mass enclosed within that radius. Fornax stands out with the largest half-light radius and largest mass enclosed, however it is in the group with the lowest average density within the half-light radius, accompanied by Carina and Sextans. At the other end of the spectrum are Draco and Leo II, which are the most compact, enclose the least mass but have the highest density within $r_{1/2}$.

Because the degree to which the stellar population is embedded in the DM potential is an important factor for the accuracy of the model, we examine this in Figure 4.11, which shows $r_{1/2}/r_s$ for the observed sample. Sculptor and Fornax have lowest inferences, with median values of 0.20 and 0.24, respectively. Though still less embedded than the most embedded

Figure 4.11: The posterior inference for $r_{1/2}/r_s$, which indicates the degree to which the stellar population is embedded in its DM halo, for the observed sample. The 68% confidence intervals are shown.



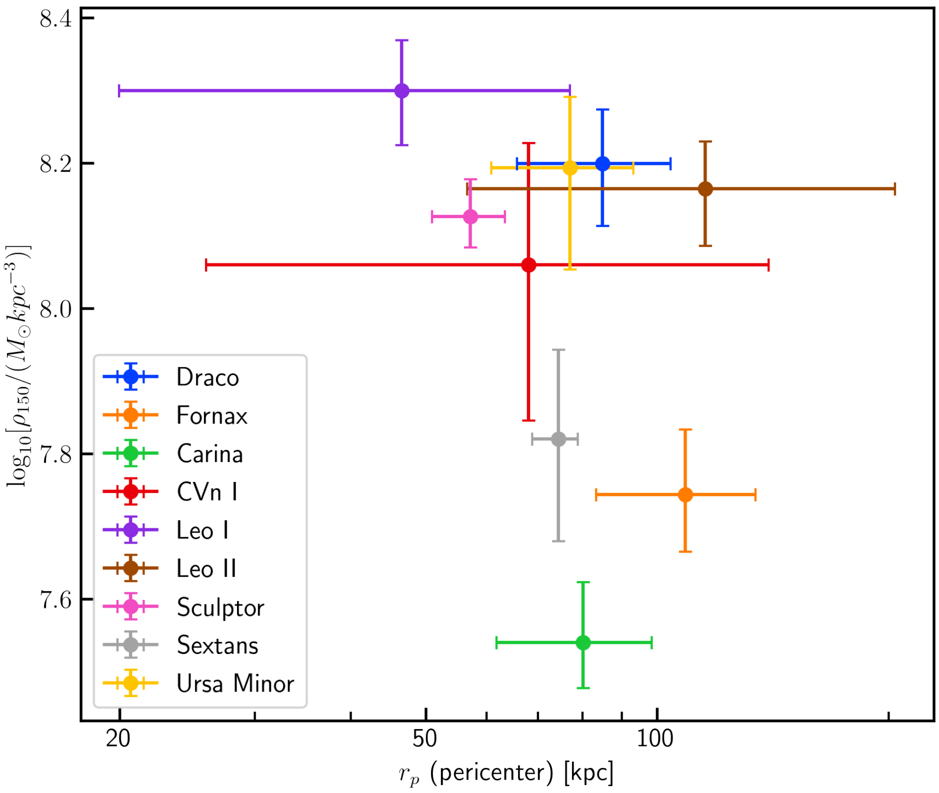
mock data sets, the inferences for Sculptor and Fornax could be vulnerable to the types of biases seen in the most embedded mock data sets.

The posteriors for the density at 150 pc (ρ_{150}) for the observed sample is shown in Figure 4.12, plotted against the orbital pericenter distance (r_p) of each dwarf. Patel et al. [139] recently studied the impact of the Large Magellanic Cloud ("LMC") on the MW dwarfs likely to be signficantly impacted by the LMC, which include Carina, Draco, Fornax, Sculptor and Ursa Minor. For the other pericenter distances we used the recent work of Battaglia et al. [6], which utilizes Gaia data release 3. We use the MW Model 1 from Patel et al. [139] and the Light MW model from Battaglia et al. [6], as their assumptions for the mass of the MW are similar ($1 \times 10^{12} M_\odot$ and $0.8 \times 10^{12} M_\odot$, respectively). There is evidence for anti-

Table 4.5: Comparison of findings for DM density at 150 pc (ρ_{150}), in units of $10^7 M_\odot$ kpc^{-3}. The mean posterior value and the 1σ uncertainties are indicated. The references for comparison are (A) Read et al. [152], (B) Kaplinghat et al. [81], isothermal case, and (C) Kaplinghat et al. [81], NFW case. Note that CVn I was not studied in reference A.

dSph Name	This Work	Ref. A	Ref. B	Ref. C
Draco	16.0 ± 2.9	23.6 ± 2.9	21.7 ± 5.0	21.9 ± 2.3
Fornax	5.8 ± 1.2	7.9 ± 2.3	3.6 ± 1.5	7.8 ± 1.7
Carina	3.6 ± 0.6	11.6 ± 2.1	6.4 ± 2.4	10.4 ± 1.0
CVn I	11.9 ± 4.8	—	12.9 ± 5.9	13.8 ± 2.8
Leo I	20.0 ± 3.2	17.7 ± 3.3	14.6 ± 5.0	15.5 ± 2.8
Leo II	14.6 ± 2.3	18.4 ± 1.7	14.8 ± 3.0	16.3 ± 3.0
Sculptor	13.5 ± 1.4	14.9 ± 2.6	15.9 ± 3.1	17.1 ± 2.1
Sextans	6.8 ± 2.0	12.8 ± 3.2	9.3 ± 4.3	11.5 ± 2.3
Ursa Minor	15.6 ± 3.7	15.3 ± 3.4	25.7 ± 5.9	24.4 ± 3.7

Figure 4.12: DM density at 150 pc (ρ_{150}) inferred from the DF fits for the bright MW dSphs vs orbital pericenter distance (r_p). The error bars indicate the 68% confidence interval.



correlation between the pericenter distance and the density at 150 pc, as might be caused by tidal stripping [81], however the correlation appears stronger in that work than it does here. We have used different pericenter data, and in particular the pericenter for Leo II is substantially larger in our work (although the uncertainty is large), which weakens the correlation between density and pericenter.

Table 4.5 compares our findings for ρ_{150} to those of Read et al. [152] and Kaplinghat et al. [81]. The results are generally comparable within errors. However, our finding for Carina at $3.6 \pm 0.6\ M_\odot\ \mathrm{kpc}^{-3}$ is substantially lower than the others, inconsistent with that of Read et al. [152] but marginally compatible with the isothermal case of Kaplinghat et al. [81]. We checked to see if excluding large values of c in the cNFW profile would change this inference significantly, but it does not; we found that if the core parameter is restricted so that $0 < c < 1$, the inference for ρ_{150} increases only approximately 0.1 dex.

For completeness, we briefly discuss the inferences for the anisotropy parameter at the half-light radius, the core parameter and the core radius, although the results of the mock data show that these inferences are of limited reliability. The posteriors for $c = r_c/r_s$ and $\log_{10}[r_{\mathrm{core}}/\mathrm{kpc}]$ are shown in Appendix C.4 in Figures C.1 and C.2. Most of the sample shows a preference for nearly zero values of the core parameter, indicating a density profile that is close to the NFW profile. As the results from mock data show limited accuracy and evident bias towards lower core sizes, especially for those data sets with anisotropic velocity distributions, we make no conclusions about core sizes in the observed sample. We note that Peñarrubia et al. [143] found cores in Fornax and Sculptor by exploiting separate chemo-dynamic subcomponents, although Strigari et al. [184] found only a weak preference for a core in Sculptor and that both cored and NFW profiles were good fits. The posteriors for $\beta(r_{1/2})$ are shown in Appendix C.4, Figure C.3. The inferences of the mock data had low accuracy and tended to understate the true value of β where the value was positive. Perhaps the most that can be said here is that Draco and Carina appear to have radial anisotropy at the half-light radius, given that the posteriors are mostly positive and a bias towards understating the true value would make the true values even more likely to be positive.

4.5 Conclusions

In this work, we presented a novel study of the internal dynamics of the brightest dSphs of the MW. Going beyond the standard spherical Jeans analysis often adopted for these systems, our method relies on a separable distribution function that describes the phase space of stellar tracers via 10 distinct parameters, shaping the energy and angular momentum functional form. The DF approach we follow here is completed by the modeling of the gravitational potential of the system, for which we adopted a 3-parameter cNFW distribution. This distribution is suitable for an investigation of both cuspy and cored DM halos. For the first time in literature, we apply such a general approach to the whole set of 9 dSphs with well-measured kinematics, performing a challenging data-driven Bayesian analysis on the photometric and spectroscopic data available for these objects.

Our analysis via DF modeling is validated by the use of mock data extracted from the Gaia Challenge project. In particular, we adopted mock data sets to test the predictive capability of our approach both for cuspy and cored DM profiles, for cuspy and cored stellar profiles, for different level of embeddedness of the stellar distribution within the DM halo of the system and for spatially-varying stellar orbital anisotropy profiles. From the study of the mock data we find that our DF approach is able to recover the true values of the $V_{\max}$ and $r_{\max}$ shape parameters of the underlying DM profile remarkably well, usually within the 68% posterior probability region (see Figure 4.2). It also has high accuracy for the recovery of key dynamical quantities such as the total mass within the half-light radius, $M_{1/2}$ (see Table 4.4 and Figure 4.10) and the inner local density of the system at 150 pc, ρ_{150} (Figure 4.5). In contrast, the mock data show us that with this approach it remains difficult to reliably determine the size of the core of the DM inner halo or to infer robust information about the orbital anisotropy profile of stellar tracers, difficulties also suffered by Jeans analysis.

Equipped with these findings, our detailed study of the MW dSphs allowed us to revisit, reiterate and reinforce some well-known conclusions already drawn in literature within the standard Jeans analysis. Our study of the Classical dSphs via DF modeling provides a state-of-the-art inference of ρ_{150} in these objects. In particular, we find a low inner density for systems like Carina, Fornax and Sextans, in contrast to galaxies like Draco, Leo I, Leo II and Ursa Minor, characterized by inner densities approximately four times larger. With the DF approach, we are then able to confirm the puzzling diversity of DM-dominated objects like MW dSphs, at the basis of the too big to fail problem. These inferences of the inner density constitute key dynamical information that needs to be captured by any successful model of galaxy formation within the ΛCDM cosmological model, or addressed by any attempt to go beyond the paradigm of cold and collisionless DM particles. In this regard, we note that the correlation between dwarf spheroidal pericenters and density at 150 pc found in Kaplinghat et al. [81] has been weakened, mainly due to the recent assessment of the pericenter determination for these systems (although that remains subject to many uncertainties, such as a correct description of the effects of the Large Magellanic Cloud on the orbits of some of the satellites considered here). Even so, the pericenter-density correlation remains a compelling and interesting clue for paradigms beyond ΛCDM to explain.

We also observe that for Fornax and Carina, the results of our analysis with the cNFW profile point to a tantalizing presence of a core in these systems. Nevertheless, extra care should be taken in considering this inference vis-a-vis the observed systematics of this approach, given our limited ability in distinguishing cored and cuspy halo profiles in the mock data sets.

The results of our study are promising in a twofold way. On one hand, our analysis showed that the DF modeling has a similar constraining power to that of the spherical Jean analysis, despite relying on the larger set of parameters varied in the MCMC analysis that are needed for a broad description the tracer phase-space distribution function. This is in contrast to previous findings in literature, that pointed to a much large degeneracy in parameter space, an observation that we do not confirm here. Second, we showed that the use of a fourth-order virial shape parameter helps in pinning down the DM content in pressure-supported systems like dSphs. A natural extension of this work will involve DF models that allow for multiple populations with individual metallicity distributions. Such models may allow a richer inference of the slope of the DM inner profile. This will eventually be key in order to reconcile the cusp/core problem in MW dSphs. The prospective ability to reconcile small-scale problems such as cusp/core and too big to fail problem make these objects among the most interesting places to look for indications of physics beyond ΛCDM.

Acknowledgements

We thank Matthew Walker and Andrew Pace for providing velocity dispersion data. We gratefully acknowledge a grant of computer time from TACC allocation AST20027.

Bibliography

[1] N. Aghanim, Y. Akrami, M. Ashdown, J. Aumont, C. Baccigalupi, M. Ballardini, A. J. Banday, R. B. Barreiro, N. Bartolo, S. Basak, R. Battye, K. Benabed, J. P. Bernard, M. Bersanelli, P. Bielewicz, J. J. Bock, J. R. Bond, J. Borrill, F. R. Bouchet, F. Boulanger, M. Bucher, C. Burigana, R. C. Butler, E. Calabrese, J. F. Cardoso, J. Carron, A. Challinor, H. C. Chiang, J. Chluba, L. P.L. Colombo, C. Combet, D. Contreras, B. P. Crill, F. Cuttaia, P. De Bernardis, G. De Zotti, J. Delabrouille, J. M. Delouis, E. Di Valentino, J. M. Diego, O. Doré, M. Douspis, A. Ducout, X. Dupac, S. Dusini, G. Efstathiou, F. Elsner, T. A. Enßlin, H. K. Eriksen, Y. Fantaye, M. Farhang, J. Fergusson, R. Fernandez-Cobos, F. Finelli, F. Forastieri, M. Frailis, A. A. Fraisse, E. Franceschi, A. Frolov, S. Galeotta, S. Galli, K. Ganga, R. T. Génova-Santos, M. Gerbino, T. Ghosh, J. González-Nuevo, K. M. Górski, S. Gratton, A. Gruppuso, J. E. Gudmundsson, J. Hamann, W. Handley, F. K. Hansen, D. Herranz, S. R. Hildebrandt, E. Hivon, Z. Huang, A. H. Jaffe, W. C. Jones, A. Karakci, E. Keihänen, R. Keskitalo, K. Kiiveri, J. Kim, T. S. Kisner, L. Knox, N. Krachmalnicoff, M. Kunz, H. Kurki-Suonio, G. Lagache, J. M. Lamarre, A. Lasenby, M. Lattanzi, C. R. Lawrence, M. Le Jeune, P. Lemos, J. Lesgourgues, F. Levrier, A. Lewis, M. Liguori, P. B. Lilje, M. Lilley, V. Lindholm, M. López-Caniego, P. M. Lubin, Y. Z. Ma, J. F. Maciás-Pérez, G. Maggio, D. Maino, N. Mandolesi, A. Mangilli, A. Marcos-Caballero, M. Maris, P. G. Martin, M. Martinelli, E. Martínez-González, S. Matarrese, N. Mauri, J. D. McEwen, P. R. Meinhold, A. Melchiorri, A. Mennella, M. Migliaccio, M. Millea, S. Mitra, M. A. Miville-Deschênes, D. Molinari, L. Montier, G. Morgante, A. Moss, P. Natoli, H. U. Nørgaard-Nielsen, L. Pagano, D. Paoletti, B. Partridge, G. Patanchon, H. V. Peiris, F. Perrotta, V. Pettorino, F. Piacentini, L. Polastri, G. Polenta, J. L. Puget, J. P. Rachen, M. Reinecke, M. Remazeilles, A. Renzi, G. Rocha, C. Rosset, G. Roudier, J. A. Rubiño-Martín, B. Ruiz-Granados, L. Salvati, M. Sandri, M. Savelainen, D. Scott, E. P.S. Shellard, C. Sirignano, G. Sirri, L. D. Spencer, R. Sunyaev, A. S. Suur-Uski, J. A. Tauber, D. Tavagnacco, M. Tenti, L. Toffolatti, M. Tomasi, T. Trombetti, L. Valenziano, J. Valiviita, B. Van Tent, L. Vibert, P. Vielva, F. Villa, N. Vittorio, B. D. Wandelt, I. K. Wehus, M. White, S. D.M. White, A. Zacchei, and A. Zonca. Planck 2018 results: VI. Cosmological parameters. *Astronomy and Astrophysics*, 641:A6, sep 2020. ISSN 14320746. doi: 10.1051/0004-6361/201833910. URL `https://www.aanda.org/articles/aa/full{_}html/2020/09/aa33910-18/aa33910-18.htmlhttps://www.aanda.org/articles/aa/abs/2020/09/aa33910-18/aa33910-18.html`.

[2] Gerardo Alvarez and Hai-Bo Yu. Astrophysical probes of inelastic dark matter with a light mediator. *Physical Review D*, 101(4):043002, February 2020. doi: 10.1103/PhysRevD.101.043002.

[3] Kevin E Andrade, Quinn Minor, Anna Nierenberg, and Manoj Kaplinghat. Detecting dark matter cores in galaxy clusters with strong lensing. *Monthly Notices of the Royal Astronomical Society*, 487(2):1905–1926, 2019. ISSN 0035-8711. doi: 10.1093/mnras/stz1360. URL `https://arxiv.org/pdf/1901.00507.pdf`.

[4] M. Annunziatella, M. Bonamigo, C. Grillo, A. Mercurio, P. Rosati, G. Caminha, A. Biviano, M. Girardi, R. Gobat, M. Lombardi, and E. Munari. Mass Profile Decomposition of the Frontier Fields Cluster MACS J0416-2403: Insights on the Dark-matter Inner Profile. *The Astrophysical Journal*, 851(2):81, dec 2017. ISSN 1538-4357. doi: 10.3847/1538-4357/aa9845. URL `http://stacks.iop.org/0004-637X/851/i=2/a=81?key=crossref.465c69331633ad0e5d8a238f05b5a4dd`.

[5] Shmuel Balberg, Stuart L. Shapiro, and Shogo Inagaki. Selfinteracting dark matter halos and the gravothermal catastrophe. *Astrophys. J.*, 568:475–487, 2002. doi: 10.1086/339038.

[6] G. Battaglia, S. Taibi, G. F. Thomas, and T. K. Fritz. Gaia early DR3 systemic motions of Local Group dwarf galaxies and orbital properties with a massive Large Magellanic Cloud, 2021. ISSN 0004-6361. URL `http://arxiv.org/abs/2106.08819`.

[7] Peter S Behroozi, Risa H Wechsler, and Charlie Conroy. The average star formation histories of galaxies in dark matter halos from z = 0-8. *Astrophysical Journal*, 770(1):57, 2013. ISSN 15384357. doi: 10.1088/0004-637X/770/1/57.

[8] Sirio Belli, Tucker Jones, Richard S. Ellis, and Johan Richard. TESTING THE UNIVERSALITY OF THE FUNDAMENTAL METALLICITY RELATION AT HIGH REDSHIFT USING LOW-MASS GRAVITATIONALLY LENSED GALAXIES. *The Astrophysical Journal*, 772(2):141, jul 2013. ISSN 0004-637X. doi: 10.1088/0004-637X/772/2/141. URL `http://stacks.iop.org/0004-637X/772/i=2/a=141?key=crossref.fdb3191b6381508cee9d298d85955977`.

[9] E. Bertin and S. Arnouts. SExtractor: Software for source extraction. *Astronomy and Astrophysics Supplement Series*, 117(2):393–404, 1996. ISSN 0365-0138. doi: 10.1051/aas:1996164.

[10] James Binney and Scott Tremaine. *Galactic dynamics*. Princeton University Press, 2008. ISBN 9780691130262. URL `http://adsabs.harvard.edu/abs/2008gady.book.....B`.

[11] Kimberly K. Boddy, Manoj Kaplinghat, Anna Kwa, and Annika H.G. Peter. Hidden sector hydrogen as dark matter: Small-scale structure formation predictions and the importance of hyperfine interactions. *Physical Review D*, 94(12):123017, dec 2016. ISSN 24700029. doi: 10.1103/PhysRevD.94.123017.

[12] M Bonamigo, C Grillo, S Ettori, G B Caminha, P Rosati, A Mercurio, E Munari, M Annunziatella, I Balestra, and M Lombardi. Dissection of the Collisional and Collisionless Mass Components in a Mini Sample of CLASH and HFF Massive Galaxy Clusters at z=0.4. *The Astrophysical Journal*, 864(1):98, sep 2018. ISSN 1538-4357. doi: 10.3847/1538-4357/aad4a7. URL `https://doi.org/10.3847/1538-4357/aad4a7`.

[13] Mario Bonamigo, Giulia Despali, Marceau Limousin, Raul Angulo, Carlo Giocoli, and Geneviève Soucail. Universality of dark matter haloes shape over six decades in mass: Insights from the Millennium XXL and SBARBINE simulations. *Monthly Notices of the Royal Astronomical Society*, 449(3):3171–3182, 2015. ISSN 13652966. doi: 10.1093/mnras/stv417. URL `https://academic.oup.com/mnras/article-abstract/449/3/3171/1117159`.

[14] Mario Bonamigo, Giulia Despali, Marceau Limousin, Raul Angulo, Carlo Giocoli, and Geneviève Soucail. Universality of dark matter haloes shape over six decades in mass: Insights from the Millennium XXL and SBARBINE simulations. *Monthly Notices of the Royal Astronomical Society*, 449(3):3171–3182, 2015. ISSN 13652966. doi: 10.1093/mnras/stv417. URL `http://www.mpa-garching.mpg.de/gadget`.

[15] Michael Boylan-Kolchin, James S Bullock, and Manoj Kaplinghat. Too big to fail? The puzzling darkness of massive Milky Way subhaloes. *Mon. Not. R. Astron. Soc*, 415:40–44, 2011. doi: 10.1111/j.1745-3933.2011.01074.x. URL `https://academic.oup.com/mnrasl/article-abstract/415/1/L40/965196`.

[16] Eric Braaten, Daekyoung Kang, and Ranjan Laha. Production of dark-matter bound states in the early universe by three-body recombination. *Journal of High Energy Physics*, 2018(11):84, November 2018. doi: 10.1007/JHEP11(2018)084.

[17] Maruša Bradač, Steven W Allen, Tommaso Treu, Harald Ebeling, Richard Massey, R Glenn Morris, Anja von der Linden, and Douglas Applegate. Revealing the Properties of Dark Matter in the Merging Cluster MACS J0025.4-1222. *The Astrophysical Journal*, 687(2):959–967, 2008. ISSN 0004-637X. doi: 10.1086/591246.

[18] Maarten A Breddels, Amina Helmi, R. C.E. Van Den Bosch, G Van De Ven, and G Battaglia. Schwarzschild models of the Sculptor dSph galaxy. In *EPJ Web of Conferences*, volume 19, 2012. ISBN 9782759807185. doi: 10.1051/epjconf/20121903009.

[19] T Broadhurst, M Takada, K Umetsu, X Kong, N Arimoto, M Chiba, and T Futamase. The Surprisingly Steep Mass Profile of A1689, from a Lensing Analysis of Subaru Images. *Astrophysical Journal Letters*, 619(2):L143–L146, feb 2005. doi: 10.1086/428122.

[20] James S Bullock and Michael Boylan-Kolchin. Small-Scale Challenges to the Λ CDM Paradigm. *Annual Review of Astronomy and Astrophysics*, 55(1):343–387, 2017. ISSN 0066-4146. doi: 10.1146/annurev-astro-091916-055313. URL `www.annualreviews.org`.

[21] James S Bullock and Michael Boylan-Kolchin. Small-Scale Challenges to the Λ CDM Paradigm. *Annual Review of Astronomy and Astrophysics*, 55(1):343–387, 2017. ISSN 0066-4146. doi: 10.1146/annurev-astro-091916-055313. URL `www.annualreviews.org`.

[22] G. B. Caminha, C. Grillo, P. Rosati, I. Balestra, W. Karman, M. Lombardi, A. Mercurio, M. Nonino, P. Tozzi, A. Zitrin, A. Biviano, M. Girardi, A. M. Koekemoer, P. Melchior, M. Meneghetti, E. Munari, S. H. Suyu, K. Umetsu, M. Annunziatella, S. Borgani, T. Broadhurst, K. I. Caputi, D. Coe, C. Delgado-Correal, S. Ettori, A. Fritz, B. Frye, R. Gobat, C. Maier, A. Monna, M. Postman, B. Sartoris, S. Seitz, E. Vanzella, and B. Ziegler. CLASH-VLT: A highly precise strong lensing model of the galaxy cluster RXC J2248.7 - 4431 (Abell S1063) and prospects for cosmography. *Astronomy and Astrophysics*, 587, mar 2016. ISSN 14320746. doi: 10.1051/0004-6361/201527670. URL `http://www.stsci.edu/hst/campaigns/frontier-fields/`.

[23] G. B. Caminha, C. Grillo, P. Rosati, M. Meneghetti, A. Mercurio, S. Ettori, I. Balestra, A. Biviano, K. Umetsu, E. Vanzella, M. Annunziatella, M. Bonamigo, C. Delgado-Correal, M. Girardi, M. Lombardi, M. Nonino, B. Sartoris, P. Tozzi, M. Bartelmann, L. Bradley, K. I. Caputi, D. Coe, H. Ford, A. Fritz, R. Gobat, M. Postman, S. Seitz, and A. Zitrin. Mass distribution in the core of MACS J1206. *Astronomy & Astrophysics*, 607:A93, nov 2017. ISSN 0004-6361. doi: 10.1051/0004-6361/201731498. URL `http://www.aanda.org/10.1051/0004-6361/201731498`.

[24] M. Cappellari, R. M. McDermid, K. Alatalo, L. Blitz, M. Bois, F. Bournaud, M. Bureau, A. F. Crocker, R. L. Davies, T. A. Davis, P. T. de Zeeuw, P.-A. Duc, E. Emsellem, S. Khochfar, D. Krajnović, H. Kuntschner, R. Morganti, T. Naab, T. Oosterloo, M. Sarzi, N. Scott, P. Serra, A.-M. Weijmans, and L. M. Young. The ATLAS^{3D} project - XX. Mass-size and mass-σ distributions of early-type galaxies: bulge fraction drives kinematics, mass-to-light ratio, molecular gas fraction and stellar initial mass function. *Monthly Notices of the Royal Astronomical Society*, 432:1862–1893, July 2013. doi: 10.1093/mnras/stt644.

[25] Michele Cappellari. Measuring the inclination and mass-to-light ratio of axisymmetric galaxies via anisotropic jeans models of stellar kinematics. *Monthly Notices of the Royal Astronomical Society*, 390(1):71–86, 2008. doi: 10.1111/j.1365-2966.2008.13754.x. URL `http://dx.doi.org/10.1111/j.1365-2966.2008.13754.x`.

[26] Catherine Cerny, Keren Sharon, Felipe Andrade-Santos, Roberto J. Avila, Maruša Bradač, Larry D. Bradley, Daniela Carrasco, Dan Coe, Nicole G. Czakon, William A. Dawson, Brenda L. Frye, Austin Hoag, Kuang-Han Huang, Traci L. Johnson, Christine Jones, Daniel Lam, Lorenzo Lovisari, Ramesh Mainali, Pascal A. Oesch, Sara Ogaz, Matthew Past, Rachel Paterno-Mahler, Avery Peterson, Adam G. Riess, Steven A. Rodney, Russell E. Ryan, Brett Salmon, Irene Sendra-Server, Daniel P. Stark, Louis-Gregory Strolger, Michele Trenti, Keiichi Umetsu, Benedetta Vulcani, and Adi Zitrin. RELICS: Strong Lens Models for Five Galaxy Clusters from the Reionization Lensing Cluster Survey. *The Astrophysical Journal*, 859(2):159, 2018. doi: 10.3847/1538-4357/aabe7b. URL `https://doi.org/10.3847/1538-4357/aabe7b`.

[27] G. Chabrier, P. Hennebelle, and S. Charlot. Variations of the Stellar Initial Mass Function in the Progenitors of Massive Early-type Galaxies and in Extreme Starburst Environments. *Astrophysical Journal*, 796:75, December 2014. doi: 10.1088/0004-637X/796/2/75.

[28] Gilles Chabrier. Galactic Stellar and Substellar Initial Mass Function. *Publications of the Astronomical Society of the Pacific*, 115(809):763–795, jul 2003. ISSN 0004-6280. doi: 10.1086/376392. URL `http://iopscience.iop.org/article/10.1086/376392`.

[29] Soo-Min Choi, Hyun Min Lee, and Min-Seok Seo. Cosmic abundances of SIMP dark matter. *Journal of High Energy Physics*, 2017(4):154, April 2017. doi: 10.1007/JHEP04(2017)154.

[30] J. Andrés Christen and Colin Fox. A general purpose sampling algorithm for continuous distributions (the t-walk). *Bayesian Analysis*, 5(2):263–281, jun 2010. ISSN 1936-0975. doi: 10.1214/10-BA603. URL `http://projecteuclid.org/euclid.ba/1340218339`.

[31] Xiaoyong Chu, Camilo Garcia-Cely, and Hitoshi Murayama. Velocity Dependence from Resonant Self-Interacting Dark Matter. *Physical Review Letters*, 122(7):071103, February 2019. doi: 10.1103/PhysRevLett.122.071103.

[32] Xiaoyong Chu, Camilo Garcia-Cely, and Hitoshi Murayama. A practical and consistent parametrization of dark matter self-interactions. *Journal of Cosmology and Astroparticle Physics*, 2020(6):043, June 2020. doi: 10.1088/1475-7516/2020/06/043.

[33] D Clowe, G. De Lucia, and L King. Effects of asphericity and substructure on the determination of cluster mass with weak gravitational lensing. *Monthly Notices of the Royal Astronomical Society*, 350(3):1038–1048, 2004. ISSN 00358711. doi: 10.1111/j.1365-2966.2004.07723.x. URL `https://academic.oup.com/mnras/article/350/3/1038/971161`.

[34] Dan Coe, Keiichi Umetsu, Adi Zitrin, Megan Donahue, Elinor Medezinski, Marc Postman, Mauricio Carrasco, Timo Anguita, Margaret J. Geller, Kenneth J. Rines, Antonaldo Diaferio, Michael J. Kurtz, Larry Bradley, Anton Koekemoer, Wei Zheng, Mario Nonino, Alberto Molino, Andisheh Mahdavi, Doron Lemze, Leopoldo Infante, Sara Ogaz, Peter Melchior, Ole Host, Holland Ford, Claudio Grillo, Piero Rosati, Yolanda Jiménez-Teja, John Moustakas, Tom Broadhurst, Begoña Ascaso, Ofer Lahav, Matthias Bartelmann, Narciso Benítez, Rychard Bouwens, Or Graur, Genevieve Graves, Saurabh Jha, Stephanie Jouvel, Daniel Kelson, Leonidas Moustakas, Dan Maoz, Massimo Meneghetti, Julian Merten, Adam Riess, Steve Rodney, and Stella Seitz. CLASH: PRECISE NEW CONSTRAINTS ON THE MASS PROFILE OF THE GALAXY CLUSTER A2261. *The Astrophysical Journal*, 757(1):22, sep 2012. ISSN 0004-637X. doi: 10.1088/0004-637X/757/1/22. URL `http://stacks.iop.org/0004-637X/757/i=1/a=22?key=crossref.2082029d0a1a51ca6641e0f508225e6d`.

[35] C. Conroy and P. G. van Dokkum. The Stellar Initial Mass Function in Early-type Galaxies From Absorption Line Spectroscopy. II. Results. *Astrophysical Journal*, 760:71, November 2012. doi: 10.1088/0004-637X/760/1/71.

[36] C. Conroy, P. G. van Dokkum, and A. Villaume. The Stellar Initial Mass Function in Early-type Galaxies from Absorption Line Spectroscopy. IV. A Super-Salpeter IMF in the Center of NGC 1407 from Non-parametric Models. *Astrophysical Journal*, 837: 166, March 2017. doi: 10.3847/1538-4357/aa6190.

[37] M. J. Currie, D. S. Berry, T. Jenness, A. G. Gibb, G. S. Bell, and P. W. Draper. *Starlink Software in 2013*, volume 485 of *Astronomical Society of the Pacific Conference Series*, page 391. 2014. Gaia is a part of this package.

[38] Philip J Davis and Philip Rabinowitz. *Methods of Numerical Integration: Second Edition (Dover Books on Mathematics)*. Dover Publications, 2007. ISBN 0486453391.

[39] William A Dawson, David Wittman, M James Jee, Perry Gee, John P Hughes, J Anthony Tyson, Samuel Schmidt, Paul Thorman, Maruša Bradač, Satoshi Miyazaki, Brian Lemaux, Yousuke Utsumi, and Vera E Margoniner. Discovery of a dissociative galaxy cluster merger with large physical separation. *Astrophysical Journal Letters*, 747(2):42, 2012. ISSN 20418205. doi: 10.1088/2041-8205/747/2/L42.

[40] A. Del Popolo. On the density-profile slope of clusters of galaxies. *Monthly Notices of the Royal Astronomical Society*, 424(1):38–51, jul 2012. ISSN 00358711. doi: 10.1111/j.1365-2966.2012.21141.x. URL `https://academic.oup.com/mnras/article-lookup/doi/10.1111/j.1365-2966.2012.21141.x`.

[41] A. Del Popolo. The flat density profiles of massive, and relaxed galaxy clusters. *Journal of Cosmology and Astroparticle Physics*, 2014(7):019–019, jul 2014. ISSN 14757516. doi: 10.1088/1475-7516/2014/07/019. URL `http://stacks.iop.org/1475-7516/2014/i=07/a=019?key=crossref.7d1849718934d2e48e177d82a9f11e4e`.

[42] Vandana Desai. *The assembly of cluster galaxies.* PhD thesis, UNIVERSITY OF WASHINGTON, jan 2004.

[43] Foivos I. Diakogiannis, Geraint F. Lewis, and Rodrigo A. Ibata. Resolving the mass-anisotropy degeneracy of the spherically symmetric jeans equation - I. Theoretical foundation. *Monthly Notices of the Royal Astronomical Society*, 443(1):598–609, sep 2014. ISSN 13652966. doi: 10.1093/mnras/stu1153. URL `https://academic.oup.com/mnras/article/443/1/598/1483112`.

[44] Benedikt Diemer. COLOSSUS: A Python Toolkit for Cosmology, Large-scale Structure, and Dark Matter Halos. *Astrophysical Journals*, 239(2):35, December 2018. ISSN 0067-0049. doi: 10.3847/1538-4365/aaee8c. URL `https://doi.org/10.3847/1538-4365/aaee8c`.

[45] Benedikt Diemer and Michael Joyce. An accurate physical model for halo concentrations. *The Astrophysical Journal*, 871(2):168, jan 2019. doi: 10.3847/1538-4357/aafad6. URL `https://doi.org/10.3847%2F1538-4357%2Faafad6`.

[46] A. Donnarumma, S. Ettori, M. Meneghetti, and L. Moscardini. X-ray and strong lensing mass estimate of MS2137.3–2353. *Monthly Notices of the Royal Astronomical*

Society, 398(1):438–450, 08 2009. ISSN 0035-8711. doi: 10.1111/j.1365-2966.2009.15165.x. URL `https://doi.org/10.1111/j.1365-2966.2009.15165.x`.

[47] A. Donnarumma, S. Ettori, M. Meneghetti, R. Gavazzi, B. Fort, L. Moscardini, A. Romano, L. Fu, F. Giordano, M. Radovich, R. Maoli, R. Scaramella, and J. Richard. Abell 611. *Astronomy & Astrophysics*, 528:A73, apr 2011. ISSN 0004-6361. doi: 10.1051/0004-6361/201014120. URL `http://www.aanda.org/10.1051/0004-6361/201014120`.

[48] Yohan Dubois, Julien Devriendt, Adrianne Slyz, and Romain Teyssier. Jet-regulated cooling catastrophe. *Monthly Notices of the Royal Astronomical Society*, 409(3):985–1001, dec 2010. ISSN 00358711. doi: 10.1111/j.1365-2966.2010.17338.x.

[49] H. S. Dúmet-Montoya, G. B. Caminha, and M. Makler. Domain of validity for pseudo-elliptical NFW lens models. *Astronomy & Astrophysics*, 544:A83, aug 2012. ISSN 0004-6361. doi: 10.1051/0004-6361/201118485. URL `http://www.aanda.org/10.1051/0004-6361/201118485`.

[50] H. Ebeling, E. Barrett, D. Donovan, C. J. Ma, A. C. Edge, and L. van Speybroeck. A Complete Sample of 12 Very X-Ray Luminous Galaxy Clusters at z > 0.5. *Astrophysical Journal Letters*, 661(1):L33–L36, May 2007. doi: 10.1086/518603.

[51] Oliver D. Elbert, James S. Bullock, Shea Garrison-Kimmel, Miguel Rocha, Jose Oñorbe, and Annika H.G. Peter. Core formation in dwarf haloes with self-interacting dark matter: No fine-tuning necessary. *Monthly Notices of the Royal Astronomical Society*, 453(1):29–37, 2015. ISSN 13652966. doi: 10.1093/mnras/stv1470.

[52] Ardís Elíasdóttir, Marceau Limousin, Johan Richard, Jens Hjorth, Jean-Paul Kneib, Priya Natarajan, Kristian Pedersen, Eric Jullo, and Danuta Paraficz. Where is the Matter in the Merging Cluster Abell 2218? Technical report, 2007.

[53] Rouven Essig, Samuel D. Mcdermott, Hai-Bo Yu, and Yi-Ming Zhong. Constraining Dissipative Dark Matter Self-Interactions. *Phys. Rev. Lett.*, 123(12):121102, 2019. doi: 10.1103/PhysRevLett.123.121102.

[54] S. M. Faber and R. E. Jackson. Velocity dispersions and mass-to-light ratios for elliptical galaxies. *The Astrophysical Journal*, 204:668, mar 1976. ISSN 0004-637X. doi: 10.1086/154215. URL `http://adsabs.harvard.edu/doi/10.1086/154215`.

[55] Jonathan L. Feng, Manoj Kaplinghat, and Hai-Bo Yu. Sommerfeld enhancements for thermal relic dark matter. *Phys. Rev. D*, 82(8):083525, Oct 2010. ISSN 15507998. doi: 10.1103/PhysRevD.82.083525. URL `https://link.aps.org/doi/10.1103/PhysRevD.82.083525`.

[56] F. Feroz, M. P. Hobson, and M. Bridges. MultiNest: An efficient and robust Bayesian inference tool for cosmology and particle physics. *Monthly Notices of the Royal Astronomical Society*, 398(4):1601–1614, oct 2009. ISSN 00358711. doi: 10.1111/j.1365-2966.2009.14548.x. URL `http://arxiv.org/abs/0809.3437http:`

//dx.doi.org/10.1111/j.1365-2966.2009.14548.xhttps://academic.oup.com/mnras/article-lookup/doi/10.1111/j.1365-2966.2009.14548.x.

[57] Catherine E Fielder, Yao-Yuan Mao, Andrew R Zentner, Jeffrey A Newman, Hao-Yi Wu, and Risa H Wechsler. Illuminating dark matter halo density profiles without subhaloes. *Monthly Notices of the Royal Astronomical Society*, 499(2):2426–2444, Sep 2020. ISSN 1365-2966. doi: 10.1093/mnras/staa2851. URL http://dx.doi.org/10.1093/mnras/staa2851.

[58] Ricardo A Flores and Joel R Primack. Observational and Theoretical Constraints on Singular Dark Matter Halos. *Astrophysical Journal Letters*, 427:L1, may 1994. ISSN 0004-637X. doi: 10.1086/187350.

[59] Daniel Foreman-Mackey, Will Farr, Manodeep Sinha, Anne Archibald, David Hogg, Jeremy Sanders, Joe Zuntz, Peter Williams, Andrew Nelson, Miguel de Val-Borro, Tobias Erhardt, Ilya Pashchenko, and Oriol Pla. emcee v3: A Python ensemble sampling toolkit for affine-invariant MCMC. *Journal of Open Source Software*, 4(43):1864, nov 2019. ISSN 2475-9066. doi: 10.21105/joss.01864.

[60] L. Gao, J. F. Navarro, C. S. Frenk, A. Jenkins, V. Springel, and S. D. M. White. The Phoenix Project: the dark side of rich Galaxy clusters. *Monthly Notices of the Royal Astronomical Society*, 425(3):2169–2186, sep 2012. ISSN 00358711. doi: 10.1111/j.1365-2966.2012.21564.x. URL https://academic.oup.com/mnras/article-lookup/doi/10.1111/j.1365-2966.2012.21564.x.

[61] L. Girardi, E. K. Grebel, M. Odenkirchen, and C. Chiosi. Theoretical isochrones in several photometric systems. II. The Sloan Digital Sky Survey ugriz system. *Astrononmy and Astrophysics*, 422:205–215, July 2004. doi: 10.1051/0004-6361:20040250.

[62] Ghislain Golse and Jean-Paul Kneib. Pseudo Elliptical Lensing Mass Model: application to the NFW mass distribution. *Astronomy & Astrophysics*, 390:821–827, 2002. doi: 10.1051/0004-6361:20020639. URL https://arxiv.org/pdf/astro-ph/0112138.pdf.

[63] Ghislain Golse and Jean-Paul J.-P Kneib. Pseudo Elliptical Lensing Mass Model: application to the NFW mass distribution. *A&A*, 390:821–827, 2002. doi: 10.1051/0004-6361:20020639. URL http://dx.doi.org/10.1051/0004-6361:20020639https://arxiv.org/pdf/astro-ph/0112138.pdf.

[64] F. Governato, A. Zolotov, A. Pontzen, C. Christensen, S. H. Oh, A. M. Brooks, T. Quinn, S. Shen, and J. Wadsley. Cuspy no more: how outflows affect the central dark matter and baryon distribution in Λ cold dark matter galaxies. *Monthly Notices of the Royal Astronomical Society*, 422(2):1231–1240, may 2012. ISSN 00358711. doi: 10.1111/j.1365-2966.2012.20696.x. URL https://academic.oup.com/mnras/article-lookup/doi/10.1111/j.1365-2966.2012.20696.x.

[65] W. J. Handley, M. P. Hobson, and A. N. Lasenby. PolyChord: next-generation nested sampling. *Monthly Notices of the Royal Astronomical Society*, 453(4):4384–4398, may

2015. ISSN 13652966. doi: 10.1093/mnras/stv1911. URL `http://arxiv.org/abs/1506.00171http://dx.doi.org/10.1093/mnras/stv1911`.

[66] David Harvey, Richard Massey, Thomas Kitching, Andy Taylor, and Eric Tittley. The nongravitational interactions of dark matter in colliding galaxy clusters. *Science*, 347(6229):1462–1465, mar 2015. ISSN 10959203. doi: 10.1126/science.1261381. URL `https://science.sciencemag.org/content/347/6229/1462https://science.sciencemag.org/content/347/6229/1462.abstract`.

[67] David Harvey, F Courbin, J P Kneib, and Ian G. McCarthy. A detection of wobbling brightest cluster galaxies within massive galaxy clusters. *Monthly Notices of the Royal Astronomical Society*, 472(2):1972–1980, 2017. ISSN 13652966. doi: 10.1093/MNRAS/STX2084. URL `https://academic.oup.com/mnras/article/472/2/1972/4082236`.

[68] David Harvey, Andrew Robertson, Richard Massey, and Ian G. McCarthy. Observable tests of self-interacting dark matter in galaxy clusters: BCG wobbles in a constant density core. *Monthly Notices of the Royal Astronomical Society*, 488(2):1572–1579, jul 2019. ISSN 13652966. doi: 10.1093/mnras/stz1816.

[69] Qiuhan He, Hongyu Li, Ran Li, Carlos S Frenk, Matthieu Schaller, David Barnes, Yannick Bahé, Scott T Kay, Liang Gao, and Claudio Dalla Vecchia. Constraining the inner density slope of massive galaxy clusters. *Monthly Notices of the Royal Astronomical Society*, 496(4):4717–4733, jun 2020. ISSN 1365-2966. doi: 10.1093/mnras/staa1769. URL `http://dx.doi.org/10.1093/mnras/staa1769`.

[70] Joseph F Hennawi, Neal Dalal, Paul Bode, and Jeremiah P Ostriker. Characterizing the Cluster Lens Population. *The Astrophysical Journal*, 654(2):714–730, 2007. ISSN 0004-637X. doi: 10.1086/497362. URL `http://healpix.jpl.nasa.gov.`

[71] Lars Hernquist. An analytical model for spherical galaxies and bulges. *The Astrophysical Journal*, 356:359, 1990. ISSN 0004-637X. doi: 10.1086/168845.

[72] Edward Higson, Will Handley, Mike Hobson, and Anthony Lasenby. Dynamic nested sampling: an improved algorithm for parameter estimation and evidence calculation. *Statistics and Computing*, 29(5):891–913, September 2019. ISSN 15731375. doi: 10.1007/s11222-018-9844-0.

[73] Yonit Hochberg, Eric Kuflik, Hitoshi Murayama, Tomer Volansky, and Jay G. Wacker. Model for Thermal Relic Dark Matter of Strongly Interacting Massive Particles. *Physical Review Letters*, 115(2):021301, July 2015. doi: 10.1103/PhysRevLett.115.021301.

[74] M. T. Hogan, B. R. McNamara, F. Pulido, P. E. J. Nulsen, H. R. Russell, A. N. Vantyghem, A. C. Edge, and R. A. Main. Mass Distribution in Galaxy Cluster Cores. *The Astrophysical Journal*, 837(1):51, mar 2017. ISSN 1538-4357. doi: 10.3847/1538-4357/aa5f56. URL `http://stacks.iop.org/0004-637X/837/i=1/a=51?key=crossref.a751b331d93554c47bf22ffd7153123a`.

[75] John R. Jardel and Karl Gebhardt. The dark matter density profile of the Fornax dwarf. *Astrophysical Journal*, 746(1):89, jan 2012. ISSN 15384357. doi: 10.1088/0004-637X/746/1/89. URL `https://iopscience.iop.org/article/10.1088/0004-637X/746/1/89https://iopscience.iop.org/article/10.1088/0004-637X/746/1/89/meta`.

[76] John R. Jardel, Karl Gebhardt, Maximilian H. Fabricius, Niv Drory, and Michael J. Williams. Measuring dark matter profiles non-parametrically in dwarf spheroidals: An application to Draco. *Astrophysical Journal*, 763(2):91, jan 2013. ISSN 15384357. doi: 10.1088/0004-637X/763/2/91. URL `https://iopscience.iop.org/article/10.1088/0004-637X/763/2/91https://iopscience.iop.org/article/10.1088/0004-637X/763/2/91/meta`.

[77] Felix Kahlhoefer, Kai Schmidt-Hoberg, Mads T. Frandsen, and Subir Sarkar. Colliding clusters and dark matter self-interactions. *Monthly Notices of the Royal Astronomical Society*, 437(3):2865–2881, jan 2014. ISSN 00358711. doi: 10.1093/mnras/stt2097.

[78] M. Kaplinghat, S. Tulin, and H.-B. Yu. Dark Matter Halos as Particle Colliders: Unified Solution to Small-Scale Structure Puzzles from Dwarfs to Clusters. *Physical Review Letters*, 116(4):041302, January 2016. doi: 10.1103/PhysRevLett.116.041302.

[79] Manoj Kaplinghat, Ryan E Keeley, Tim Linden, and Hai Bo Yu. Tying dark matter to baryons with self-interactions. *Physical Review Letters*, 113(2), 2014. ISSN 10797114. doi: 10.1103/PhysRevLett.113.021302.

[80] Manoj Kaplinghat, Sean Tulin, and Hai Bo Yu. Dark Matter Halos as Particle Colliders: Unified Solution to Small-Scale Structure Puzzles from Dwarfs to Clusters. *Physical Review Letters*, 116(4):041302, jan 2016. ISSN 10797114. doi: 10.1103/PhysRevLett.116.041302.

[81] Manoj Kaplinghat, Tao Ren, and Hai-Bo Yu. Dark matter cores and cusps in spiral galaxies and their explanations, 2019. ISSN 13652966.

[82] Manoj Kaplinghat, Tao Ren, and Hai Bo Yu. Dark matter cores and cusps in spiral galaxies and their explanations. *Journal of Cosmology and Astroparticle Physics*, 2020 (6), 2020. ISSN 14757516. doi: 10.1088/1475-7516/2020/06/027. URL `http://arxiv.org/abs/1911.00544`.

[83] G. Kauffmann, S. D. M. White, and B. Guiderdoni. The formation and evolution of galaxies within merging dark matter haloes*. *Monthly Notices of the Royal Astronomical Society*, 264(1):201–218, sep 1993. ISSN 0035-8711. doi: 10.1093/mnras/264.1.201.

[84] Charles R Keeton. Computational Methods for Gravitational Lensing. Technical report, February 2001. URL `http://cfa-www.harvard.edu/castles.`

[85] Charles R Keeton. A Catalog of Mass Models for Gravitational Lensing. Technical report, 2001. URL `http://arxiv.org/abs/astro-ph/0102341`.

[86] Stacy Y Kim, Annika H.G. Peter, and David Wittman. In the wake of dark giants: New signatures of dark matter self-interactions in equal-mass mergers of galaxy clusters. *Monthly Notices of the Royal Astronomical Society*, 469(2):1414–1444, 2017. ISSN 13652966. doi: 10.1093/MNRAS/STX896.

[87] Anatoly Klypin, Andrey V. Kravtsov, Octavio Valenzuela, and Francisco Prada. Where Are the Missing Galactic Satellites? *The Astrophysical Journal*, 522(1):82–92, sep 1999. ISSN 0004-637X. doi: 10.1086/307643.

[88] Jean-Paul Kneib and Priyamvada Natarajan. Cluster lenses. *The Astronomy and Astrophysics Review*, 19(1):47, dec 2011. ISSN 0935-4956. doi: 10.1007/s00159-011-0047-3. URL `http://link.springer.com/10.1007/s00159-011-0047-3`.

[89] Jean-Paul Kneib, Patrick Hudelot, Richard S Ellis, Tommaso Treu, Graham P Smith, Phil Marshall, Oliver Czoske, Ian Smail, and Priyamvada Natarajan. A Wide-Field Hubble Space Telescope Study of the Cluster Cl 0024+1654 at z=0.4. II. The Cluster Mass Distribution. *Astrophysical Journal*, 598(2):804–817, dec 2003. doi: 10.1086/378633.

[90] C. S. Kochanek and N. Dalal. Tests for Substructure in Gravitational Lenses. *Astrophysical Journal*, 610:69–79, July 2004. doi: 10.1086/421436.

[91] Klaudia Kowalczyk, Ewa L. Łokas, and Monica Valluri. Recovering the mass profile and orbit anisotropy of mock dwarf galaxies with Schwarzschild modelling. *Monthly Notices of the Royal Astronomical Society*, 470(4):3959–3969, oct 2017. ISSN 13652966. doi: 10.1093/mnras/stx1520. URL `https://academic.oup.com/mnras/article/470/4/3959/3869625`.

[92] Pavel Kroupa. The Initial Mass Function of Stars: Evidence for Uniformity in Variable Systems. *Science*, 295(5552):82–91, jan 2002. ISSN 0036-8075. doi: 10.1126/SCIENCE.1067524. URL `http://science.sciencemag.org/content/295/5552/82.full`.

[93] F. La Barbera, I. Ferreras, A. Vazdekis, I. G. de la Rosa, R. R. de Carvalho, M. Trevisan, J. Falcón-Barroso, and E. Ricciardelli. SPIDER VIII - constraints on the stellar initial mass function of early-type galaxies from a variety of spectral features. *Monthly Notices of the Royal Astronomical Society*, 433:3017–3047, August 2013. doi: 10.1093/mnras/stt943.

[94] F. La Barbera, A. Vazdekis, I. Ferreras, A. Pasquali, M. Cappellari, I. Martín-Navarro, F. Schönebeck, and J. Falcón-Barroso. Radial constraints on the initial mass function from TiO features and Wing-Ford band in early-type galaxies. *Monthly Notices of the Royal Astronomical Society*, 457:1468–1489, April 2016. doi: 10.1093/mnras/stv2996.

[95] D. Leier, I. Ferreras, P. Saha, S. Charlot, G. Bruzual, and F. La Barbera. Strong gravitational lensing and the stellar IMF of early-type galaxies. *Monthly Notices of the Royal Astronomical Society*, 459:3677–3692, July 2016. doi: 10.1093/mnras/stw885.

[96] Antony Lewis. GetDist: a Python package for analysing Monte Carlo samples. 2019. URL `https://getdist.readthedocs.io`.

[97] M Limousin, J P Kneib, S Bardeau, P Natarajan, O Czoske, I Smail, H Ebeling, and G P Smith. Truncation of galaxy dark matter halos in high density environments. *Astronomy & Astrophysics*, 461(3):881–891, 2007. ISSN 0004-6361. doi: 10.1051/0004-6361:20065543. URL `http://dx.doi.org/10.1051/0004-6361:20065543`.

[98] Marceau Limousin, Jean-Paul Paul Kneib, and Priyamvada Natarajan. Constraining the mass distribution of galaxies using galaxy-galaxy lensing in clusters and in the field. *Mon. Not. R. Astron. Soc*, 000(1):1–12, 2004. ISSN 00358711. doi: 10.1111/j.1365-2966.2004.08449.x. URL `http://webast.ast.obs-mip.fr/zCosmos`.

[99] Marceau Limousin, Johan Richard, Eric Jullo, Jean-Paul Kneib, Bernard Fort, Genevieve Soucail, Ardis Eliasdottir, Priyamvada Natarajan, Richard S Ellis, Ian Smail, Oliver Czoske, Graham P Smith, Patrick Hudelot, Sebastien Bardeau, Harald Ebeling, Eiichi Egami, and Kirsten K Knudsen. Combining Strong and Weak Gravitational Lensing in Abell 1689. *The Astrophysical Journal*, 668(2):643–666, 2007. ISSN 0004-637X. doi: 10.1086/521293.

[100] Abraham Loeb and Neal Weiner. Cores in Dwarf Galaxies from Dark Matter with a Yukawa Potential. *Physics Review Letters*, 106(17):171302, apr 2011. doi: 10.1103/PhysRevLett.106.171302.

[101] Ewa L. Łokas and Gary A. Mamon. Dark matter distribution in the Coma cluster from galaxy kinematics: Breaking the mass-anisotropy degeneracy. *Monthly Notices of the Royal Astronomical Society*, 343(2):401–412, aug 2003. ISSN 00358711. doi: 10.1046/j.1365-8711.2003.06684.x. URL `https://academic.oup.com/mnras/article/343/2/401/1038976`.

[102] M. Lyubenova, I. Martín-Navarro, G. van de Ven, J. Falcón-Barroso, L. Galbany, A. Gallazzi, R. García-Benito, R. González Delgado, B. Husemann, F. La Barbera, R. A. Marino, D. Mast, J. Mendez-Abreu, R. F. P. Peletier, P. Sánchez-Blázquez, S. F. Sánchez, S. C. Trager, R. C. E. van den Bosch, A. Vazdekis, C. J. Walcher, L. Zhu, S. Zibetti, B. Ziegler, J. Bland-Hawthorn, and CALIFA Collaboration. IMF shape constraints from stellar populations and dynamics from CALIFA. *Monthly Notices of the Royal Astronomical Society*, 463:3220–3225, December 2016. doi: 10.1093/mnras/stw2434.

[103] Rachel Mandelbaum, Uroš Seljak, Richard J Cool, Michael Blanton, Christopher M Hirata, and Jonathan Brinkmann. Density profiles of galaxy groups and clusters from SDSS galaxy-galaxy weak lensing. *Monthly Notices of the Royal Astronomical Society*, 372(2):758–776, oct 2006. doi: 10.1111/j.1365-2966.2006.10906.x.

[104] Andrew W. Mann and Harald Ebeling. X-ray-optical classification of cluster mergers and the evolution of the cluster merger fraction. *Monthly Notices of the Royal Astronomical Society*, 420(3):2120–2138, Mar 2012. doi: 10.1111/j.1365-2966.2011.20170.x.

[105] Federico Marinacci, Mark Vogelsberger, Rüdiger Pakmor, Paul Torrey, Volker Springel, Lars Hernquist, Dylan Nelson, Rainer Weinberger, Annalisa Pillepich, Jill Naiman, and et al. First results from the illustristng simulations: radio haloes and magnetic fields. *Monthly Notices of the Royal Astronomical Society*, Aug 2018. ISSN 1365-2966. doi: 10.1093/mnras/sty2206. URL `http://dx.doi.org/10.1093/mnras/sty2206`.

[106] Davide Martizzi, Romain Teyssier, Ben Moore, and Tina Wentz. The effects of baryon physics, black holes and active galactic nucleus feedback on the mass distribution in clusters of galaxies. *Monthly Notices of the Royal Astronomical Society*, 422(4):3081–3091, jun 2012. ISSN 00358711. doi: 10.1111/j.1365-2966.2012.20879.x. URL `https://academic.oup.com/mnras/article-lookup/doi/10.1111/j.1365-2966.2012.20879.x`.

[107] Davide Martizzi, Romain Teyssier, and Ben Moore. Cusp-core transformations induced by AGN feedback in the progenitors of cluster galaxies. *Monthly Notices of the Royal Astronomical Society*, 432(3):1947–1954, July 2013. doi: 10.1093/mnras/stt297.

[108] Mario Mateo, Edward W. Olszewski, and Matthew G. Walker. The Velocity Dispersion Profile of the Remote Dwarf Spheroidal Galaxy Leo I: A Tidal Hit and Run? *The Astrophysical Journal*, 675(1):201–233, mar 2008. ISSN 0004-637X. doi: 10.1086/522326. URL `https://iopscience.iop.org/article/10.1086/522326https://iopscience.iop.org/article/10.1086/522326/meta`.

[109] Alan W. McConnachie. The observed properties of dwarf galaxies in and around the Local Group. *Astronomical Journal*, 144(1):4–36, jul 2012. ISSN 00046256. doi: 10.1088/0004-6256/144/1/4. URL `https://iopscience.iop.org/article/10.1088/0004-6256/144/1/4https://iopscience.iop.org/article/10.1088/0004-6256/144/1/4/meta`.

[110] James M.G. Mead, Lindsay J. King, Debora Sijacki, Adrienne Leonard, Ewald Puchwein, and Ian G. McCarthy. The impact of AGN feedback and baryonic cooling on galaxy clusters as gravitational lenses. *Monthly Notices of the Royal Astronomical Society*, 406(1):434–444, 2010. ISSN 00358711. doi: 10.1111/j.1365-2966.2010.16674.x. URL `http://dx.doi.org/10.1111/j.1365-2966.2010.16674.x`.

[111] M. Meneghetti, C. Fedeli, F. Pace, S. Gottlöber, and G. Yepes. Strong lensing in the MARENOSTRUM UNIVERSE. *Astronomy and Astrophysics*, 519:A90, 2010. ISSN 0004-6361. doi: 10.1051/0004-6361/201014098. URL `https://ui.adsabs.harvard.edu/abs/2010A{%}26A...519A..90M/abstract`.

[112] Michael R Merrifield and Stephen M Kent. Fourth Moments and the Dynamics of Spherical Systems. *Astronomical Journal*, 99:1548, may 1990. doi: 10.1086/115438.

[113] J Merten, D. Coe, R Dupke, R Massey, A Zitrin, E S Cypriano, N Okabe, B Frye, F G Braglia, Y Jiménez-Teja, N Benítez, T Broadhurst, J Rhodes, M Meneghetti, L A Moustakas, L Sodré, J Krick, and J N Bregman. Creation of cosmic structure in the complex galaxy cluster merger Abell 2744. *Monthly Notices of the Royal Astronomical*

Society, 417(1):333–347, 2011. ISSN 00358711. doi: 10.1111/j.1365-2966.2011.19266.x. URL `https://academic.oup.com/mnras/article/417/1/333/979730`.

[114] J. Merten, M. Meneghetti, M. Postman, K. Umetsu, A. Zitrin, E. Medezinski, M. Nonino, A. Koekemoer, P. Melchior, D. Gruen, L. A. Moustakas, M. Bartelmann, O. Host, M. Donahue, D. Coe, A. Molino, S. Jouvel, A. Monna, S. Seitz, N. Czakon, D. Lemze, J. Sayers, I. Balestra, P. Rosati, N. Benítez, A. Biviano, R. Bouwens, L. Bradley, T. Broadhurst, M. Carrasco, H. Ford, C. Grillo, L. Infante, D. Kelson, O. Lahav, R. Massey, J. Moustakas, E. Rasia, J. Rhodes, J. Vega, and W. Zheng. CLASH: The Concentration-Mass Relation of Galaxy Clusters. *Astrophysical Journal*, 806(1): 4, June 2015. ISSN 15384357. doi: 10.1088/0004-637X/806/1/4.

[115] Jordi Miralda-Escude. Gravitational Lensing by a Cluster of Galaxies and the Central cD Galaxy: Measuring the Mass Profile. *Astrophysical Journal*, 438:514, jan 1995. doi: 10.1086/175098.

[116] A. Monna, S. Seitz, A. Zitrin, M. J. Geller, C. Grillo, A. Mercurio, N. Greisel, A. Halkola, S. H. Suyu, M. Postman, P. Rosati, I. Balestra, A. Biviano, D. Coe, D. G. Fabricant, H. S. Hwang, and A. Koekemoer. Constraining the galaxy mass content in the core of A383 using velocity dispersion measurements for individual cluster members. *Monthly Notices of the Royal Astronomical Society*, 447(2):1224–1241, Feb 2015. doi: 10.1093/mnras/stu2534.

[117] A. Monna, S. Seitz, I. Balestra, P. Rosati, C. Grillo, A. Halkola, S. H. Suyu, D. Coe, G. B. Caminha, B. Frye, A. Koekemoer, A. Mercurio, M. Nonino, M. Postman, and A. Zitrin. Precise strong lensing mass profile of the CLASH galaxy cluster MACS 2129. *Monthly Notices of the Royal Astronomical Society*, 466(4):4094–4106, Apr 2017. doi: 10.1093/mnras/stx015.

[118] A. Monna, S. Seitz, M. J. Geller, A. Zitrin, A. Mercurio, S. H. Suyu, M. Postman, D. G. Fabricant, H. S. Hwang, and A. Koekemoer. Separating galaxies from the cluster dark matter halo in Abell 611. *Monthly Notices of the Royal Astronomical Society*, 465(4):4589–4601, 2017. ISSN 13652966. doi: 10.1093/mnras/stw3048. URL `https://arxiv.org/pdf/1602.08491.pdf`.

[119] A Monna, S Seitz, I Balestra, P Rosati, C Grillo, A Halkola, S H Suyu, D Coe, G B Caminha, B Frye, A Koekemoer, A Mercurio, M Nonino, M Postman, and A Zitrin. Precise strong lensing mass profile of the CLASH galaxy cluster MACS 2129. Technical Report 0000, 2018. URL `http://archive.stsci.edu/prepds/clash,`.

[120] Ben Moore. The Nature Of Dark Matter. *arXiv e-prints*, 370(6491):astro–ph/9402009, feb 1994. ISSN 00280836. doi: 10.1038/370629a0. URL `https://www.nature.com/articles/370629a0`.

[121] J. A. Munoz, C S Kochanek, and C R Keeton. Cusped Mass Models of Gravitational Lenses. *The Astrophysical Journal*, 558(2):657–665, 2001. ISSN 0004-637X. doi: 10.1086/322314. URL `https://arxiv.org/pdf/astro-ph/0103009.pdfhttp://stacks.iop.org/0004-637X/558/i=2/a=657`.

[122] Ricardo R Muñoz, Patrick Côté, Felipe A Santana, Marla Geha, Joshua D Simon, Grecco A Oyarzún, Peter B Stetson, and S G Djorgovski. A MegaCam Survey of Outer Halo Satellites. III. Photometric and Structural Parameters * †. *The Astrophysical Journal*, 860:66, 2018. doi: 10.3847/1538-4357/aac16b. URL https://doi.org/10.3847/1538-4357/aac16b.

[123] Jill P Naiman, Annalisa Pillepich, Volker Springel, Enrico Ramirez-Ruiz, Paul Torrey, Mark Vogelsberger, Rüdiger Pakmor, Dylan Nelson, Federico Marinacci, Lars Hernquist, and et al. First results from the illustristng simulations: a tale of two elements – chemical evolution of magnesium and europium. *Monthly Notices of the Royal Astronomical Society*, 477(1):1206–1224, Mar 2018. ISSN 1365-2966. doi: 10.1093/mnras/sty618. URL http://dx.doi.org/10.1093/mnras/sty618.

[124] Priyamvada Natarajan and Jean-Paul Kneib. Probing the dynamics of cluster-lenses. *Monthly Notices of the Royal Astronomical Society*, 283(3):1031–1046, dec 1996. doi: 10.1093/mnras/283.3.1031.

[125] Julio F. Navarro, Carlos S. Frenk, and Simon D. M. White. The Structure of Cold Dark Matter Halos. *The Astrophysical Journal*, 462:563, aug 1996. ISSN 0004-637X. doi: 10.1086/177173. URL http://arxiv.org/abs/astro-ph/9508025http://dx.doi.org/10.1086/177173http://adsabs.harvard.edu/doi/10.1086/177173.

[126] Julio F. Navarro, Aaron Ludlow, Volker Springel, Jie Wang, Mark Vogelsberger, Simon D. M. White, Adrian Jenkins, Carlos S. Frenk, and Amina Helmi. The diversity and similarity of simulated cold dark matter haloes. *Monthly Notices of the Royal Astronomical Society*, 402(1):21–34, feb 2010. ISSN 00358711. doi: 10.1111/j.1365-2966.2009.15878.x. URL https://academic.oup.com/mnras/article-lookup/doi/10.1111/j.1365-2966.2009.15878.x.

[127] Dylan Nelson, Annalisa Pillepich, Volker Springel, Rainer Weinberger, Lars Hernquist, Rüdiger Pakmor, Shy Genel, Paul Torrey, Mark Vogelsberger, Guinevere Kauffmann, and et al. First results from the illustristng simulations: the galaxy colour bimodality. *Monthly Notices of the Royal Astronomical Society*, 475(1):624–647, Nov 2017. ISSN 1365-2966. doi: 10.1093/mnras/stx3040. URL http://dx.doi.org/10.1093/mnras/stx3040.

[128] Dylan Nelson, Volker Springel, Annalisa Pillepich, Vicente Rodriguez-Gomez, Paul Torrey, Shy Genel, Mark Vogelsberger, Ruediger Pakmor, Federico Marinacci, Rainer Weinberger, Luke Kelley, Mark Lovell, Benedikt Diemer, and Lars Hernquist. The IllustrisTNG Simulations: Public Data Release. *Computational Astrophysics and Cosmology*, 6(1), dec 2018. URL http://arxiv.org/abs/1812.05609.

[129] A. F. Neto, L. Gao, P. Bett, S. Cole, J. F. Navarro, C. S. Frenk, S. D. M. White, V. Springel, and A. Jenkins. The statistics of CDM halo concentrations. *Monthly Notices of the Royal Astronomical Society*, 381(4):1450–1462, nov 2007. ISSN 0035-8711. doi: 10.1111/j.1365-2966.2007.12381.x. URL https://academic.oup.com/mnras/article-lookup/doi/10.1111/j.1365-2966.2007.12381.x.

[130] Andrew B. Newman, Tommaso Treu, Richard S. Ellis, David J. Sand, Johan Richard, Philip J. Marshall, Peter Capak, and Satoshi Miyazaki. THE DISTRIBUTION OF DARK MATTER OVER THREE DECADES IN RADIUS IN THE LENSING CLUSTER ABELL 611. *The Astrophysical Journal*, 706(2):1078–1094, dec 2009. ISSN 0004-637X. doi: 10.1088/0004-637X/706/2/1078. URL `http://stacks.iop.org/0004-637X/706/i=2/a=1078?key=crossref.6b569ac352c61b00e064c226dbdc7022`.

[131] Andrew B Newman, Tommaso Treu, Richard S Ellis, and David J Sand. THE DARK MATTER DISTRIBUTION IN A383: EVIDENCE FOR A SHALLOW DENSITY CUSP FROM IMPROVED LENSING, STELLAR KINEMATIC, AND X-RAY DATA. *The Astrophysical Journal Letters*, 728(6pp):39, 2011. doi: 10.1088/2041-8205/728/2/L39. URL `http://www.oamp.fr/cosmology/lenstool/`.

[132] Andrew B. Newman, Tommaso Treu, Richard S. Ellis, and David J. Sand. THE DENSITY PROFILES OF MASSIVE, RELAXED GALAXY CLUSTERS. II. SEPARATING LUMINOUS AND DARK MATTER IN CLUSTER CORES. *The Astrophysical Journal*, 765(1):25, feb 2013. ISSN 0004-637X. doi: 10.1088/0004-637X/765/1/25. URL `http://stacks.iop.org/0004-637X/765/i=1/a=25?key=crossref.7a6560a366829f87d7e609dbd6d266cb`.

[133] Andrew B. Newman, Tommaso Treu, Richard S. Ellis, David J. Sand, Carlo Nipoti, Johan Richard, and Eric Jullo. THE DENSITY PROFILES OF MASSIVE, RELAXED GALAXY CLUSTERS. I. THE TOTAL DENSITY OVER THREE DECADES IN RADIUS. *The Astrophysical Journal*, 765(1):24, feb 2013. ISSN 0004-637X. doi: 10.1088/0004-637X/765/1/24. URL `http://stacks.iop.org/0004-637X/765/i=1/a=24?key=crossref.813111976314f342ad821f273765617a`.

[134] Anna Niemiec, Eric Jullo, Marceau Limousin, Carlo Giocoli, Thomas Erben, Hendrik Hildebrant, Jean-Paul Kneib, Alexie Leauthaud, Martin Makler, Bruno Moraes, and et al. Stellar-to-halo mass relation of cluster galaxies. *Monthly Notices of the Royal Astronomical Society*, 471(1):1153–1166, Jul 2017. ISSN 1365-2966. doi: 10.1093/mnras/stx1667. URL `http://dx.doi.org/10.1093/mnras/stx1667`.

[135] Hiroya Nishikawa, Kimberly K. Boddy, and Manoj Kaplinghat. Accelerated core collapse in tidally stripped self-interacting dark matter halos. *Phys. Rev. D*, 101(6): 063009, 2020. doi: 10.1103/PhysRevD.101.063009.

[136] Masamune Oguri, Matthew B Bayliss, Håkon Dahle, Keren Sharon, Michael D Gladders, Priyamvada Natarajan, Joseph F Hennawi, and Benjamin P Koester. Combined strong and weak lensing analysis of 28 clusters from the Sloan Giant Arcs Survey. *Monthly Notices of the Royal Astronomical Society*, 420(4):3213–3239, 2012. ISSN 00358711. doi: 10.1111/j.1365-2966.2011.20248.x. URL `https://academic.oup.com/mnras/article/420/4/3213/973116`.

[137] Nobuhiro Okabe, Masahiro Takada, Keiichi Umetsu, Toshifumi Futamase, and Graham P Smith. LoCuSS: Subaru Weak Lensing Study of 30 Galaxy Clusters. *Publications of the Astronomical Society of Japan*, 62:811, jun 2010. doi: 10.1093/pasj/62.3.811.

[138] Andrew B. Pace, Manoj Kaplinghat, Evan Kirby, Joshua D. Simon, Erik Tollerud, Ricardo R. Muñoz, Patrick Côté, S. G. Djorgovski, and Marla Geha. Multiple chemodynamic stellar populations of the Ursa Minor dwarf spheroidal galaxy. *Monthly Notices of the Royal Astronomical Society*, 495(3):3022–3040, may 2020. ISSN 13652966. doi: 10.1093/mnras/staa1419. URL `https://academic.oup.com/mnras/article/495/3/3022/5841952`.

[139] Ekta Patel, Nitya Kallivayalil, Nicolas Garavito-Camargo, Gurtina Besla, Daniel R. Weisz, Roeland P. van der Marel, Michael Boylan-Kolchin, Marcel S. Pawlowski, and Facundo A. Gómez. The Orbital Histories of Magellanic Satellites Using Gaia DR2 Proper Motions. *The Astrophysical Journal*, 893 (2):121, apr 2020. ISSN 0004-637X. doi: 10.3847/1538-4357/ab7b75. URL `https://iopscience.iop.org/article/10.3847/1538-4357/ab7b75https://iopscience.iop.org/article/10.3847/1538-4357/ab7b75/meta`.

[140] Fabian Pedregosa, Vincent Michel, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Jake Vanderplas, David Cournapeau, Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Bertrand Thirion, Olivier Grisel, Vincent Dubourg, Alexandre Passos, Matthieu Brucher, Matthieu Perrot andÉdouardand, AndÉdouard Duchesnay, and FRÉdouard Duchesnay. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12(85):2825–2830, 2011. ISSN 1533-7928. URL `http://scikit-learn.sourceforge.net.`

[141] S Peirani, S Kay, and J Silk. Active galactic nuclei and massive galaxy cores. *Astronomy & Astrophysics*, 479(1):123–129, February 2008. ISSN 0004-6361. doi: 10.1051/0004-6361:20077956. URL `http://dx.doi.org/10.1051/0004-6361:20077956`.

[142] Sébastien Peirani, Yohan Dubois, Marta Volonteri, Julien Devriendt, Kevin Bundy, Joe Silk, Christophe Pichon, Sugata Kaviraj, Raphaël Gavazzi, and Mélanie Habouzit. Density profile of dark matter haloes and galaxies in the horizon–agn simulation: the impact of agn feedback. *Monthly Notices of the Royal Astronomical Society*, 472(2): 2153–2169, Aug 2017. ISSN 1365-2966. doi: 10.1093/mnras/stx2099. URL `http://dx.doi.org/10.1093/mnras/stx2099`.

[143] J Peñarrubia, A Pontzen, M.˜G. Walker, and S.˜E. Koposov. The Coupling between the Core/Cusp and Missing Satellite Problems. *The Astrophysical Journal*, 759(2): L42, nov 2012. ISSN 20418205. doi: 10.1088/2041-8205/759/2/L42.

[144] Annalisa Pillepich, Dylan Nelson, Lars Hernquist, Volker Springel, Rüdiger Pakmor, Paul Torrey, Rainer Weinberger, Shy Genel, Jill P Naiman, Federico Marinacci, and et al. First results from the illustristng simulations: the stellar mass content of groups and clusters of galaxies. *Monthly Notices of the Royal Astronomical Society*, 475(1): 648–675, Dec 2017. ISSN 1365-2966. doi: 10.1093/mnras/stx3112. URL `http://dx.doi.org/10.1093/mnras/stx3112`.

[145] H. C. Plummer. On the Problem of Distribution in Globular Star Clusters: (Plate 8.). *Monthly Notices of the Royal Astronomical Society*, 71(5):460–470, mar 1911.

ISSN 0035-8711. doi: 10.1093/MNRAS/71.5.460. URL `https://academic.oup.com/mnras/article/71/5/460/972150`.

[146] Andrew Pontzen and Fabio Governato. How supernova feedback turns dark matter cusps into cores. *Monthly Notices of the Royal Astronomical Society*, 421(4):3464–3471, 2012. ISSN 00358711. doi: 10.1111/j.1365-2966.2012.20571.x. URL `https://academic.oup.com/mnras/article-abstract/421/4/3464/1097213`.

[147] Marc Postman, Dan Coe, Narciso Benitez, Larry Bradley, Tom Broadhurst, Megan Donahue, Holland Ford, Or Graur, Genevieve Graves, Stephanie Jouvel, Anton Koekemoer, Doron Lemze, Elinor Medezinski, Alberto Molino, Leonidas Moustakas, Sara Ogaz, Adam Riess, Steve Rodney, Piero Rosati, Keiichi Umetsu, Wei Zheng, Adi Zitrin, Matthias Bartelmann, Rychard Bouwens, Nicole Czakon, Sunil Golwala, Ole Host, Leopoldo Infante, Saurabh Jha, Yolanda Jimenez-Teja, Daniel Kelson, Ofer Lahav, Ruth Lazkoz, Dani Maoz, Curtis McCully, Peter Melchior, Massimo Meneghetti, Julian Merten, John Moustakas, Mario Nonino, Brandon Patel, EnikÃ¶ RegÃ¶s, Jack Sayers, Stella Seitz, and Arjen Van der Wel. THE CLUSTER LENSING AND SUPERNOVA SURVEY WITH HUBBLE: AN OVERVIEW. *The Astrophysical Journal Supplement Series*, 199(2):25, mar 2012. doi: 10.1088/0067-0049/199/2/25.

[148] C Power, J F Navarro, A Jenkins, C S Frenk, S. D.M. White, V Springel, J Stadel, and T Quinn. The inner structure of Λ CDM haloes -I. A numerical convergence study. *Monthly Notices of the Royal Astronomical Society*, 338(1):14–34, 2003. ISSN 00358711. doi: 10.1046/j.1365-8711.2003.05925.x. URL `www.mpa-garching.mpg.de/gadget`.

[149] Cinthia Ragone-Figueroa, Gian Luigi Granato, and Mario G. Abadi. Effects of baryon mass loss on profiles of large galactic dark matter haloes. *Monthly Notices of the Royal Astronomical Society*, 423(4):3243–3250, jul 2012. ISSN 00358711. doi: 10.1111/j.1365-2966.2012.21113.x.

[150] Scott W. Randall, Maxim Markevitch, Douglas Clowe, Anthony H. Gonzalez, and Marusa Bradač. Constraints on the Self-Interaction Cross Section of Dark Matter from Numerical Simulations of the Merging Galaxy Cluster 1E 0657-56. *The Astrophysical Journal*, 679(2):1173–1180, jun 2008. ISSN 0004-637X. doi: 10.1086/587859.

[151] J I Read, G. Iorio, O Agertz, and F Fraternali. The stellar mass-halo mass relation of isolated field dwarfs: A critical test of ΛCDM at the edge of galaxy formation. *Monthly Notices of the Royal Astronomical Society*, 467(2):2019–2038, 2017. ISSN 13652966. doi: 10.1093/mnras/stx147.

[152] J I Read, M G Walker, and P Steger. Dark matter heats up in dwarf galaxies. *MNRAS*, 484:1401–1420, 2019. doi: 10.1093/mnras/sty3404. URL `https://academic.oup.com/mnras/article/484/1/1401/5265085`.

[153] J.˜I. Read, O Agertz, and M.˜L.˜M. Collins. Dark matter cores all the way down. *Monthly Notices of the Royal Astronomical Society*, 459(3):2573–2590, jul 2016. doi: 10.1093/mnras/stw713.

[154] Justin Read, Mark Gieles, and Daisuke Kawata. The Gaia Challenge, 2019. URL `http://astrowiki.ph.surrey.ac.uk/dokuwiki/doku.php`.

[155] Tao Ren, Anna Kwa, Manoj Kaplinghat, and Hai Bo Yu. Reconciling the Diversity and Uniformity of Galactic Rotation Curves with Self-Interacting Dark Matter. *Physical Review X*, 9(3), 2019. ISSN 21603308. doi: 10.1103/PhysRevX.9.031020. URL `https://arxiv.org/pdf/1808.05695.pdfhttp://arxiv.org/abs/1808.05695`.

[156] Johan Richard, Graham P. Smith, Jean-Paul Kneib, Richard S. Ellis, A. J. R. Sanderson, L. Pei, T. A. Targett, D. J. Sand, A. M. Swinbank, H. Dannerbauer, P. Mazzotta, M. Limousin, E. Egami, E. Jullo, V. Hamilton-Morris, and S. M. Moran. LoCuSS: first results from strong-lensing analysis of 20 massive galaxy clusters at <i>z</i> = 0.2. *Monthly Notices of the Royal Astronomical Society*, 404(1):325–349, mar 2010. ISSN 00358711. doi: 10.1111/j.1365-2966.2009.16274.x. URL `http://mnras.oxfordjournals.org/cgi/doi/10.1111/j.1365-2966.2009.16274.x`.

[157] Thomas Richardson and Malcolm Fairbairn. On the dark matter profile in sculptor: Breaking the β degeneracy with virial shape parameters. *Monthly Notices of the Royal Astronomical Society*, 441(2):1584–1600, jun 2014. ISSN 13652966. doi: 10.1093/mnras/stu691. URL `https://academic.oup.com/mnras/article/441/2/1584/1077653`.

[158] Andrew Robertson, Richard Massey, and Vincent Eke. What does the Bullet Cluster tell us about self-interacting dark matter? *Monthly Notices of the Royal Astronomical Society*, 465(1):569–587, 2017. ISSN 13652966. doi: 10.1093/mnras/stw2670. URL `https://academic.oup.com/mnras/article/465/1/569/2417398`.

[159] Andrew Robertson, Richard Massey, and Vincent Eke. Cosmic particle colliders: Simulations of self-interacting dark matter with anisotropic scattering. *Monthly Notices of the Royal Astronomical Society*, 467(4):4719–4730, jun 2017. ISSN 13652966. doi: 10.1093/mnras/stx463.

[160] Andrew Robertson, David Harvey, Richard Massey, Vincent Eke, Ian G McCarthy, Mathilde Jauzac, Baojiu Li, and Joop Schaye. Observable tests of self-interacting dark matter in galaxy clusters: cosmological simulations with SIDM and baryons. *Monthly Notices of the Royal Astronomical Society*, 488(3):3646–3662, sep 2019. ISSN 0035-8711. doi: 10.1093/mnras/stz1815. URL `https://academic.oup.com/mnras/article/488/3/3646/5530780`.

[161] Andrew Robertson, Richard Massey, Vincent Eke, Joop Schaye, and Tom Theuns. The surprising accuracy of isothermal Jeans modelling of self-interacting dark matter density profiles. *arXiv e-prints*, art. arXiv:2009.07844, September 2020.

[162] Victor H. Robles, Tyler Kelley, James S. Bullock, and Manoj Kaplinghat. The Milky Way's halo and subhaloes in self-interacting dark matter. *Monthly Notices of the Royal Astronomical Society*, 490(2):2117–2123, December 2019. doi: 10.1093/mnras/stz2345.

[163] Miguel Rocha, Annika H.G. Peter, James S Bullock, Manoj Kaplinghat, Shea Garrison-kimmel, Jose Oñorbe, and Leonidas A Moustakas. Cosmological simulations with self-interacting dark matter - I. Constant-density cores and substructure. *Monthly Notices of the Royal Astronomical Society*, 430(1):81–104, 2013. ISSN 00358711. doi: 10.1093/mnras/sts514. URL `https://academic.oup.com/mnras/article-abstract/430/1/81/984169`.

[164] A. Romano, L. Fu, F. Giordano, R. Maoli, P. Martini, M. Radovich, R. Scaramella, V. Antonuccio-Delogu, A. Donnarumma, S. Ettori, K. Kuijken, M. Meneghetti, L. Moscardini, S. Paulin-Henriksson, E. Giallongo, R. Ragazzoni, A. Baruffolo, A. DiPaola, E. Diolaiti, J. Farinato, A. Fontana, S. Gallozzi, A. Grazian, J. Hill, F. Pedichini, R. Speziali, R. Smareglia, and V. Testa. Abell 611. *Astronomy and Astrophysics*, 514:A88, may 2010. ISSN 0004-6361. doi: 10.1051/0004-6361/200912654. URL `http://www.aanda.org/10.1051/0004-6361/200912654`.

[165] Laura Sagunski, Sophia Gad-Nasr, Brian Colquhoun, Andrew Robertson, and Sean Tulin. Velocity-dependent self-interacting dark matter from groups and clusters of galaxies. *Journal of Cosmology and Astroparticle Physics*, 2021(1), 2021. ISSN 14757516. doi: 10.1088/1475-7516/2021/01/024.

[166] Edwin E. Salpeter. The Luminosity Function and Stellar Evolution. *The Astrophysical Journal*, 121:161, jan 1955. ISSN 0004-637X. doi: 10.1086/145971. URL `http://adsabs.harvard.edu/doi/10.1086/145971`.

[167] David J. Sand, Tommaso Treu, Graham P. Smith, and Richard S. Ellis. The Dark Matter Distribution in the Central Regions of Galaxy Clusters: Implications for Cold Dark Matter. *The Astrophysical Journal*, 604(1):88–107, 2004. ISSN 0004-637X. doi: 10.1086/382146. URL `http://iopscience.iop.org/article/10.1086/382146/pdf`.

[168] David J. Sand, Tommaso Treu, Richard S. Ellis, Graham P. Smith, and Jean-Paul Kneib. Separating Baryons and Dark Matter in Cluster Cores: A Full Two-dimensional Lensing and Dynamic Analysis of Abell 383 and MS 2137-23. *The Astrophysical Journal*, 674(2):711–727, February 2008. ISSN 0004-637X. doi: 10.1086/524652. URL `http://www.oamp.fr`.

[169] D.˜J. Sand, T Treu, and R.˜S. Ellis. The Matter Distribution in Galaxy Clusters. In *American Astronomical Society Meeting Abstracts*, volume 201 of *American Astronomical Society Meeting Abstracts*, page 58.07, dec 2002.

[170] Matthieu Schaller, Carlos S. Frenk, Richard G. Bower, Tom Theuns, James Trayford, Robert A. Crain, Michelle Furlong, Joop Schaye, Claudio Dalla Vecchia, and I. G. McCarthy. The effect of baryons on the inner density profiles of rich clusters. *Monthly Notices of the Royal Astronomical Society*, 452(1):343–355, sep 2015. ISSN 0035-8711. doi: 10.1093/mnras/stv1341. URL `https://academic.oup.com/mnras/article-lookup/doi/10.1093/mnras/stv1341`.

[171] Joop Schaye, Robert A Crain, Richard G Bower, Michelle Furlong, Matthieu Schaller, Tom Theuns, Claudio Dalla Vecchia, Carlos S Frenk, I.˜G. McCarthy, John C Helly,

Adrian Jenkins, Y.~M. Rosas-Guevara, Simon D.~M. White, Maarten Baes, C.~M. Booth, Peter Camps, Julio F Navarro, Yan Qu, Alireza Rahmati, Till Sawala, Peter A Thomas, and James Trayford. The EAGLE project: simulating the evolution and assembly of galaxies and their environments. *Monthly Notices of the Royal Astronomical Society*, 446(1):521–554, jan 2015. doi: 10.1093/mnras/stu2058.

[172] R. W. Schmidt and S. W. Allen. The dark matter haloes of massive, relaxed galaxy clusters observed with Chandra. *Monthly Notices of the Royal Astronomical Society*, 379(1):209–221, jul 2007. ISSN 0035-8711. doi: 10.1111/j.1365-2966.2007.11928.x. URL `https://academic.oup.com/mnras/article-lookup/doi/10.1111/j.1365-2966.2007.11928.x`.

[173] T Schramm. Realistic elliptical potential wells for gravitational lens models. *aap*, 231:19–24, may 1990.

[174] Marc Schumann. Direct detection of WIMP dark matter: Concepts and status, 2019. ISSN 13616471.

[175] M. Schwarzchild. A numerical model for a triaxial stellar system in dynamical equilibrium. II - Some dynamical features of the model. *The Astrophysical Journal Supplement Series*, 43:435, 1980. ISSN 0067-0049. doi: 10.1086/190676.

[176] Mauro Sereno, Carlo Giocoli, Stefano Ettori, and Lauro Moscardini. The mass–concentration relation in lensing clusters: the role of statistical biases and selection effects. *Monthly Notices of the Royal Astronomical Society*, 449(2):2024–2039, Mar 2015. ISSN 0035-8711. doi: 10.1093/mnras/stv416. URL `http://dx.doi.org/10.1093/mnras/stv416`.

[177] Joshua D. Simon. The Faintest Dwarf Galaxies, aug 2019. ISSN 00664146. URL `https://doi.org/10.1146/annurev-astro-091918-`.

[178] Joshua D Simon, Marla Geha, Quinn E Minor, Gregory D Martinez, Evan N Kirby, James S Bullock, Manoj Kaplinghat, Louis E Strigari, Beth Willman, Philip I Choi, Erik J Tollerud, and Joe Wolf. A COMPLETE SPECTROSCOPIC SURVEY OF THE MILKY WAY SATELLITE SEGUE 1: THE DARKEST GALAXY *. *The Astrophysical Journal*, 733:46, 2011. doi: 10.1088/0004-637X/733/1/46.

[179] Anastasia Sokolenko, Kyrylo Bondarenko, Thejs Brinckmann, Jesús Zavala, Mark Vogelsberger, Torsten Bringmann, and Alexey Boyarsky. Towards an improved model of self-interacting dark matter haloes. *Journal of Cosmology and Astroparticle Physics*, 2018(12):038–038, dec 2018. ISSN 1475-7516. doi: 10.1088/1475-7516/2018/12/038. URL `http://stacks.iop.org/1475-7516/2018/i=12/a=038?key=crossref.b535e575bf6718838f7aab8e0ddff69f`.

[180] Joshua S Speagle. dynesty: a dynamic nested sampling package for estimating Bayesian posteriors and evidences. *Monthly Notices of the Royal Astronomical Society*, 493(3):3132–3158, apr 2020. ISSN 0035-8711. doi: 10.1093/mnras/staa278.

[181] Meghin E. Spencer, Mario Mateo, Matthew G. Walker, and Edward W. Olszewski. A Multi-epoch Kinematic Study of the Remote Dwarf Spheroidal Galaxy Leo II. *The Astrophysical Journal*, 836(2):202, feb 2017. ISSN 0004-637X. doi: 10.3847/1538-4357/836/2/202. URL `https://iopscience.iop.org/article/10.3847/1538-4357/836/2/202https://iopscience.iop.org/article/10.3847/1538-4357/836/2/202/meta`.

[182] David N. Spergel and Paul J. Steinhardt. Observational evidence for self-interacting cold dark matter. *Physical Review Letters*, 84(17):3760–3, apr 2000. ISSN 10797114. doi: 10.1103/PhysRevLett.84.3760. URL `http://arxiv.org/abs/astro-ph/9909386{%}0Ahttp://dx.doi.org/10.1103/PhysRevLett.84.3760https://link.aps.org/doi/10.1103/PhysRevLett.84.3760http://arxiv.org/abs/astro-ph/9909386http://dx.doi.org/10.1103/PhysRevLett.84.3760http://www.ncbi.nlm.nih.gov/pubm`.

[183] Volker Springel, Rüdiger Pakmor, Annalisa Pillepich, Rainer Weinberger, Dylan Nelson, Lars Hernquist, Mark Vogelsberger, Shy Genel, Paul Torrey, Federico Marinacci, and Jill Naiman. First results from the IllustrisTNG simulations: matter and galaxy clustering. *Monthly Notices of the Royal Astronomical Society*, 475(1):676–698, March 2018. doi: 10.1093/mnras/stx3304.

[184] Louis E. Strigari, Carlos S. Frenk, and Simon D. M. White. Dynamical Models for the Sculptor Dwarf Spheroidal in a ΛCDM Universe. *The Astrophysical Journal*, 838(2):123, mar 2017. ISSN 1538-4357. doi: 10.3847/1538-4357/aa5c8e. URL `http://stacks.iop.org/0004-637X/838/i=2/a=123?key=crossref.91c268875e443faecc7d603d5418223c`.

[185] Elmo Tempel and Noam I Libeskind. Galaxy spin alignment in filaments and sheets: Observational evidence. *Astrophysical Journal Letters*, 775(2):42, 2013. ISSN 20418205. doi: 10.1088/2041-8205/775/2/L42.

[186] Romain Teyssier, Ben Moore, Davide Martizzi, Yohan Dubois, and Lucio Mayer. Mass distribution in galaxy clusters: The role of Active Galactic Nuclei feedback. *Monthly Notices of the Royal Astronomical Society*, 414(1):195–208, jun 2011. ISSN 00358711. doi: 10.1111/j.1365-2966.2011.18399.x.

[187] Sean Tulin and Hai Bo Yu. Dark matter self-interactions and small scale structure, 2018. ISSN 03701573.

[188] Sean Tulin, Hai Bo Yu, and Kathryn M Zurek. Beyond collisionless dark matter: Particle physics dynamics for dark matter halo structure. *Physical Review D - Particles, Fields, Gravitation and Cosmology*, 87(11), 2013. ISSN 15507998. doi: 10.1103/PhysRevD.87.115007.

[189] Sean Tulin, Hai Bo Yu, and Kathryn M. Zurek. Beyond collisionless dark matter: Particle physics dynamics for dark matter halo structure. *Physical Review D - Particles, Fields, Gravitation and Cosmology*, 87(11):115007, jun 2013. ISSN 15507998. doi: 10.1103/PhysRevD.87.115007.

[190] Keiichi Umetsu, Tom Broadhurst, Adi Zitrin, Elinor Medezinski, and Li-Yen Hsu. CLUSTER MASS PROFILES FROM A BAYESIAN ANALYSIS OF WEAK-LENSING DISTORTION AND MAGNIFICATION MEASUREMENTS: APPLICATIONS TO SUBARU DATA. *The Astrophysical Journal*, 729(2):127, mar 2011. ISSN 0004-637X. doi: 10.1088/0004-637X/729/2/127. URL `http://stacks.iop.org/0004-637X/729/i=2/a=127?key=crossref.156d4032184a5bd6d7fb4d2b20028063`.

[191] Pieter van Dokkum, Charlie Conroy, Alexa Villaume, Jean Brodie, and Aaron J. Romanowsky. The Stellar Initial Mass Function in Early-type Galaxies from Absorption Line Spectroscopy. III. Radial Gradients. *The Astrophysical Journal*, 841(2):68, may 2017. ISSN 1538-4357. doi: 10.3847/1538-4357/aa7135. URL `http://stacks.iop.org/0004-637X/841/i=2/a=68?key=crossref.74dc2647931cb69db24ffe9568d476fc`.

[192] A. Vazdekis, E. Casuso, R. F. Peletier, and J. E. Beckman. A New Chemo-evolutionary Population Synthesis Model for Early-Type Galaxies. I. Theoretical Basis. *Astrophysical Journals*, 106:307, October 1996. doi: 10.1086/192340.

[193] Jesús Vega-Ferrero, Gustavo Yepes, and Stefan Gottlöber. On the shape of dark matter haloes from MultiDark Planck simulations. *Monthly Notices of the Royal Astronomical Society*, 467(3):3226–3238, 2017. ISSN 13652966. doi: 10.1093/mnras/stx282. URL `http://www.multidark.org`.

[194] M Vogelsberger, S Genel, V Springel, P Torrey, D Sijacki, D Xu, G Snyder, S Bird, D Nelson, and L Hernquist. Properties of galaxies reproduced by a hydrodynamic simulation. *Nature*, 509(7499):177–182, may 2014. doi: 10.1038/nature13316.

[195] Mark Vogelsberger, Jesus Zavala, and Abraham Loeb. Subhaloes in self-interacting galactic dark matter haloes. *Monthly Notices of the Royal Astronomical Society*, 423(4):3740–3752, jul 2012. ISSN 00358711. doi: 10.1111/j.1365-2966.2012.21182.x. URL `https://academic.oup.com/mnras/article-lookup/doi/10.1111/j.1365-2966.2012.21182.x`.

[196] Mark Vogelsberger, Jesús Zavala, Katelin Schutz, and Tracy R. Slatyer. Evaporating the Milky Way halo and its satellites with inelastic self-interacting dark matter. *Monthly Notices of the Royal Astronomical Society*, 484(4):5437–5452, apr 2019. ISSN 13652966. doi: 10.1093/mnras/stz340.

[197] L M Voigt and S L Bridle. Limitations of model-fitting methods for lensing shear estimation. *Monthly Notices of the Royal Astronomical Society*, 404(1):458–467, 2010. ISSN 00358711. doi: 10.1111/j.1365-2966.2010.16300.x. URL `http://www.darkenergysurvey.org`.

[198] Matthew G Walker and Jorge Penarrubia. A method for measuring (slopes of) the mass profiles of dwarf spheroidal galaxies. *Astrophysical Journal*, 742(1):20, 2011. ISSN 15384357. doi: 10.1088/0004-637X/742/1/20.

[199] Matthew G Walker, Mario Mateo, Edward W Olszewski, Jorge P Narrubia, N Wyn Evans, and Gerard Gilmore. A UNIVERSAL MASS PROFILE FOR DWARF SPHEROIDAL GALAXIES? *. *The Astrophysical Journal*, 704:1274–1287, 2009. doi: 10.1088/0004-637X/704/2/1274. URL `http://iopscience.iop.org/article/10.1088/0004-637X/704/2/1274/pdf`.

[200] Matthew G. Walker, Edward W. Olszewski, and Mario Mateo. Bayesian analysis of resolved stellar spectra: Application to MMT/Hectochelle observations of the Draco dwarf spheroidal. *Monthly Notices of the Royal Astronomical Society*, 448(3):2717–2732, apr 2015. ISSN 13652966. doi: 10.1093/mnras/stv099. URL `https://academic.oup.com/mnras/article/448/3/2717/1091040`.

[201] Coral Wheeler, Andrew B. Pace, James S. Bullock, Michael Boylan-Kolchin, Jose Oñorbe, Oliver D. Elbert, Alex Fitts, Philip F. Hopkins, and Dušan Kereš. The no-spin zone: Rotation versus dispersion support in observed and simulated dwarf galaxies. *Monthly Notices of the Royal Astronomical Society*, 465(2):2420–2431, feb 2017. ISSN 13652966. doi: 10.1093/mnras/stw2583. URL `https://academic.oup.com/mnras/article/465/2/2420/2680382`.

[202] David Wittman, Nathan Golovich, and William A. Dawson. The Mismeasure of Mergers: Revised Limits on Self-interacting Dark Matter in Merging Galaxy Clusters. *Astrophys. J.*, 869(2):104, 2018. doi: 10.3847/1538-4357/aaee77.

[203] Joe Wolf, Gregory D. Martinez, James S. Bullock, Manoj Kaplinghat, Marla Geha, Ricardo R. Muñoz, Joshua D. Simon, and Frank F. Avedo. Accurate masses for dispersion-supported galaxies. *Monthly Notices of the Royal Astronomical Society*, 406(2):no–no, may 2010. ISSN 00358711. doi: 10.1111/j.1365-2966.2010.16753.x. URL `https://academic.oup.com/mnras/article-lookup/doi/10.1111/j.1365-2966.2010.16753.x`.

[204] Xiaoan Wu and Scott Tremaine. Deriving the Mass Distribution of M87 from Globular Clusters. *The Astrophysical Journal*, 643(1):210–221, may 2006. ISSN 0004-637X. doi: 10.1086/501515. URL `http://stacks.iop.org/0004-637X/643/i=1/a=210`.

[205] J. Yoo, C. S. Kochanek, E. E. Falco, and B. A. McLeod. The Lens Galaxy in PG 1115+080 is an Ellipse. *Astrophysical Journal*, 626:51–57, June 2005. doi: 10.1086/429959.

[206] J. Yoo, C. S. Kochanek, E. E. Falco, and B. A. McLeod. Halo Structures of Gravitational Lens Galaxies. *Astrophysical Journal*, 642:22–29, May 2006. doi: 10.1086/500968.

[207] J Zavala, M Vogelsberger, and M.˜G. Walker. Constraining self-interacting dark matter with the Milky way's dwarf spheroidals. *Monthly Notices of the Royal Astronomical Society*, 431:L20–L24, apr 2013. doi: 10.1093/mnrasl/sls053.

[208] Hongsheng Zhao. Analytical models for galactic nuclei. *Monthly Notices of the Royal Astronomical Society*, 278(2):488–496, 1996. ISSN 00358711. doi: 10.1093/mnras/278.2.488. URL `https://academic.oup.com/mnras/article/278/2/488/951933`.

[209] S. Zieleniewski, R. C. W. Houghton, N. Thatte, R. L. Davies, and S. P. Vaughan. Radial gradients in initial mass function sensitive absorption features in the Coma brightest cluster galaxies. *Monthly Notices of the Royal Astronomical Society*, 465: 192–212, February 2017. doi: 10.1093/mnras/stw2712.

[210] A. Zitrin, T. Broadhurst, D. Coe, K. Umetsu, M. Postman, N. Benítez, M. Meneghetti, E. Medezinski, S. Jouvel, L. Bradley, A. Koekemoer, W. Zheng, H. Ford, J. Merten, D. Kelson, O. Lahav, D. Lemze, A. Molino, M. Nonino, M. Donahue, P. Rosati, A. Van der Wel, M. Bartelmann, R. Bouwens, O. Graur, G. Graves, O. Host, L. Infante, S. Jha, Y. Jimenez-Teja, R. Lazkoz, D. Maoz, C. McCully, P. Melchior, L. A. Moustakas, S. Ogaz, B. Patel, E. Regoes, A. Riess, S. Rodney, and S. Seitz. The Cluster Lensing and Supernova Survey with Hubble (CLASH): Strong-lensing Analysis of A383 from 16-band HST/WFC3/ACS Imaging. *Astrophysical Journal*, 742(2):117, December 2011. doi: 10.1088/0004-637X/742/2/117.

[211] Adi Zitrin, Felipe Menanteau, John P. Hughes, Dan Coe, L. Felipe Barrientos, Leopoldo Infante, and Rachel Mandelbaum. A HIGHLY ELONGATED PROMINENT LENS AT <i>z</i> = 0.87: FIRST STRONG-LENSING ANALYSIS OF EL GORDO. *The Astrophysical Journal*, 770(1):L15, may 2013. ISSN 2041-8205. doi: 10.1088/2041-8205/770/1/L15. URL `http://stacks.iop.org/2041-8205/770/i=1/a=L15?key=crossref.493ea493c1d04f1a3966a55c5754dc8f`.

[212] Adi Zitrin, Agnese Fabris, Julian Merten, Peter Melchior, Massimo Meneghetti, Anton Koekemoer, Dan Coe, Matteo Maturi, Matthias Bartelmann, Marc Postman, Keiichi Umetsu, Gregor Seidel, Irene Sendra, Tom Broadhurst, Italo Balestra, Andrea Biviano, Claudio Grillo, Amata Mercurio, Mario Nonino, Piero Rosati, Larry Bradley, Mauricio Carrasco, Megan Donahue, Holland Ford, Brenda L. Frye, and John Moustakas. HUBBLE SPACE TELESCOPE COMBINED STRONG AND WEAK LENSING ANALYSIS OF THE CLASH SAMPLE: MASS AND MAGNIFICATION MODELS AND SYSTEMATIC UNCERTAINTIES. *The Astrophysical Journal*, 801(1):44, mar 2015. ISSN 1538-4357. doi: 10.1088/0004-637X/801/1/44. URL `http://stacks.iop.org/0004-637X/801/i=1/a=44?key=crossref.b1c220b5c2cae8f5bea3b87ec7883bee`.

[213] F. Zwicky. Die Rotverschiebung von extragalaktischen Nebeln. *Helvetica Physica Acta*, 6:110–127, 1933. ISSN 0018-0238. doi: 10.5169/seals-110267. URL `https://ui.adsabs.harvard.edu/abs/1933AcHPh...6..110Z/abstract`.

Appendix A

Appendices for Chapter 2

A.1 Relevant lensing formulas for cored halo models

A.1.1 cNFW halo model

The cNFW (cNFW) model is defined by modifying the NFW profile as follows:

$$\rho = \frac{\rho_s r_s^3}{\left(r_c + r\right)\left(r_s + r\right)^2}. \tag{A.1}$$

Defining $x = r/r_s$ and $\beta = r_c/r_s$, by integrating the density profile along the line of sight we find an analytic expression for the projected density profile,

$$\kappa(x) = \frac{2\kappa_s}{(\beta-1)^2}\left\{\frac{1}{x^2-1}\left[1-\beta-(1-x^2\beta)\mathcal{F}(x)\right] - \mathcal{F}\left(\frac{x}{\beta}\right)\right\} \tag{A.2}$$

, where we have defined $\kappa_s = \rho_s r_s/\Sigma_{cr}$, and

$$\mathcal{F}(x) = \begin{cases} \frac{1}{\sqrt{x^2-1}} \tan^{-1}\sqrt{x^2-1} & (x > 1) \\ \frac{1}{\sqrt{1-x^2}} \tanh^{-1}\sqrt{1-x^2} & (x < 1) \\ 1 & (x = 1) \end{cases} \tag{A.3}$$

When using the pseudo-elliptical approximation (discussed in Section 2.3.4), it is useful to have an analytic formula for the deflection angle generated by a spherical cNFW lens. By integrating Eq. A.2, we obtain

$$\alpha(x) = \frac{2\kappa_s r_s}{(1-\beta)^2 x}\Bigg\{(1-\beta)^2 \ln\left(\frac{x^2}{4}\right) - \beta^2 \ln\beta^2 + \tag{A.4}$$

$$2(\beta^2 - x^2)\mathcal{F}\left(\frac{x}{\beta}\right) + 2[1+\beta(x^2-2)]\mathcal{F}(x)\Bigg\}. \tag{A.5}$$

It can be easily verified that in the limit $\beta \to 0$, these formulae reduce to the usual analytic formulas for an NFW profile [62]. Numerical convergence of these formulae becomes difficult in the neighborhood of either $x \approx \beta$, $x \approx 1$ or $\beta \approx 1$; in each of these cases, series expansions can be used for greater accuracy, all of which have been implemented and tested in the `QLens` code.

A.1.2 Corecusp halo model

The Corecusp model is generated by including a core in the “cusped halo model” from [121], such that the density profile has the form

$$\rho = \frac{\rho_s r_s^n}{\left(r^2 + r_c^2\right)^{\gamma/2} \left(r^2 + r_s^2\right)^{(n-\gamma)/2}} \tag{A.6}$$

where r_c is the core radius and r_s acts as the scale radius where the power law “turns over”; it can also act as a tidal radius if the outer slope n is chosen to be steep enough. Choosing $n = 3, \gamma = 1$ corresponds to a cored Pseudo-NFW profile, while $n = 4, \gamma = 2$ corresponds to the dual pseudo-isothermal ellipsoid (dPIE) profile. If we allow γ to vary but set $n = \gamma$, the model reduces to the often-used softened power-law model (Barkana 1998). The advantage of this profile is that a scale radius (or tidal radius) is included, while still allowing for a variable inner slope γ and core radius r_c.

If the density profile is integrated over the line-of-sight to obtain $\kappa(R|r_s, r_c)$ where R is the projected radius, then it can be shown that this is equivalent to

$$\kappa(R|r_s, r_c) = \left(\frac{r_s}{r_s'}\right)^n \kappa_0(R'|r_s') \tag{A.7}$$

where

$$R' = \sqrt{R^2 + r_c^2}, \quad r_s' = \sqrt{r_s^2 - r_c^2}. \tag{A.8}$$

and κ_0 is defined as the coreless model, in other words, $\kappa_0(R|r_s) \equiv \kappa(R|r_s, r_c = 0)$. Thus, the cored kappa profile can be obtained from the coreless ($r_c = 0$) profile, for which the kappa and deflection formulas are known and given in Munoz et al. 2001, using the above substitutions. To simplify the notation, we will simply write $\kappa(R|r_s, r_c)$ as $\kappa(R)$, and define

$$\tilde{\kappa}(R) \equiv \kappa_0(R|r'_s = \sqrt{r_s^2 - r_c^2}), \tag{A.9}$$

so using this notation we rewrite eq. A.7 as

$$\kappa(R) = \left(1 - \frac{r_c^2}{r_s^2}\right)^{-\frac{n}{2}} \tilde{\kappa}(\sqrt{R^2 + r_c^2}) \tag{A.10}$$

The corresponding radial deflections will be referred to as $\tilde{\alpha}(R')$ and $\alpha(R)$, again using the same variable substitutions; in other words, we define

$$\tilde{\alpha}(R) \equiv \alpha_0(R|r'_s = \sqrt{r_s^2 - r_c^2}). \tag{A.11}$$

It is important to keep in mind that whenever we evaluate $\tilde{\kappa}$ and $\tilde{\alpha}$, we must make the transformation $r_s \to \sqrt{r_s^2 - r_c^2}$ in the formulas for the corresponding coreless model.

To obtain the formula for the radial deflection, we use

$$\begin{aligned}
\alpha(R) &= \frac{2}{R}\int_0^R u\kappa(u)du \\
&= \left(\frac{r_s}{r'_s}\right)^n \frac{2}{R}\int_0^R u\tilde{\kappa}(\sqrt{u^2+r_c^2})du \quad \text{(using eqs. A.7, A.8)} \\
&= \left(\frac{r_s}{r'_s}\right)^n \frac{2}{R}\int_{r_c}^{\sqrt{R^2+r_c^2}} w\tilde{\kappa}(w)dw \qquad \text{(A.12)}
\end{aligned}$$

and hence,

$$\alpha(R) = \left(1-\frac{r_c^2}{r_s^2}\right)^{-\frac{n}{2}}\left[\frac{\sqrt{R^2+r_c^2}}{R}\tilde{\alpha}(\sqrt{R^2+r_c^2}) - \frac{r_c}{R}\tilde{\alpha}(r_c)\right]. \qquad \text{(A.13)}$$

Thus we find that the radial deflection of the cored profile can be expressed as a linear combination of radial deflections from the corresponding coreless profile, again using the substitutions in eq. A.8. Thus, the same formulas for κ and α from Munoz et al. (2001) can be employed for the cored model using the above transformations.

The above transformations can be easily verified for the dPIE case ($n = 4$, $\gamma = 2$), where r_s is interpreted as a tidal radius, yielding:

$$\kappa = \frac{b}{2}\left[\frac{1}{(r_c^2 + R^2)^{1/2}} - \frac{1}{(r_s^2 + R^2)^{1/2}}\right], \tag{A.14}$$

$$\alpha = \frac{b}{R}\left[\sqrt{r_c^2 + R^2} - r_c - \sqrt{r_s^2 + R^2} + r_s\right] \tag{A.15}$$

where

$$b \equiv \frac{2\pi\rho_s}{\Sigma_{cr}}\frac{r_s^4}{r_s^2 - r_c^2}. \tag{A.16}$$

As an quick check, note that if the core radius r_c is set to zero, we can apply the transformations in eqs. A.7 and A.8 to recover the same formulas for the cored profile and its corresponding deflection angle.

A.2 Tables of Model Parameters

Table A.1: The values of the fixed parameters for the BCG and seven perturbers in the Abell 611 Lens Model. The magnitude values are in the ST magnitude system. The magnitude of the object is used to determine its core radius, cutoff radius and mass parameters via the scaling relations described in the text.

Cluster Member No.	Core Radius (arc sec)	Axis Ratio	Orientation (degrees)	x (arc sec)	y (arc sec)	m_{606w} (mag)
BCG	0.0555	0.70	132.5	0.0	0.0	17.0
1	0.0101	0.83	112.8	2.33	-7.85	20.7
2	0.0067	0.92	13.1	3.14	-10.05	21.6
3	0.0101	0.50	78.7	-5.15	17.42	20.7
4	0.0096	0.67	80.6	-10.88	10.22	20.8
5	0.0073	0.84	128.4	-16.79	0.60	21.4
6	0.0055	0.90	131.6	1.13	-2.78	22.0
7	0.0084	0.79	61.7	-13.68	12.87	21.1
ref. galaxy	0.0350	-	-	-	-	18.0

Table A.2: Median posterior parameter values for the Abell 611 models. The bounds of the 68% confidence interval are also shown. For bimodal parameters, the parameter statistics are derived from output chains that are separated into small-core ($<3''$) and large-core ($>10''$) components.

Halo profile			cNFW		Corecusp		NFW
Mode			Sm. Core	Lg. Core	Sm. Core	Lg. Core	(no core)
Bimodal Parameters	Sampled or derived	Units					
M_{200} (cNFW, NFW)	sampled	$10^{14}M_\odot$	$12.7^{+1.4}_{-1.1}$	$8.4^{+1.7}_{-0.9}$			$12.8^{+1.6}_{-1.3}$
M_{200} (Corecusp)	derived	$10^{14}M_\odot$			$7.2^{+0.7}_{-0.6}$	$5.4^{+0.8}_{-0.6}$	
c	sampled	-	$4.15^{+0.35}_{-0.35}$	$7.91^{+3.55}_{-2.16}$			$4.13^{+0.39}_{-0.41}$
κ_0	sampled	-			$1.04^{+0.08}_{-0.09}$	$1.68^{+0.38}_{-0.31}$	
γ	sampled	-			$1.04^{+0.05}_{-0.03}$	$0.71^{+0.16}_{-0.19}$	
Corecusp halo scale	sampled	$''$			$62.9^{+6.5}_{-5.7}$	$40.6^{+8.5}_{-6.5}$	
log(s/arcsec) (Corecusp core scale)	sampled	-			$-2.63^{+1.33}_{-0.96}$	$-1.60^{+1.62}_{-1.63}$	
$log(r_c/kpc)$ (cNFW core scale)	sampled	-	$-2.12^{+1.36}_{-1.29}$	$1.61^{+0.34}_{-0.37}$			
BCG mass scale	sampled	$''$	$3.12^{+0.74}_{-0.64}$	$4.63^{+0.75}_{-0.65}$	$2.99^{+0.64}_{-0.60}$	$4.35^{+0.77}_{-0.68}$	$3.13^{+0.81}_{-0.69}$
Perturber 1 mass scale	sampled	$''$	$0.94^{+2.85}_{-0.20}$	$0.92^{+2.97}_{-0.20}$	$0.83^{+3.17}_{-0.12}$	$0.87^{+2.73}_{-0.15}$	$0.94^{+2.41}_{-0.20}$
Perturber 1 r_{eff}	sampled	$''$	$2.12^{+4.84}_{-1.85}$	$2.16^{+4.75}_{-1.90}$	$3.12^{+4.35}_{-2.86}$	$2.69^{+4.44}_{-2.41}$	$2.24^{+4.88}_{-1.93}$
Perturber 2 mass scale	sampled	$''$	$0.39^{+0.21}_{-0.17}$	$0.36^{+0.23}_{-0.17}$	$0.36^{+0.17}_{-0.16}$	$0.32^{+0.20}_{-0.16}$	$0.42^{+0.31}_{-0.20}$
Perturber 2 r_{eff}	sampled	$''$	$4.81^{+3.45}_{-3.18}$	$4.41^{+3.62}_{-3.10}$	$4.76^{+3.36}_{-3.11}$	$4.48^{+3.60}_{-3.11}$	$4.14^{+3.95}_{-3.29}$
Perturber 3+ mass scale	sampled	$''$	$0.31^{+0.05}_{-0.05}$	$0.32^{+0.09}_{-0.06}$	$0.32^{+0.05}_{-0.05}$	$0.33^{+0.09}_{-0.06}$	$0.31^{+0.06}_{-0.05}$
Perturber 3+ r_{eff}	sampled	$''$	$5.09^{+2.39}_{-1.85}$	$3.57^{+2.48}_{-1.75}$	$4.49^{+1.91}_{-1.51}$	$3.12^{+2.17}_{-1.51}$	$5.00^{+2.55}_{-1.98}$
$\hat{c}_{200}$	derived	-	$4.1^{+0.3}_{-0.4}$	$6.1^{+0.6}_{-0.9}$	$6.3^{+0.4}_{-0.4}$	$7.6^{+0.6}_{-0.6}$	$4.1^{+0.4}_{-0.4}$
r_{core}	derived	$''$	$0.3^{+1.2}_{-0.2}$	$13.3^{+1.3}_{-2.0}$	$0.01^{+0.6}_{-0.01}$	$15.6^{+1.5}_{-2.6}$	(0 by def.)
M_{BCG}	derived	$10^{12}M_\odot$	$4.9^{+1.2}_{-1.0}$	$7.2^{+1.2}_{-1.0}$	$4.7^{+1.0}_{-0.9}$	$6.8^{+1.2}_{-1.0}$	$4.9^{+1.3}_{-1.1}$
$\kappa_{tot}(5'')$	derived	-	$1.32^{0.01}_{-0.01}$	$1.30^{+0.01}_{-0.01}$	$1.32^{+0.01}_{-0.01}$	$1.29^{+0.01}_{-0.01}$	$1.32^{+0.01}_{-0.01}$
$\kappa_{DM}(5'')$	derived	-	$1.12^{+0.04}_{-0.04}$	$1.02^{+0.04}_{-0.04}$	$1.13^{+0.03}_{-0.03}$	$1.03^{+0.04}_{-0.04}$	$1.12^{+0.04}_{-0.04}$
$\kappa_{DM}(20'')$	derived	-	$0.60^{+0.01}_{-0.01}$	$0.59^{+0.01}_{-0.01}$	$0.60^{+0.01}_{-0.01}$	$0.59^{+0.01}_{-0.01}$	$0.60^{+0.01}_{-0.01}$
Unimodal Parameters	Sampled or derived	Units					
position angle	sampled	degrees	$133.3^{+0.2}_{-0.2}$		$133.3^{+0.3}_{-0.2}$		$133.3^{+0.3}_{-0.2}$
axis ratio	sampled	-	$0.67^{+0.01}_{-0.01}$		$0.67^{+0.01}_{-0.01}$		$0.67^{+0.01}_{-0.01}$
x-center	sampled	$''$	$-0.2^{+0.3}_{-0.3}$		$-0.3^{+0.3}_{-0.3}$		$-0.2^{+0.2}_{-0.2}$
y-center	sampled	$''$	$0.7^{+0.3}_{-0.3}$		$1.0^{+0.3}_{-0.3}$		$0.7^{+0.3}_{-0.2}$

A.3 Posterior Distributions for Mock Data and Abell 611

We show the 1-d and 2-d posteriors for the mock data inferences (both cored and cuspy mocks) and Abell 611 data (main halo modeled with cNFW profile and corecusp profile).

Figure A.1: Posterior distributions and two-dimensional correlations using mock data for the cored case without central images. True parameter values are indicated in orange.

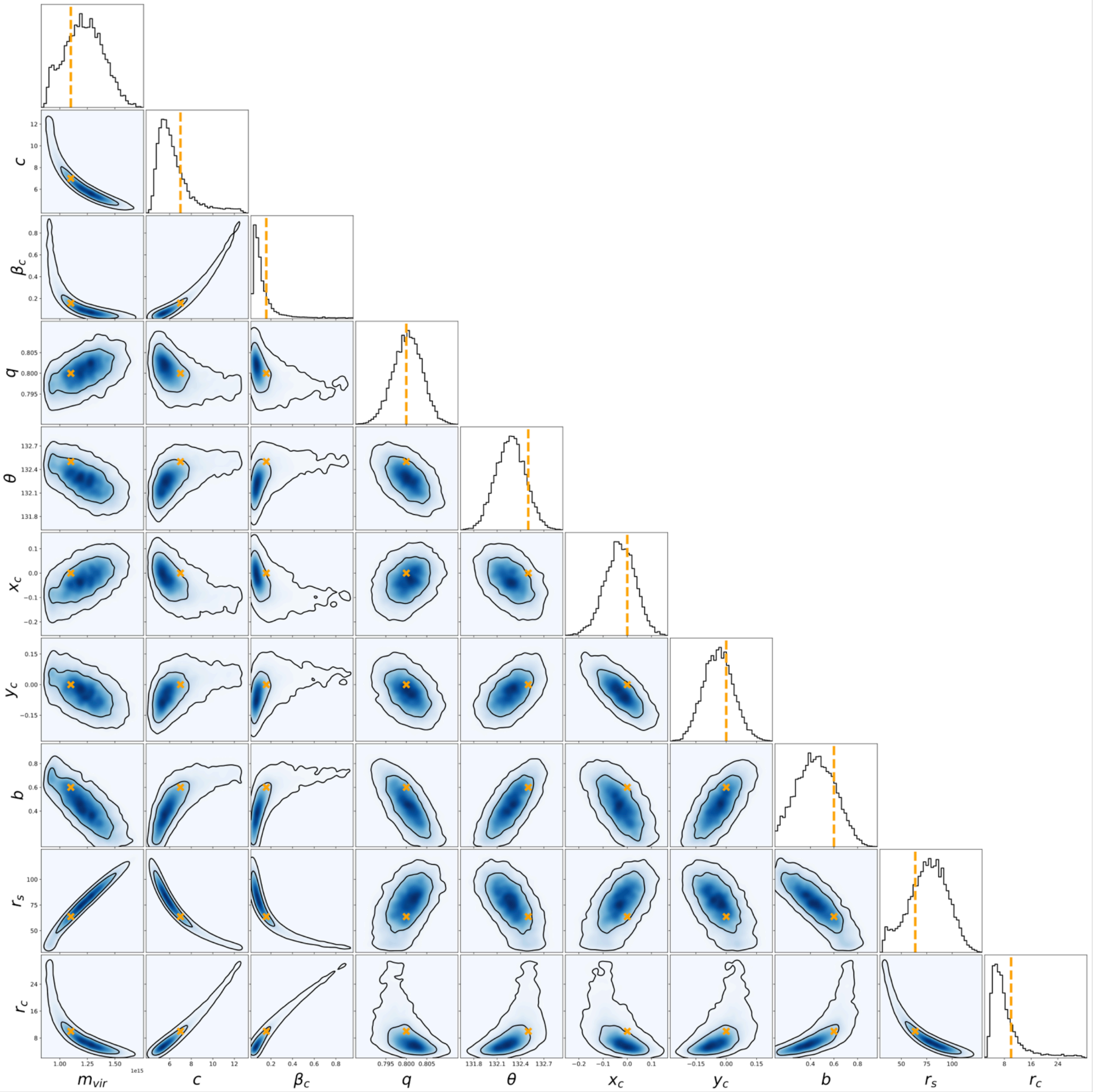


Figure A.2: Posterior distributions and two-dimensional correlations using mock data for the cuspy case without central images. True parameter values are indicated in orange.

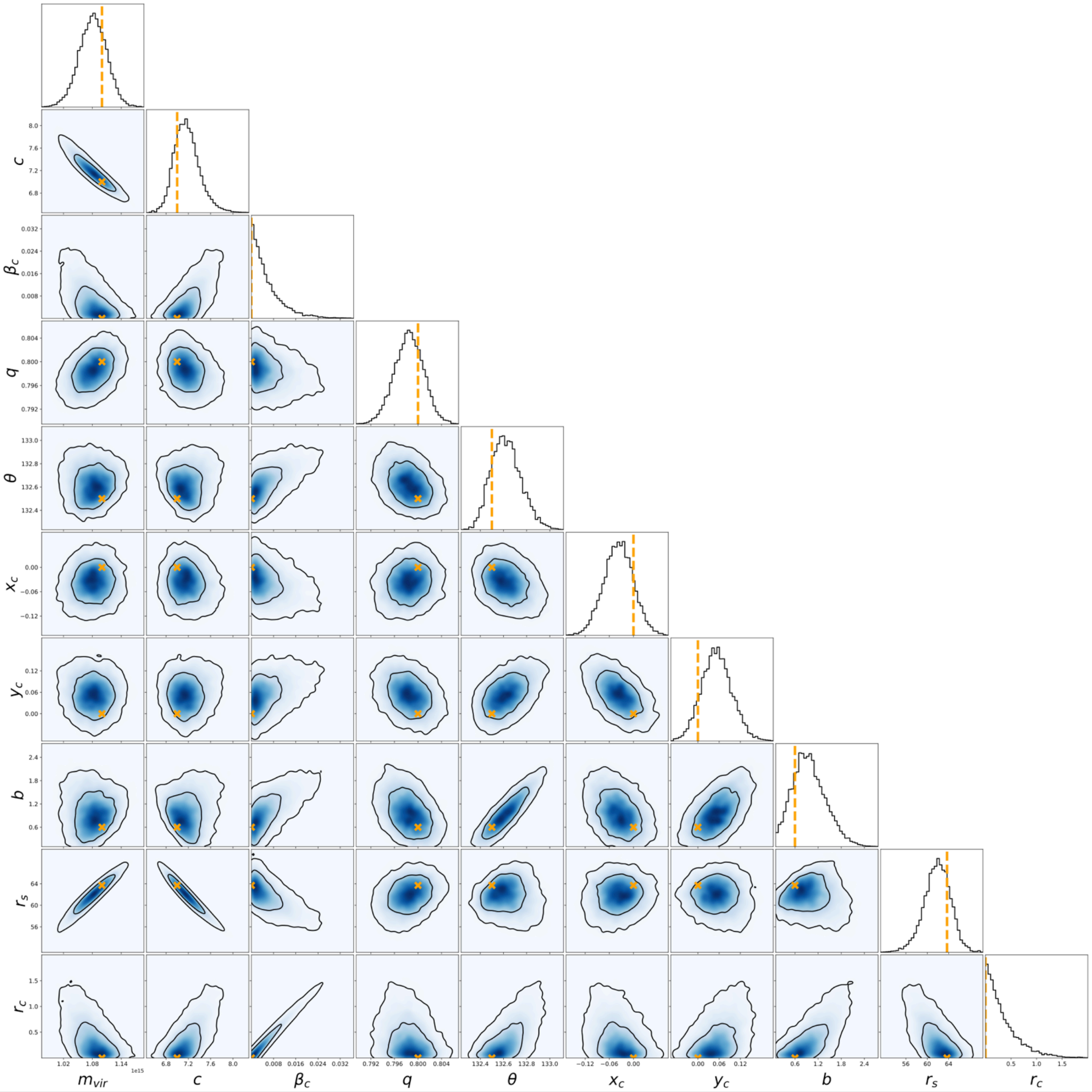
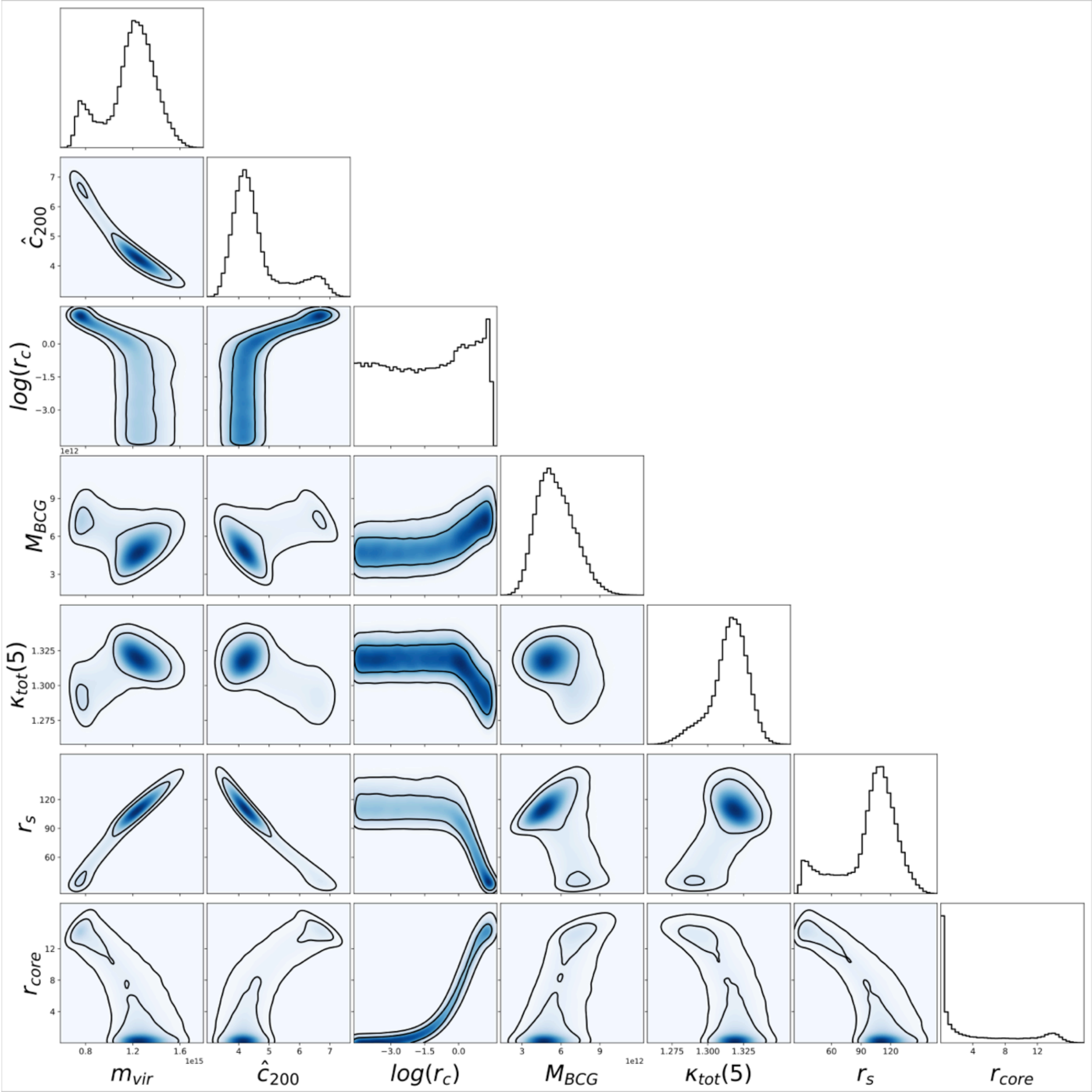


Figure A.3: Selected Abell 611 posterior distributions and two-dimensional correlations for the cNFW density profile.



‘’

Figure A.4: Selected Abell 611 posterior distributions and two-dimensional correlations for the Corecusp density profile.

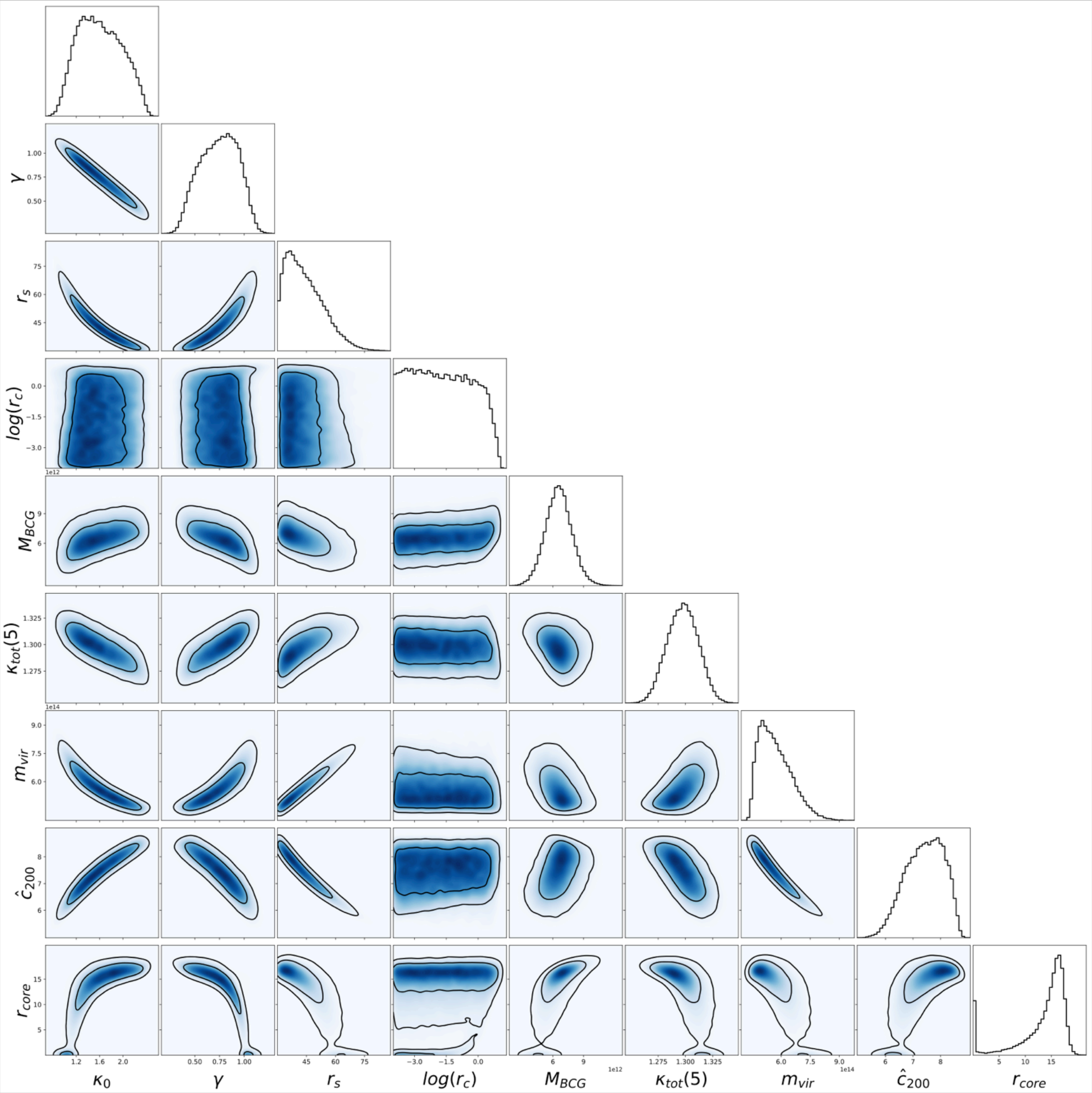
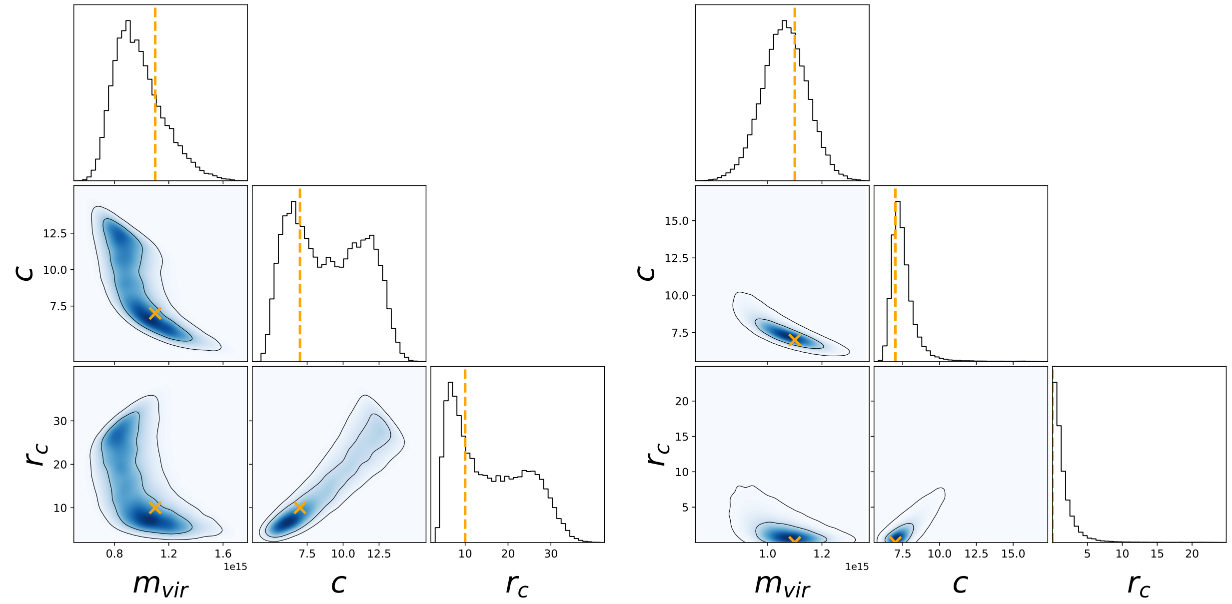


Figure A.5: Selected posterior distributions for a cNFW model of the mock data set with perturbers. This data set has only two unique redshifts and contains two perturbing galaxies. [*Left:*] cored case. [*Right:*] cuspy case. Orange lines and 'x' markers indicate the true parameter values. The units for m_{vir} are $10^{15} M_{\odot}$, and for core radius, arcseconds.



Appendix B

Appendix for Chapter 3

B.1 Relevant lensing formulas for the cNFW halo model

The cNFW (cNFW) model is defined by modifying the NFW profile as follows:

$$\rho = \frac{\rho_s r_s^3}{\left(r_c + r\right)\left(r_s + r\right)^2}. \tag{B.1}$$

Defining $x = r/r_s$ and $\beta = r_c/r_s$, by integrating the density profile along the line of sight we find an analytic expression for the projected density profile,

$$\kappa(x) = \frac{2\kappa_s}{(\beta - 1)^2}\left\{\frac{1}{x^2 - 1}\left[1 - \beta - (1 - x^2\beta)\mathcal{F}(x)\right] - \mathcal{F}\left(\frac{x}{\beta}\right)\right\} \tag{B.2}$$

, where we have defined $\kappa_s = \rho_s r_s/\Sigma_{cr}$, and

$$\mathcal{F}(x) = \begin{cases} \frac{1}{\sqrt{x^2-1}}\tan^{-1}\sqrt{x^2-1} & (x > 1) \\ \frac{1}{\sqrt{1-x^2}}\tanh^{-1}\sqrt{1-x^2} & (x < 1) \\ 1 & (x = 1) \end{cases} \tag{B.3}$$

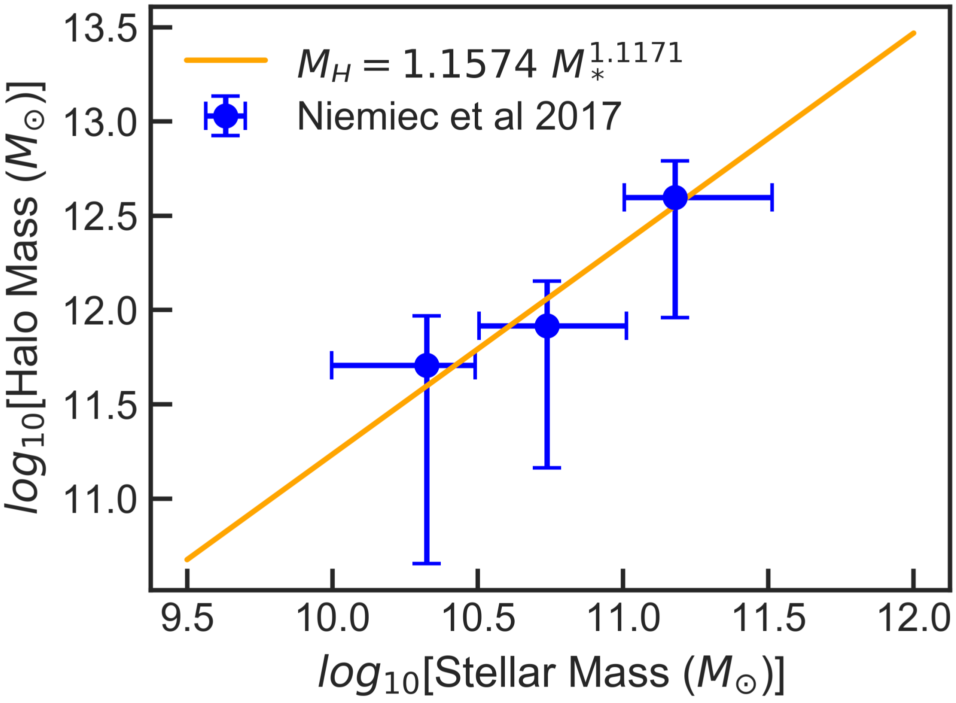
When using the pseudo-elliptical approximation, it is useful to have an analytic formula for the deflection angle generated by a spherical cNFW lens. By integrating Equation B.2, we obtain

$$\alpha(x) = \frac{2\kappa_s r_s}{(1-\beta)^2 x}\Bigg\{(1-\beta)^2 \ln\left(\frac{x^2}{4}\right) - \beta^2 \ln\beta^2 \quad + \tag{B.4}$$

$$2(\beta^2 - x^2)\mathcal{F}\left(\frac{x}{\beta}\right) + 2[1+\beta(x^2-2)]\mathcal{F}(x)\Bigg\}. \tag{B.5}$$

It can be easily verified that in the limit $\beta \to 0$, these formulae reduce to the usual analytic formulas for an NFW profile [63]. Numerical convergence of these formulae becomes difficult in the neighborhood of either $x \approx \beta$, $x \approx 1$ or $\beta \approx 1$; in each of these cases, series expansions can be used for greater accuracy, all of which have been implemented and tested in the `QLens` code.

Figure B.1: Stellar mass-halo mass data from Niemiec et al. [134], and the best fit power law which was used to determine the mass for DM subhalos in the lens models.



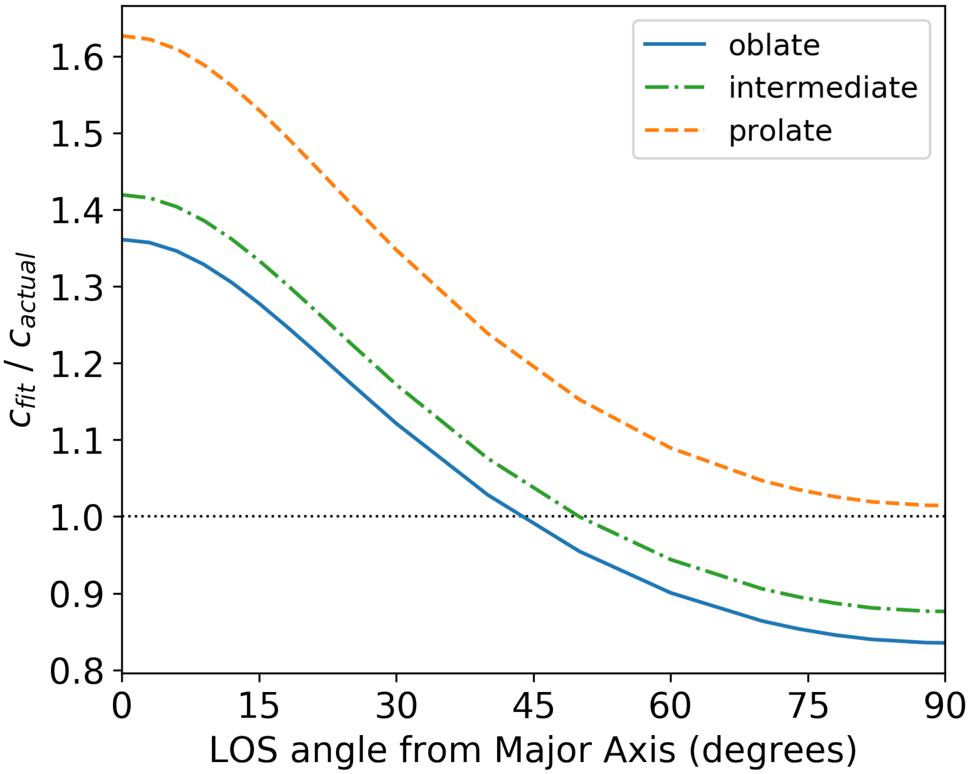
B.2 Stellar Mass–Halo Mass Relation

Niemiec et al. [134] examined the stellar mass-halo mass relation for cluster galaxies. A power law fit to their data is shown in Figure B.1. The resulting power law is $M_{\rm halo} = 1.1574 M_*^{1.1171}$.

B.3 Concentration Bias in Triaxial Halos

Since strong lensing analysis is based on 2-D projections of 3-D bodies, effectively hiding one dimension from direct measurement, biases can arise in model parameters such as concentration and scale radius. Halo elongation parallel (perpendicular) to the line of sight can increase (decrease) the strength of the lensing. Moreover, alignment of the major axis with

Figure B.2: Bias in concentration when fitting 2-D projections of triaxial halos. The bias is $\sim +60\%$ for prolate halos oriented along the LOS.



the line of sight has shown to increase as lensing cross section increases, and leads to an upward bias in concentration [111].

Galaxies and galaxy clusters are thought to form along filaments of the cosmic web. The major axis of elliptical galaxies tend to be parallel to the direction of the host filament [185]. In cosmological N-body simulations, prolate halos are more common than oblate ones, especially so for massive halos [13]. We therefore expect the preponderance of our sample to be prolate halos with their major axis oriented preferentially along the line of sight, creating significant upward bias in concentration. To quantify the effect of concentration bias in our modeling pipeline, we created simulated triaxial halos, projected them in various orientations, and fitted the resulting 2D shapes with our fitting code. The simulated halos were NFW ellipsoids, with triaxiality introduced by scaling the x, y and z axes by factors a, b and c. The triaxiality is thus defined by the ellipsoid with form

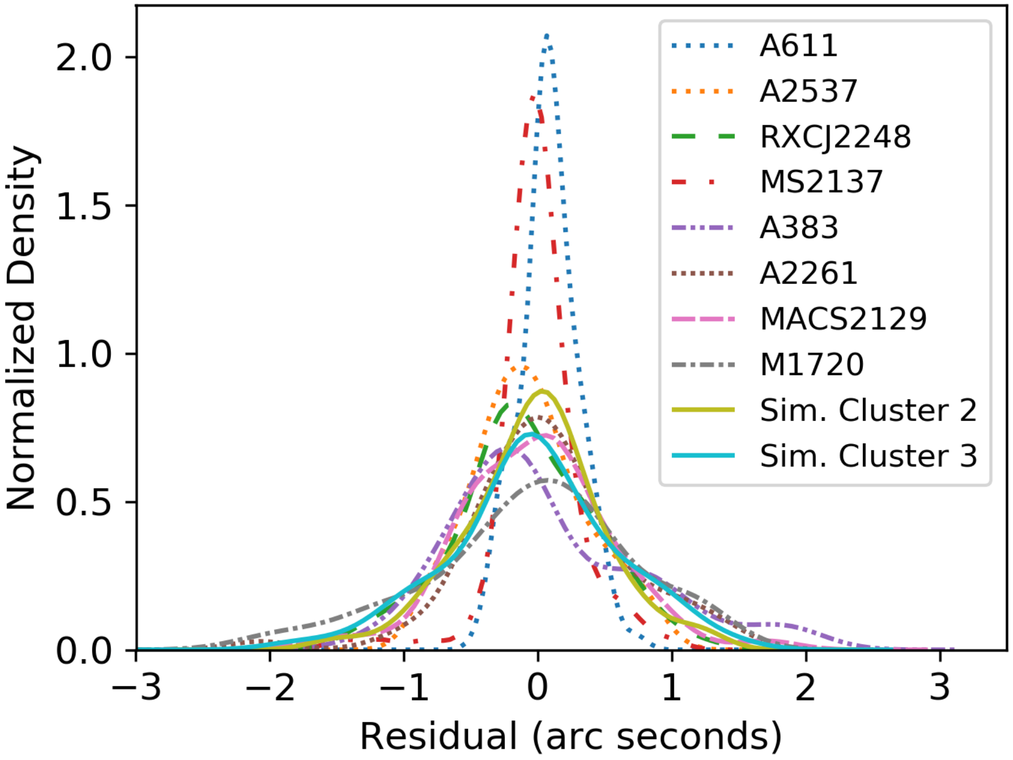
$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1. \tag{B.6}$$

For the oblate case, we chose $a = 0.5,\ b = c = 1.0$. For the intermediate case, $a = 0.5,\ b = 0.75,\ c = 1.0$. The prolate case used $a = b = 0.5,\ c = 1.0$. Thus the z-axis is the major axis in all cases. Figure B.2 shows the results of that analysis. Biases in concentration range from approximately +60% to -20%. The strongest bias of $\sim$ +60% occurs in the common case of prolate halos with major axis orientation nearly along the line of sight. We therefore expect that our modeled concentrations are often biased high by as much as 60%.

B.4 Strong Lensing Error Residuals

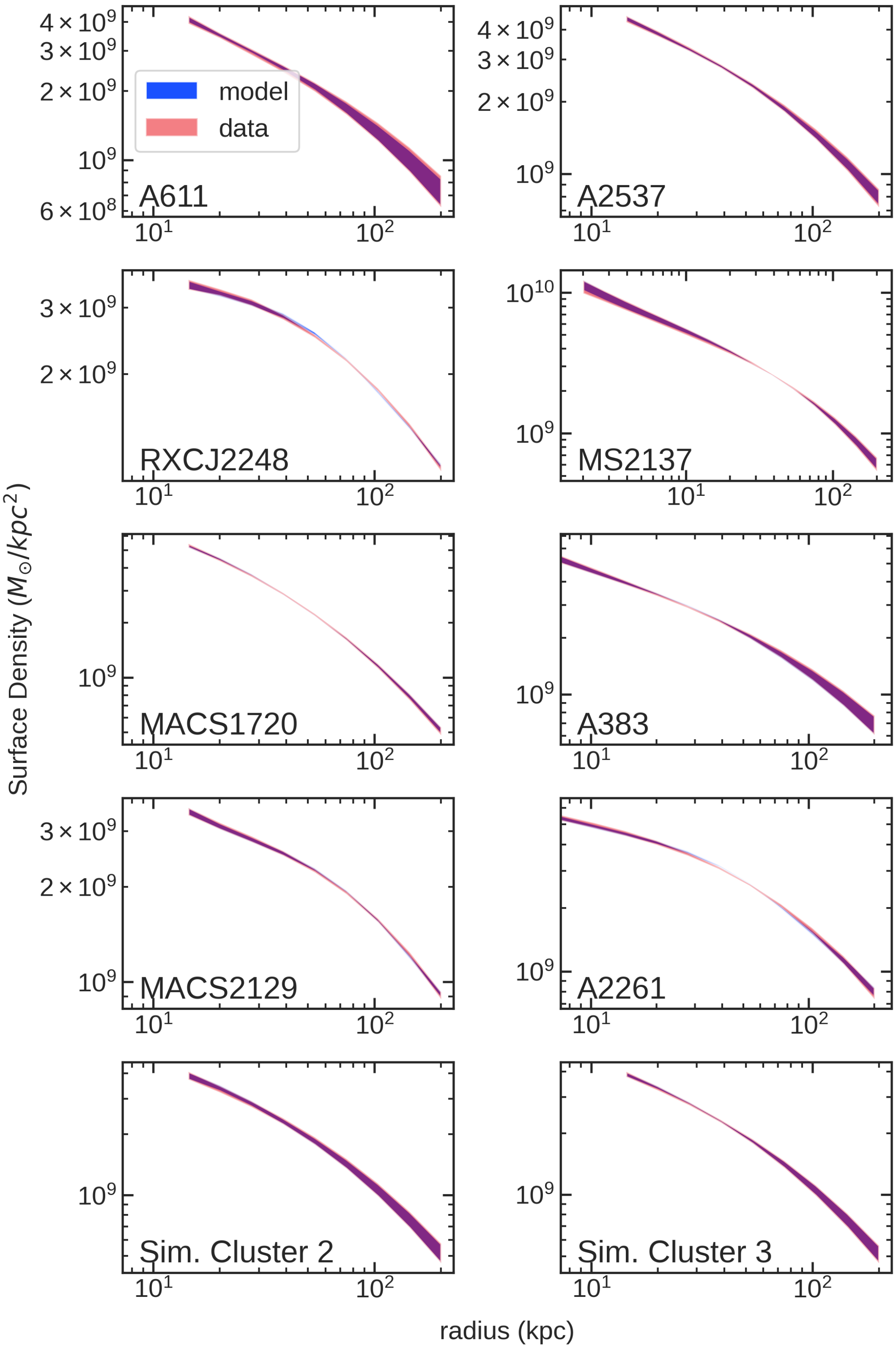
In the Figure B.3 we show the distribution of residual errors in image locations from the strong lensing analysis. For the simulated clusters, random Gaussian error of $0\farcs50$ was added to each of the x and y components, resulting in a total position error of $0\farcs71$. In the strong lensing χ^2, we have again assumed image position errors of $0\farcs50$ for each of the x and y components. The figure demonstrates that the image position errors are distributed in a Gaussian shape, as would be expected for measurement errors.

Figure B.3: Residual position errors in image locations from strong lensing. Both the x and y error components are included; positive residuals mean errors in the westerly direction for x-components and the northerly direction for y-components.



B.5 Surface Density Match between Strong Lensing and SIDM Models

Figure B.4: Plots of the 68% confidence intervals for strong lensing ("data") and SIDM ("model") surface densities for each cluster. The model is shown in blue and the data in red, and the regions appear purple where they overlap.



Appendix C

Appendices for Chapter 4

C.1 Virial Shape Parameter

The virial shape parameter is derived from the fourth-order projected virial theorem [112], and for approximately spherical systems it can take two forms [157]. Following Kaplinghat et al. [81], we utilize the first form, which we label here as the "vsp":

$$\text{vsp} = \frac{1}{2}\int_0^\infty dR^2 \Sigma \langle v_{\text{los}}^4 \rangle = \frac{G}{5}\int_0^\infty dr^2 M(5-2\beta)\nu\sigma_r^2, \tag{C.1}$$

where M denotes the mass distribution function, β is the anisotropy parameter, ν is the stellar density and $\langle v_{\text{los}}^4 \rangle$ is the fourth moment of the line-of-sight velocity distribution.

To calculate the vsp from the DF, we integrate the fourth velocity moment as follows:

$$(2\pi)\text{vsp} = \int d^3r d^3v \; v_z^4 f(E, L).$$

Note that there is no factor of N_* because the DF is normalized to unity over the entire phase space. Now write

$$v_r = v\cos\eta, v_\theta = v\sin\eta\cos\psi.$$

Then

$$v_z = \hat{z}\cdot\vec{v} = v_r\cos\theta - v_\theta\sin\theta = v(\cos\theta\cos\eta - \sin\theta\sin\eta\cos\psi)$$

and

$$(2\pi)\mathrm{vsp} = (2\pi)\int r^2dr\ d\mathrm{cos}\theta\ v^2dv\ d\psi d\mathrm{cos}\eta\ v_z^4 f(E,L)$$

But $E = \Psi - v^2/2$ and $L = rv_t = rv\sin\eta$. So we can first do the θ and ψ integrals over v_z^4.

It can be shown that

$$\int_0^{2\pi} d\psi \int_{-1}^{1} d\mathrm{cos}\theta(\cos\theta\cos\eta - \sin\theta\sin\eta\cos\psi)^4 = 4\pi/5,$$

hence,

$$\mathrm{vsp} = (4\pi/5)\int_0^{r_t} r^2dr\ \int_0^{\sqrt{2\Psi(r)}} v^6dv\ \int_{-1}^{1} d\mathrm{cos}\eta\ f(E,L) \tag{C.2}$$

For data sets with measured line-of-sight velocities, the vsp can be calculated as follows. In our coordinate system, the z-axis is the line-of-sight. First, the mean value of v_z is subtracted from each $v_{z,i}$ to remove bulk motion of the galaxy. The vsp is then

$$\mathrm{vsp} = \frac{1}{2\pi N_*}\sum_{i=1}^{N_*} v_{z,i}^4 \tag{C.3}$$

For the mock data sets, we wish to find the an estimate of the distribution of the vsp given the one set of sampled velocities. We do so by generating 10,000 ensembles of binned velocity data, each with length N_*, from a Pearson distribution of Type VII, with the same star count and velocity dispersion in each bin as the original data set. To simulate measurement uncertainty, we add Gaussian error with a standard deviation of 2 km/s. The kurtosis of the Pearson distribution is adjustable via a parameter, and that parameter is iteratively varied until the kurtosis of the entire ensemble matches that of the original data set. We then tabulate the 15.9, 50 and 84.1 percentile values of the vsp of the entire ensemble, which are used as estimators for the mean and standard deviation of the vsp. Those values are used as data for the DF model and are tabulated for the mock data sets in Table C.1.

C.2 Mock Data Characteristics

Table C.1: Mock data characteristics. The columns, from left, are (1) ID number, (2) ID string, (3) the number of stars in the data set, (4) DM profile type, (5) log slope of the inner stellar profile, (4) scale radius of the stellar profile, (6) anisotropy radius of the stellar profile, (7) 3D half-light radius, and (8) virial shape parameter of the data set.

Number	ID	w	DM Profile	γ_*	r_*	r_a	$r_{1/2}$	VSP (10^3 km^4/sec^4)
1	aaaO_4639	4639	cored	0.1	0.1	1	0.122	$4.25^{+1.59}_{-0.75}$
2	aabO_4941	4941	cored	0.1	0.1	10000	0.119	$3.08^{+0.62}_{-0.43}$
3	acaO_3904	3904	cored	0.1	0.5	1	0.596	$94.09^{+10.67}_{-7.99}$
4	acbO_2607	2607	cored	0.1	0.5	10000	0.588	$84.75^{+12.34}_{-9.30}$
5	adaO_1980	1980	cored	0.1	1.0	1	1.233	$231.66^{+25.55}_{-22.09}$
6	adbO_1441	1441	cored	0.1	1.0	10000	1.251	$172.54^{+15.29}_{-14.22}$
7	baaO_1826	1826	cored	1.0	0.1	1	0.093	$2.89^{+2.26}_{-0.80}$
8	babO_2156	2156	cored	1.0	0.1	10000	0.090	$1.80^{+0.72}_{-0.40}$
9	bcaO_2107	2107	cored	1.0	0.5	1	0.463	$70.06^{+12.83}_{-9.49}$
10	bcbO_2349	2349	cored	1.0	0.5	10000	0.464	$50.70^{+8.25}_{-5.63}$
11	bdaO_2677	2677	cored	1.0	1.0	1	0.913	$164.09^{+17.59}_{-14.81}$
12	bdbO_2456	2456	cored	1.0	1.0	10000	0.914	$113.42^{+8.75}_{-8.47}$
13	aaaN_2358	2358	NFW	0.1	0.1	1	0.121	$3.73^{+0.45}_{-0.39}$
14	aabN_3539	3539	NFW	0.1	0.1	10000	0.122	$3.04^{+0.23}_{-0.21}$
15	acaN_1088	1088	NFW	0.1	0.5	1	0.600	$20.78^{+2.64}_{-2.22}$
16	acbN_550	550	NFW	0.1	0.5	10000	0.603	$15.12^{+2.46}_{-2.17}$
17	adaN_1860	1860	NFW	0.1	1.0	1	1.238	$30.86^{+3.40}_{-2.96}$
18	adbN_826	826	NFW	0.1	1.0	10000	1.226	$23.92^{+3.38}_{-2.76}$
19	baaN_1533	1533	NFW	1.0	0.1	1	0.096	$2.96^{+0.49}_{-0.36}$
20	babN_1491	1491	NFW	1.0	0.1	10000	0.092	$2.43^{+0.37}_{-0.28}$
21	bcaN_2054	2054	NFW	1.0	0.5	1	0.453	$16.46^{+1.61}_{-1.40}$
22	bcbN_1222	1222	NFW	1.0	0.5	10000	0.434	$14.89^{+1.57}_{-1.38}$
23	bdaN_2912	2912	NFW	1.0	1.0	1	0.953	$25.33^{+2.03}_{-1.80}$
24	bdbN_1524	1524	NFW	1.0	1.0	10000	0.925	$24.24^{+2.18}_{-2.07}$

C.3 Full Likelihood Function

Consider a population of w stars in a potential Φ and with a distribution function f. Our goal is to estimate Φ and f based on the star population. For star i, we have position coordinates $R_i = \sqrt{x_i^2 + y_i^2}$, and we have velocity coordinate $v_{z,i}$ (but we do not generally know $z_i, v_{x,i}$ or $v_{y,i}$). The best estimate of Φ and f is the one that maximizes the likelihood function

$$LH(\Phi, f|[R_i, v_{z,i}]) \tag{C.4}$$

By Bayes Theorem, we instead estimate the posterior and prior probabilities

$$LH(\Phi, f|[R_i, v_{z,i}]) = LH([R_i, v_{z,i}]|\Phi, f)\frac{P(\Phi, f)}{P([R_i, v_{z,i}])} \tag{C.5}$$

where $LH([R_i, v_{z,i}]|\Phi, f)$ is the posterior probability of observing the given data with a particular Φ and f, and $P(\Phi, f)$ is the prior probability for observing Φ and f, and incorporates any prior beliefs. The probability of observing the data for our model, $P([R_i, v_{z,i}])$, also known as the "evidence", is not generally known, but as it is a constant factor it will not affect our attempts to maximize the likelihood function.

We wish to employ the distribution function as a probability of finding a star i at radius R and line-of-sight velocity v_z. The probability can be written as

$$p_{*,i}(R_i, v_{z,i}) = 2\pi \int_{-\infty}^{\infty} v_R dv_R \int_{-\infty}^{\infty} dz\ f(E_i, L_i). \tag{C.6}$$

The composite likelihood for all stars in the data set is then

$$LH([R, v_z]|\Phi, f) = \prod_{i=1}^{w} p_{*,i}(R_i, v_{z,i}) \tag{C.7}$$

and the log likelihood is then

$$LLH \triangleq log(LH([R, v_z]|\Phi, f))) = \sum_{i=1}^{w} log(p_{*,i}(R_i, v_{z,i})). \tag{C.8}$$

Computationally, we have a vector of parameters $\mathbf{p} = \{r_s, v_s, \Phi_{\text{lim}}, e, a, q, E_c, d, L_\beta, b_{\text{in}}, b_{\text{out}}, w\}$ for which we want to calculate a given likelihood. The normalization factor n_f may be factored out of the sum, and LLH becomes

$$\begin{aligned} LLH(\mathbf{p}) =& n_f(\mathbf{p}) \sum_{i=1}^{w} \log \Big(\int_{-\infty}^{\infty} dv_R \int_{-\infty}^{\infty} dz \ h\big(\mathbf{p}, E(\mathbf{x_i}, \mathbf{v_i})\big) \\ & \cdot g\big(\mathbf{p}, L(\mathbf{x_i}, \mathbf{v_i})\big)\Big), \end{aligned} \tag{C.9}$$

where the functions $h()$ and $g()$ are given in Equations 4.13 and 4.14.

C.4 Core Parameter, Core Size, and Anisotropy Parameter Posteriors of the Observed Sample

Figure C.1: Posteriors of the core parameter, $c = r_c/r_s$ for the bright MW dSphs. Only Draco and Carina show a significant tail above $c = 1$.

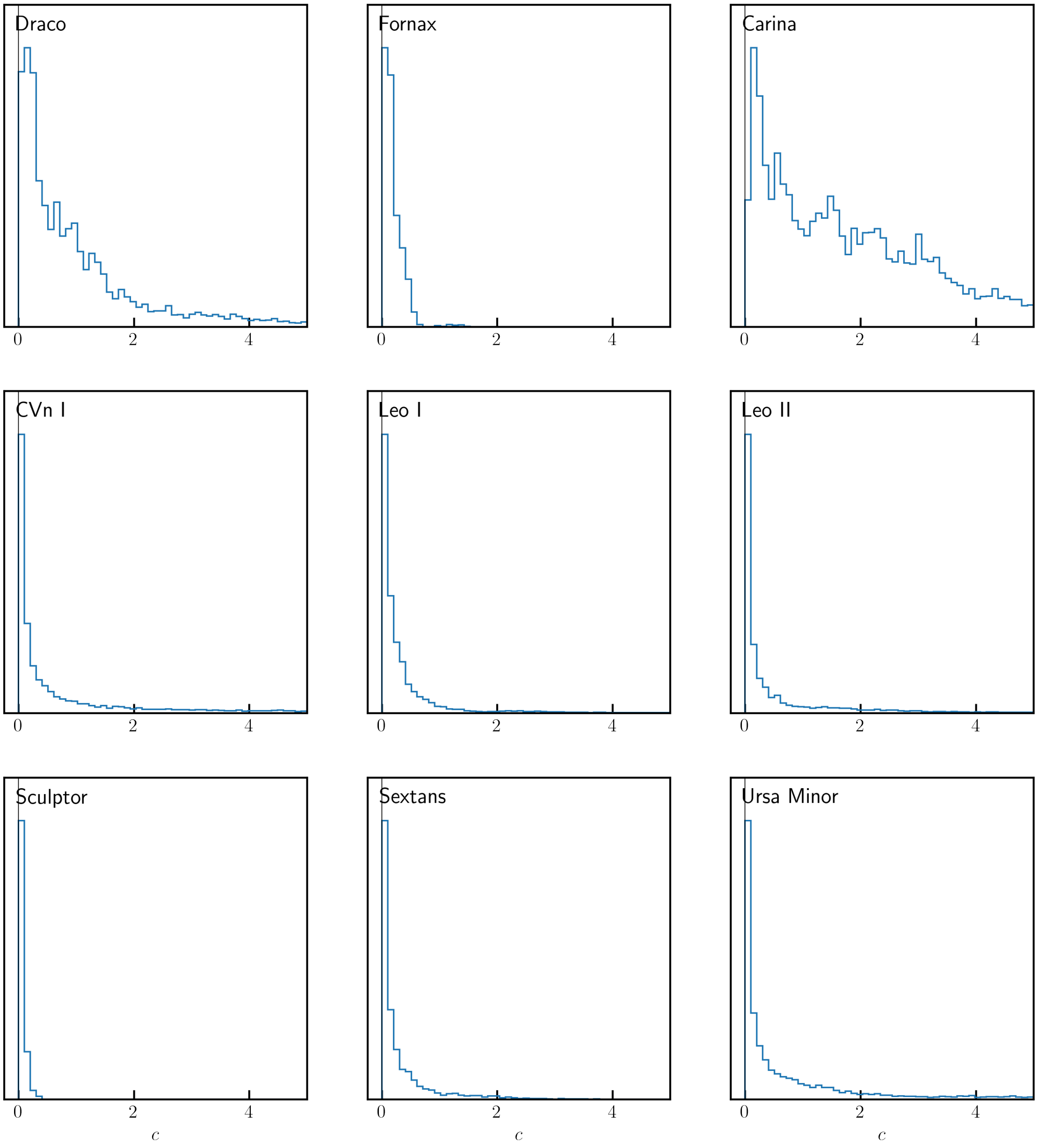


Figure C.2: Posteriors of r_{core} (kpc) for the bright MW dSphs.

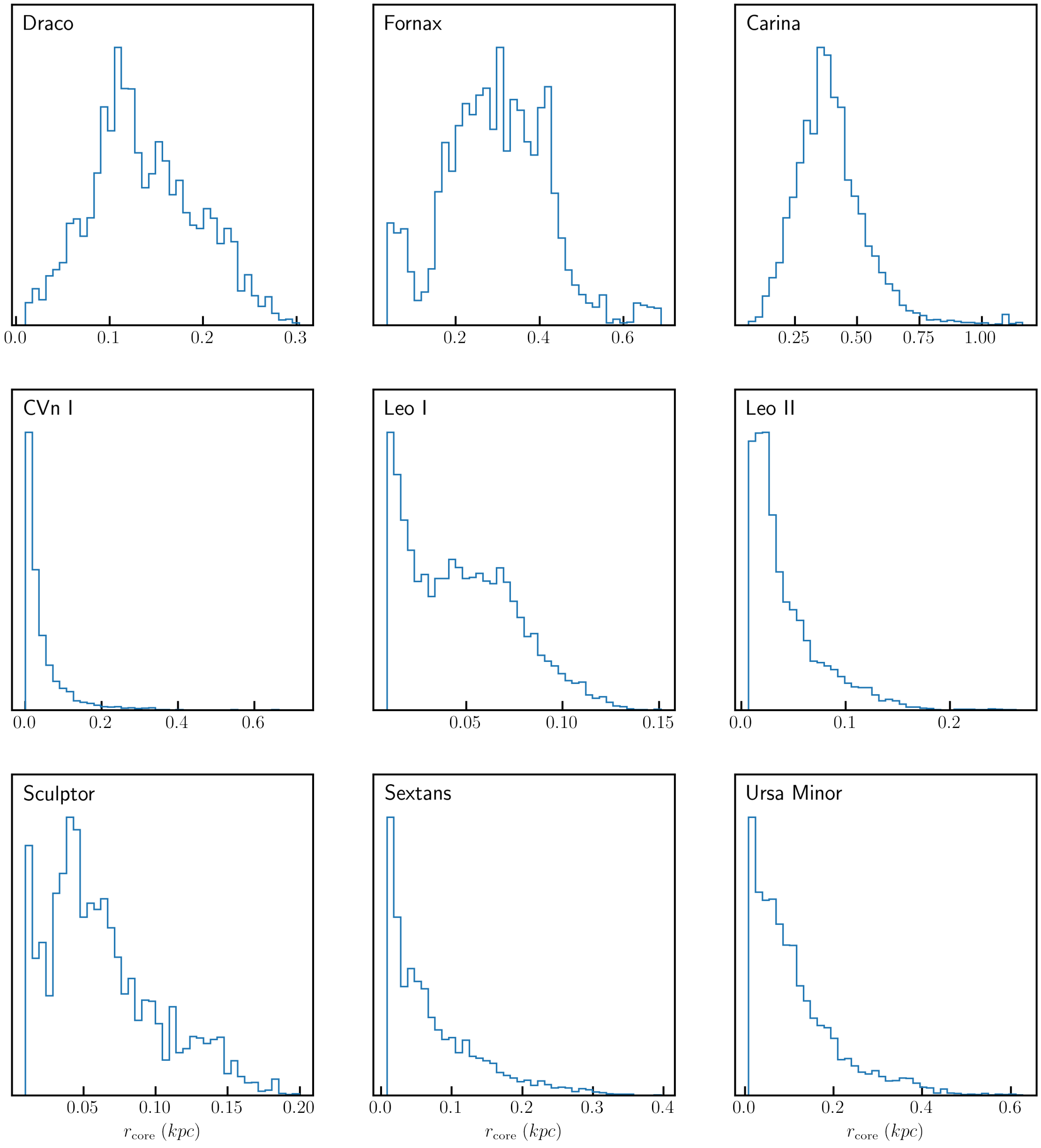


Figure C.3: Posteriors of anisotropy parameter β at the half-light radius for the bright MW dSphs.