

Understanding dark matter halos with tidal
caustics

by

Robyn Ellyn Sanderson

B.S., University of Maryland (2003)

Submitted to the Department of Physics
in partial fulfillment of the requirements for the degree of

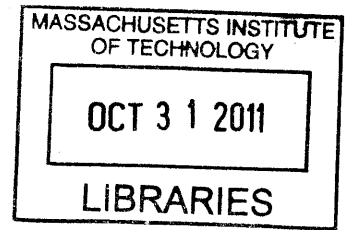
Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 2011


© Massachusetts Institute of Technology 2011. All rights reserved.



ARCHIVES

Author 

Department of Physics

January 25, 2011 

Certified by

.....

Edmund Bertschinger


Professor of Physics

Chair, Department of Physics

Thesis Supervisor

Accepted by

.....

 Krishna Rajagopal

Professor of Physics

Associate Head for Education, Department of Physics

Understanding dark matter halos with tidal caustics

by

Robyn Ellyn Sanderson

Submitted to the Department of Physics
on January 25, 2010, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

The products of interactions between galaxies with a high mass ratio and low orbital angular momentum are studied. The interactions scatter the material from the smaller galaxy into structures with distinctive dynamics and morphology, including high local densities and a simple density profile related to properties of the participating galaxies. The role of the larger galaxy's tides in creating these structures and their relation to a well-studied class of mathematical objects motivates us to name them "tidal caustics".

We study the densities achievable in tidal caustics for a typical merger of this type using an example from the Andromeda galaxy to determine whether they are sufficient to produce a detectable gamma-ray signal from self-interactions in the dark matter component, for likely particle models of dark matter. We find that the expected signal is an order of magnitude too low to be detected with current instruments.

We also study the constraints that can be placed on the properties of the participating galaxies by observing the surface brightness profiles of the tidal caustics. We find that the local gravity and gravity gradient of the larger galaxy, and the energy spread and initial phase space density of the smaller galaxy, can be jointly constrained by fitting this profile. The constraints are degenerate but model-independent. We find that measurements of multiple caustics and the velocity of the material in each caustic along the line of sight give information about the orbital angular momentum and the deviations from spherical symmetry in the larger galaxy, though this information is somewhat model-dependent. We discuss the main technical difficulty in fitting the surface brightness profile: determining the inclination angle of the caustic. We demonstrate that a simple model can successfully recover the necessary parameters for some cases, and that a simple modification to this model will improve its success rate.

Thesis Supervisor: Edmund Bertschinger
Title: Professor of Physics
Chair, Department of Physics

Acknowledgments

TBA

Contents

1	Introduction	23
1.1	Dark matter and galaxy formation	23
1.1.1	Particle candidates for dark matter and indirect detection . .	24
1.1.2	Dark matter and galaxy formation	25
1.1.3	Dark halos and tidal streams	27
1.2	Radial mergers and caustics	28
1.2.1	Caustics	31
1.2.2	Caustics created by gravity	33
1.3	Goals of this work	34
2	Calculation of the gamma-ray flux from tidal caustics in the Andromeda galaxy	37
2.1	Introduction	37
2.2	The Optimal Estimator for High Density Contrast	41
2.2.1	Eliminating Poisson Bias	45
2.2.2	Performance of Estimators on the Uniform Distribution	48
2.2.3	Additional Bias for Systems with High Density Contrast	50
2.2.4	Understanding the Undersampling Bias	53
2.2.5	Performance of Estimators on the Non-uniform Distribution . .	56
2.3	Calculation of the Boost Factor and Gamma-ray Flux	67
2.3.1	N-body Model	67
2.3.2	Formulae	68
2.3.3	Calibration of the Rate Estimate	72

2.3.4	Boost factor	81
2.3.5	Astrophysical factor	83
2.3.6	Gamma-ray Signal	84
2.4	Conclusions	84
2.5	Future Work	87
3	Derivation of an analytic, one-dimensional caustic	89
3.1	Density Profile	89
3.2	Gravitational caustic model	95
3.3	Integrated density-squared	97
4	Constraining properties of dark matter halos with tidal caustics	101
4.1	Introduction	101
4.2	A universal density profile	104
4.3	Initial conditions	107
4.4	Spatial distribution	110
4.5	Accounting for projection effects	110
4.6	Phase space distribution	111
4.7	Multiple caustics	115
4.8	Simulations with a live halo	116
4.9	Future work	122
4.10	Conclusion	124
5	Constraints on halo mass profiles from surface brightness measurements of tidal shells	125
5.1	Introduction	125
5.2	Deprojection of two-dimensional measurements	126
5.2.1	Equations for cone approximation	126
5.2.2	Example fit and limitations of the model	131
5.3	Fitting an N-body model of two tidal shells	143
5.3.1	Preparation of sample data	143

5.3.2	Fitting the reoriented shells	144
5.3.3	Comparison with full three-dimensional fit	145
5.3.4	Systematic error: the tilt	149
5.3.5	Fit to Caustic 2 in its original orientation	152
5.4	Future work	154
5.5	Conclusion	156
6	Conclusions	157
A	Rate Estimators	161
A.1	Nearest Neighbor	161
A.2	FiEstAS	163
A.3	Kernel-based	164
B	Calculation of the Minimum Nearest-neighbor Distance	167
C	Technical information	171
C.1	N-body integration	171
C.1.1	Smoothing	171
C.1.2	Force computation	172
C.1.3	Time integration	173
C.2	Computer programs used in this thesis	174
C.2.1	Serial tree code	174
C.2.2	Axisymmetric potentials	174
C.2.3	Self-consistent N-body realizations	176
C.2.4	GADGET	177
C.2.5	Beowulf Cluster	178
C.3	Downhill Simplex Optimization Method	179

List of Figures

- 1-1 Material in a small satellite galaxy (blue points) falling into the potential of a large host galaxy (not shown) initially has a relatively small spread in both velocity and position compared to the entire accessible region (left panel). In order to conserve phase space volume, therefore, the stripped material must reduce its velocity dispersion as it spreads out in position (right panel). In the right panel, the satellite has passed near the center of the galaxy once and some material was stripped while some remains bound to the satellite. 29
- 1-2 Computer simulation of a nearly-radial merger of a small satellite (blue points) with a large host galaxy (not shown, but the center is marked with a green target in the left panel). The merger creates distinctive tidal features (left panel) that reflect the shape of the highly correlated phase-space distribution (right panel). In this example, the satellite's initial velocity vector was pointed 15 degrees away from the center of the galaxy (located at the origin) and the response of the host galaxy was ignored for the reasons stated in the text. The satellite galaxy experienced two pericenter passages; during the first one about half the material was stripped off by tides and in the second one the remaining material was stripped. The two tidal tails extending from the center correspond to the two epochs of stripping. 30
- 1-3 An example of the formation of a caustic in a one-dimensional force-free system. The first column shows the initial state of the system, and time increases to the right. See the text for a discussion. 32

2-1	A three-dimensional random N-body realization of the uniform distribution. All the points (particles in the realization) in the full volume are used to estimate the density; to avoid edge effects, the volume tessellation and Riemann sum are performed over the volume indicated by the yellow box.	46
2-2	Example of adaptive volume tessellation referred to in the text, produced by the space-filling tree in the FiEstAS estimator. The colors, from red to blue, indicate the way in which the tree divides the space (the spatial distribution of the inorder traversal). For clarity, not all the boxes with edges on the boundaries are shown.	58
2-3	All the rate estimators except for $\hat{\Gamma}_u$ (red circles) require slightly different but similarly shaped bias corrections. Equation (2.7) is an adequate ansatz for the N_s -dependence of the correction: an example corresponding to Equation (2.6) is shown as the black dashed line and the green filled points show the effect of using Equation (2.6) to correct $\hat{\Gamma}_f$. Fit parameters are listed in Table 2.2; Table 2.1 explains the label abbreviations.	59
2-4	Once estimators are corrected for Poisson bias using the fitting formula in Equation (2.7), the corrected versions all have a bias much smaller than the typical standard deviation of a few percent. The residual bias is comparable to that of $\hat{\Gamma}_u$. Table 2.1 explains the label abbreviations.	60
2-5	The RMS error of the corrected estimators is around five percent, and declines with increasing N_s . $\hat{\Gamma}_f$ (green diamonds) and $\hat{\Gamma}_u$ (red circles) have the smallest RMS error at low N_s . Introducing an adaptive Riemann volume (orange squares) increases the RMS error significantly at low N_s . The RMS error starts at a higher value and scales more slowly with N_s for the kernel-based estimators (blue triangles and purple stars) than for the nearest-neighbors estimators, an effect of the weighting function used to estimate the density. Table 2.1 explains the label abbreviations.	61

- 2-6 An example of a random N -body realization of the analytic one-dimensional caustic used to test estimators for undersampling bias. The points are the locations of the particles in the realization. The yellow box indicates the integration volume, positioned to include the complete caustic shape and avoid edge effects. For this sample, $N_p = 10^4$, $\sigma = 0.01$, and the dimensions of the box are given in Table 2.3. 62
- 2-7 Using a larger number of particles resolves sharper caustics in all three cases: $\hat{\Gamma}_u$ (a), $\hat{\Gamma}_f$ (b), and $\hat{\Gamma}_n$ (not shown, but very similar to the two others). Regions colored dark or bright blue are considered fully resolved because the bias is less than the typical standard deviation of about 2 percent. The different colors indicate the magnitude of the undersampling bias b_{us} : blue indicates $|b_{us}| < 0.02$, aqua $|b_{us}| \sim 0.05$, yellow $|b_{us}| \sim 0.1$, and red $|b_{us}| > 0.2$ 63
- 2-8 Using a lower value of N_s leads to lower bias whether or not the realization fully resolves all the features in the underlying distribution (a, main figure and inset). The increase in the standard deviation, and therefore RMS error, from using a smaller N_s is small compared to the improvement in the bias when there are under-resolved features (b). If the realization is fully resolved, the increase in RMS error is detectable but extremely small (b, inset). 63
- 2-9 The adaptive Riemann volume algorithm is sensitive to the treatment of the boundaries of the integration volume relative to the geometry of the density distribution. Thanks to the location of the caustic parallel to one boundary of the integration volume, the rate is overestimated (solid blue lines) at marginal resolutions unless the Riemann volumes on that boundary are trimmed to fit the face of the caustic (dot-dashed green lines), but trimming all the boundaries in the same manner underestimates the rate (dashed cyan lines). 64

- 2-10 All three estimators tested on non-uniform distributions achieve a convergence rate of $N_p^{-1/2}$ (dashed line) as predicted by prior analytical work. Tests with $N_s = 10$ and $\log_{10} \sigma = -0.75$ are shown. 65
- 2-11 All the estimators have very similar performance. $\hat{\Gamma}_u$ (red circles) converges faster in the marginally-resolved regime than the estimators that use adaptive Riemann volumes (orange squares and green diamonds). In the fully resolved regime the RMS errors of the three are nearly identical (inset). 66
- 2-12 Sorting particles by the number of pericenter passages, N_{peri} (the inset gives the color scheme used here), easily identifies the main dynamical structures in the tidal debris. This x - y projection shows the N-body model as it would be seen from Earth, with x and y measured relative to M31's center and aligned with the east and north directions on the sky, respectively. In this projection the edge of the younger (red) shell is sharper than that of the older (green) shell, but both are nearly spherical relative to M31's center. 74
- 2-13 Looking at the projected phase space of r and v_r , measured relative to M31's center, lets us determine the order in which the shells formed. In the self-similar model, the outermost caustic forms first: we see here that this corresponds to the bound particles with lowest N_{peri} (colors are the same as in Figure 2-12). The youngest caustic contains the particles that have undergone the largest number of orbits. 75
- 2-14 Each caustic surface was corrected for asphericity by fitting the position of the peak bin in r as a function of angle, first in the ϕ direction (top row) and then in the θ direction (bottom row). The points show the peak bin location and the lines indicate the best fit for the position of the caustic surface. In the background is the two-dimensional binned density map of the caustic. A linear fit was sufficient for all but the ϕ -dependence of Caustic 2 (top right panel) which used a quadratic fit. 76

2-15	The radial density profiles of caustics 1 (left upper panel; green) and 2 (right upper panel; red) can be fit surprisingly well with the functional form in Equation (2.16) after they have been corrected for asphericity with the process illustrated in Figure 2-14. The insets in the upper panels show how well the function fits the region right around the peak of each caustic, which is the most important region for determining the true width. The residuals (lower panels) from the data used in the fit (shown in a darker shade) are evenly scattered around zero in each case, indicating that the peak's relative height is accurately determined. The fit parameters are given in Table 2.5.	79
2-16	The largest boost factors, shown in red, are from the edges of the two shells. The contrast in this figure is independent of the parameters for the particle physics model (summarized as Φ^{SUSY}).	82
2-17	Gamma-ray emission from the dark halo, though faint, dominates over emission from the shell even at large radius (left panel). If $\Phi_{\gamma, hh}$ is removed, the remaining emission from the shell is too faint to detect with Fermi even for the most optimistic parameters in the set of benchmarks (right panel). In these images the pixels are 0.1 degree on a side to imitate the approximate degree resolution of Fermi at the energy scale of interest, and the zero of the degree scale is centered on M31's center.	86
3-1	Universal density profile for a one-dimensional caustic from with initial Gaussian velocity dispersion. The parameters are defined in the text.	98

4-1	Left: Universal caustic form for radial infall from a population with initial Gaussian velocity dispersion. r_c and δr are defined and formulae for the peak radius and density are given in the text. Right: An example phase space distribution taken from a simulation of a minor merger. Near the peak of one of the caustics (black), the phase space distribution follows the curve $r = r_c - \kappa(v_r - v_c)^2$ (dashed line), where v_c is the radial velocity of particles at the caustic surface. κ is defined in the text.	104
4-2	The appearance of tidal debris depends on the angle ϑ between the initial velocity and radius vectors of the satellite galaxy's center of mass. Caustics (sharp edges) are prominent for $\vartheta \lesssim 35^\circ$	108
4-3	Examining the projected phase space shows that when ϑ is larger about 30 degrees, the angular momentum is large enough to significantly deform the shape of the phase space distribution in this projection. .	109
4-4	Angles describing the orientation of a cone representing the spatial extent of the tidal debris relative to the line of sight (z axis). As in standard spherical polar coordinates, θ_c is the angle between the line of symmetry of the cone and the positive z axis, which points away from the observer, and ϕ_c is the projected angle of the line of symmetry in the x - y plane (the plane of the image). α is the opening angle of the cone. If θ_c is near $\pi/2$, α can be approximately measured in the x - y plane.	112
4-5	Using the simple geometric model described in the section (top), the surface brightness profile in a given direction (red dashed line in top panel) for a given orientation of the cone with respect to the line of sight can be derived analytically (bottom). The model pictured has $\alpha = 0.5$, $\theta_c = 1.3$, $\phi_c = 3.4$, and $\delta r/r_c = 0.3$	113

- 4-6 In a spherical potential, the scattering approximation can successfully predict the approximate relative angles of the tidal stream (red), first caustic (orange), and second caustic (green), although by the third caustic (blue) the additive errors in angle begin to dominate. Here the mass profile of the host is the isochrone potential. As discussed in the text, the second tidal stream (the orange material extending beyond the shell) is an indicator that two stripping episodes took place; here we consider only material stripped in the first episode. 117
- 4-7 An N-body simulation of a minor merger in a flattened cored isothermal potential, $V(\vec{r}) = \frac{1}{2}v_0^2 \ln(R_c^2 + x^2 + y^2 + z^2/q^2) + V_0$, with lines showing the predicted angles of the debris at each time based on the center-of-mass orbit, shows that this technique successfully predicts the relative angles of caustics. The potential in this example had the parameters $v_0 = 450$ km/s, $V_0 = -0.4$ kpc² Myr⁻², $R_c = 8$ kpc, and $q = 0.5$, and is oriented as if the host galaxy were face-on (the x - y plane is the plane of the sky). The same parameters were used for the N-body simulation and the integration of the center of mass motion. Notice that the shells are no longer all equally sharp, thanks to the oscillation of the orbit in the z direction (along the line of sight). 118
- 4-8 Comparing the result of a simulation with a live host halo (blue points) with one using a static halo (red points), we see that dynamical friction affects the orbit of the material still bound in the satellite, but does not disrupt the streaming of the unbound material. The relative sizes of the two bound structures differ because the particles in the live halo were plotted using larger points to highlight them; in fact a comparable amount of mass is stripped from the satellite in both cases. 123
- 4-9 The density profile of the caustic from the live halo (blue points) is wider than that in the static halo (red points), but both can be fit with the same density profile (blue and red lines). The points used in the fit are shown in a darker shade. 123

5-1	N-body realization of the model described in the previous section. One-fifth of the particles are plotted. The parameters used to generate the realization are given in Table 5.1.	133
5-2	Sample data used to test the fitting program. On the left, the color bar shows the magnitudes of the measured “surface brightness” (definition is in the text). On the right, it shows the relative error.	134
5-3	Example of convergence of the fitted parameters for the sample data shown in Figure 5-2. The dashed line indicates the restart of the fit after a false minimum. A large value of χ^2 per degree of freedom, as shown in the bottom panel, indicates that a false minimum may have been identified and the fit algorithm should be restarted to check. . .	136
5-4	Estimates and error contours for correlated parameters δr and $f_0/\sqrt{2\pi\kappa}$. The cross indicates the true values of the parameters; the ellipse shows the 68% confidence interval.	137
5-5	Fit residuals scaled by the error on the data points for the best-fit model. The fit is within a few sigma of the data points for nearly the whole fitted region, except at the edges where the surface brightness is least well-determined.	138
5-6	Two different tests for correlations between parameters, the matrix of eigenvectors of the Hessian (left) and the correlation matrix R (right), are defined and discussed in the text.	141
5-7	Eigenvalues of the Hessian for the sample data.	142
5-8	Data sets for Caustics 1 (left) and 2 (right) submitted to the fitting program, with the cuts made as described in the text. The color bar now indicates the normalized surface brightness.	144

5-9	Progress of the iterative fit method for Caustics 1 (left) and 2 (right). Successive restarts are denoted with vertical dashed lines. In the first two steps, half the parameters are held constant and the other half allowed to vary; in the final step, all the parameters are allowed to vary. All parameters are compared to the values given in the last column of Tables 5.3 (for Caustic 1) and 5.4 (for Caustic 2) except for the normalization, which is compared to its final value.	146
5-10	68% confidence ellipses for the correlated parameters δr and $f_0/\sqrt{2\pi\kappa}$ in Caustic 1 (left) and Caustic 2 (right).	147
5-11	Matrices of eigenvectors of the Hessian at the best-fit point for Caustics 1 (left) and 2 (right).	148
5-12	Correlation matrices R for Caustics 1 (left) and 2 (right).	148
5-13	Eigenvalues of the Hessian at the best-fit point for Caustics 1 (left) and 2 (right).	149
5-14	Recovery of the orientation angles of the caustics is sensitive to irregularities in the distribution of material, which is also seen in Figure 5-8. The recovered values (solid vertical lines, with uncertainties denoted by dotted lines) are often offset from the midpoint (dashed line) thanks to the irregular distribution of material. The angular extent α (delimited by dot-dashed lines) is assumed in the model to be the same in θ and ϕ , but in reality the two ranges differ. The recovered value must intermediate between them and generally leans low.	150
5-15	Energy (in $(\text{km/s})^2$) varies in Caustics 1 (left) and 2 (right) as some function of the angles θ and ϕ , producing a variation in r_c	152

5-16	The surface brightness map of the N-body model of Caustic 2, viewed in its proper orientation with respect to the observer, has a bifurcation in the shell edge that is not observed in the data used to create the model (left panel). In addition to attempting a fit of the entire shell, we also tried to fit the upper and lower pieces (delineated by the dashed line in the left panel) separately. This was motivated by the double-peaked structure in the surface brightness profile (right panel).	153
5-17	The fit to each half of the caustic attempts to fit the background density rather than account for the structure in the shell. The top row shows the upper half of the caustic; the bottom row shows the lower half (as in Figure 5-16). The leftmost column shows the surface brightness data being fitted on a normalized scale from 0 (blue) to 1 (red). The center column shows the “best-fit” model on the same scale. The rightmost column shows $(\Sigma(\vec{x}_i) - \hat{\Sigma}_i)/\sigma_i$ at each point, on a scale from -3 (blue) to 3 (red), with zero colored green.	155
B-1	The N_s -dependence of $E(\hat{r}_{N_s, \min})$ (a, thick line) is not close to the asymptotic prediction (dashed line) in the range of interest, and varies somewhat with N_p . We take an average slope of $1/2$ (solid thin line) for the power law index in N_s . However, the N_p -dependence of $E(\hat{r}_{N_s, \min})$ (b) is close to and approaches the asymptotic prediction (dashed line), so we use the asymptotic slope of $-1/3$ for the N_p scaling relation. . .	170
C-1	Relaxation time in the core of a Milky-Way-like halo as a function of resolution.	177

List of Tables

2.1	Various estimators of the gamma-ray emissivity Γ . See Appendix A for longer descriptions of the various estimators.	47
2.2	Best-fit bias corrections. Error ranges indicate the 95-percent confidence level. The form of the bias correction is given in Equation (2.7). See Appendix A and Table 2.1 for descriptions of the various estimators.	48
2.3	Parameters for the analytic one-dimensional caustic. These quantities are defined in more detail in Chapter 3. The units are given as dimensions only since they may be scaled as needed.	54
2.4	Parameter space for testing undersampling bias.	54
2.5	Density profile parameters for caustics 1 and 2. Error ranges on fitted parameters indicate the 95 percent confidence interval.	80
2.6	Values of the astrophysical factor Φ^{cosmo} for various configurations of the tidal debris, calculated using Equation (2.30).	83
2.7	Contribution to the flux of gamma-rays above 1 GeV for various MSSM benchmarks.	85
4.1	Comparison of fit parameters for a caustic in a live versus a static halo.	121
5.1	Parameters for sample data. The choice of units is arbitrary for this test data, but dimensions are given in the table.	134
5.2	Summary of fit to sample data. Standard error is calculated using Equation (5.41). For error contours for correlated parameters, see Figure 5-4.	137

5.3	Fit information for Caustic 1. The uncertainties for the fitted parameters are calculated from the Hessian as described in the text. Contours for the correlated parameters are shown in the left panel of Figure 5-10. True values of the parameters were calculated using the full N-body representation in three dimensions.	146
5.4	Fit information for Caustic 2. The uncertainties for the fitted parameters are calculated from the Hessian as described in the text. Contours for the correlated parameters are shown in the right panel of Figure 5-10. True values of the parameters were calculated using the full N-body representation in three dimensions.	147
5.5	Energy variation and corresponding radial variation in caustics 1 and 2, compared with fit results.	151

Chapter 1

Introduction

1.1 Dark matter and galaxy formation

Dark matter is an invisible substance that is thought to make up most of the Universe's mass. Unlike matter comprised of Standard-Model particles like protons, neutrons, and electrons, dark matter does not interact appreciably with light. All evidence of its existence to date comes from observing its gravitational influence on Standard-Model (often called “luminous”) matter. The first evidence that a significant amount of the gravitating matter in the universe might not interact with light came from comparing mass estimates based on the velocities of galaxies orbiting each other in a nearby cluster (the Coma cluster) with an estimate of the total amount of mass in the cluster based on the galaxies' brightnesses. The two methods disagreed by several orders of magnitude [1]. The same discrepancy between the inferred mass from the velocities of orbiting material and the observed mass estimated from brightness measurements was identified in the orbits of galaxies in other groups and clusters [2] and in the rotation speeds of disk galaxies [3]. Today dark matter is generally acknowledged to be necessary to explain a wide range of well-measured phenomena, from the small-scale patterns in the cosmic microwave background [4] to the clustering of galaxies in the present-day universe [5] to the rate at which the universe is expanding [6, 7]. It is thought to make up about 90 percent of the gravitating mass and 26 percent of the total energy content of the universe [4]. However, despite

its crucial role in shaping our universe’s contents and history, no experiment has yet successfully confirmed dark matter’s true identity.

1.1.1 Particle candidates for dark matter and indirect detection

The nature of dark matter is one of the most compelling questions in physics today, and countless theories have been proposed though none have yet been confirmed. The most popular theory at present, and the one that is relevant to this thesis, is that dark matter is indeed a particle but not part of the Standard Model. Instead it is thought to be part of the extended hierarchy of new particles proposed under supersymmetry, a well-studied family of theories in particle physics that seeks to answer several important questions posed by the Standard Model [8]. In most versions of the theory, the lightest of the new particles predicted by supersymmetry (called the “lightest supersymmetric partner”, or LSP) has properties that match those of dark matter. Such a particle is long-lived: being the lightest supersymmetric particle, there is nothing for it to decay into. Aside from gravitational interactions, an LSP interacts very infrequently with Standard-Model particles and then only via the weak force. It is also its own anti-particle, an important property that sets the total amount of dark matter in the universe. LSPs that have a mass in the range commonly predicted by supersymmetric theories (50 - 5000 GeV) are produced in the early universe in the correct amount to account for most or all of the dark matter that exists today [9, 10].

Attempts to detect supersymmetric dark matter particles fall into two categories: direct detections, where particles of dark matter interact with regular matter (usually the nuclei of large atoms) in a detector, and indirect detections, which search for the products of interactions between two dark matter particles. Direct detection experiments seek to observe dark matter either by creating it (as is being attempted at the Large Hadron Collider) or by observing it as it passes the Earth [11, 12], while most indirect detection experiments observe astronomical objects thought to contain a high concentration of dark matter [13, 14, 15, 16], since the rate of interaction between

dark matter particles scales as the dark matter density squared. The first part of this thesis evaluates a possible scenario for indirect detection that would avoid one of the primary challenges of this method: confusion between the products of interacting dark matter particles (which are high in energy thanks to the relatively large mass of most candidates) and the products of other common high-energy processes that involve luminous matter (like the accretion of material onto a black hole). An ideal situation for indirect detection would have a high density of dark matter but not of luminous matter. Such situations are rare since the distribution of all kinds of matter in the universe is primarily determined by gravity.

1.1.2 Dark matter and galaxy formation

Through gravity, dark matter also plays a very important role in the formation of galaxies. In supersymmetric models of dark matter this role starts in the first moments of the universe’s history. Just as with luminous matter, quantum fluctuations are magnified by inflation and imprinted as density perturbations in the dark matter. These perturbations start to grow in amplitude through gravitational instability while the luminous matter is still a pressure-supported plasma of photons, electrons and protons. By the time the electrons and protons assemble into neutral atoms, the dark matter has already started to form structures—dense patches and rarefied areas—into which the luminous matter falls [17]. The dark matter gets a “head start” on the formation of the cosmic web, without which the luminous matter would not have had time to assemble into the large number of structures we observe in the universe today [e.g. 18].

Dark matter’s extremely small probability of interacting with anything, even other dark matter, makes it a good candidate for N-body simulations. Compared to simulations of luminous matter, in which gas pressure, chemistry, phase transitions, the formation of stars, and the force of gravity must all be somehow taken into account, dark matter is relatively straightforward to simulate since gravity alone usually determines its behavior. The exception to this is at extremely high densities where the rate of dark matter self-interactions is high enough to produce a pressure, but although

such densities might have been attained in some tiny regions in the early universe [19] we can safely ignore this possibility in simulations of the large-scale structure of dark matter or of modern-day galaxy interactions.

In an N-body simulation, dark matter is usually represented by a distribution of bodies, each representing a huge number of dark matter particles, that interact only via a modified form of Newton’s law of gravity. The modification weakens or “softens” the gravitational force at very small distances, avoiding the divergence at zero distance and also preventing unphysical close encounters between pairs of bodies in the simulation. Because this modification is not physical, the results of a dark matter simulation are valid only at length scales much larger than the scale at which the force is softened. More information about the technical aspects of simulating dark matter is provided in Appendix C.

Large-scale N-body simulations of the growth of dark matter structures from small density fluctuations consistent with those measured in the cosmic microwave background show the formation of a cosmic web and predict that in today’s universe, dark matter structures have formed at a wide range of scales [e.g. 20]. Roughly spherical, gravitationally bound structures about the size and mass of a large galaxy are referred to as “halos”, since most of them are thought to contain a galaxy made of luminous matter [21, e.g.]. They tend to contain both a smooth distribution of dark matter that increases in density toward the center and a large number of smaller structures known as “subhalos” that have approximately the same density profile but are much smaller in mass and size [22, 23]. These subhalos are of diverse origin: some are formed within the parent, or “host” halo, while others are accreted continuously as the host halo grows in mass [24]. Based on comparisons with observations of galaxies, they are also of diverse content: in our galaxy, for instance, it seems that the largest subhalos all contain small galaxies (known as “dwarf galaxies” or “satellite galaxies”; [25]) but simulations also predict many smaller subhalos that either contain very little luminous matter (called “ultra-faint dwarfs”, [26, 27]) or none at all (and are thus virtually impossible to detect). Reconciling the amount of small-scale structure predicted by simulations of dark matter with the number of galaxies we

observe orbiting the Milky Way is an open question and an important test of both dark matter’s properties and theories of galaxy formation.

1.1.3 Dark halos and tidal streams

The hierarchy of structures of different sizes seen in simulations of dark matter also indicates that a significant portion of the material, both dark and luminous, in a Milky-Way-sized galaxy came from previously formed structures that merged together. The latest simulations estimate that the fraction of contributed material is 35–60%, with approximately half of the contributed material still gravitationally bound in subhalos and the other half spread across the host halo in gravitationally unbound, but still spatially or dynamically distinct structures known as “tidal streams” [24]. Tidal streams are made of material that once belonged to a satellite galaxy but has since been stripped away by the tide of the host galaxy so that now it orbits the host instead. At least five tidal streams of stars have been identified in the Milky Way [28] and about a dozen more are seen around other nearby galaxies [29, 30]. Besides serving as evidence that galaxies are indeed assembled by accreting many smaller structures [31], tidal streams are useful for studying the properties of their host halos because the material in them behaves like a set of test particles orbiting in the host halo’s gravitational potential. Modeling the orbits can constrain the mass profile and three-dimensional shape of the host halo. Known tidal streams are usually quite far from most of the other luminous material in a galaxy (otherwise, they would be indistinguishable in images), so their orbits are primarily sensitive to the dark halo [32, 33, 34]: they are one of the few ways to understand the distribution of dark matter in galaxies.

Unfortunately, it is difficult to construct a model of the interaction that created a particular tidal stream from the limited data we have in most cases. Because tidal streams are relatively faint, distant structures even in our own galaxy, it is difficult to measure all components of the position and velocity of the stream material. Distances to astronomical objects are nearly always difficult to measure, as are the components of the velocity in the plane of the sky (the proper motions). The velocity component

along the line of sight, measured by taking a spectrum of the integrated light from the stars in the stream and determining its Doppler shift, cannot be measured either for some streams, since they are too faint to obtain a good spectrum. So for many tidal streams, the only really well-determined measurements are the relative sky positions of material along the stream. On the other hand, the initial position and velocity of the satellite galaxy that created the tidal stream and the mass and shape of the host halo are all virtually unconstrained, and the only way to compare the data with a model is to perform a computationally intensive N-body simulation of the merger. Given that the problem is so underdetermined, any model that successfully reproduces measurements of a tidal stream must be considered only as one of a family of equally valid solutions in a space of parameters with many partial degeneracies.

These difficulties in modeling tidal streams have led to a search for alternative ways of studying them. The second half of this thesis presents and demonstrates a new method for studying tidal streams created in mergers with very little orbital angular momentum. The symmetry of this type of merger gives rise to structures known as caustics, for which theoretical tools have already been developed [35, 36, 37, 38, 39]. This thesis applies these tools to the type of merger that is likely to create a tidal stream.

1.2 Radial mergers and caustics

Images of interacting pairs of galaxies abound in the astronomical literature [40, 41, 29, 42, are some examples]. There is an incredible amount of variety in the features shown in these images, which range from long tails and wide loops to fan-shaped or shell-like structures. The type of feature created by a particular merger depends mainly on the ratio of the masses of the interacting galaxies and on the orbital angular momentum of the system. This thesis is concerned with studies of the tidal streams created when a large galaxy accretes smaller ones, so we may assume that the mass ratio (host to satellite) is large, of order 100 or more. This allows us to simplify the system by ignoring the response of the host galaxy to the gravitational pull of the

satellite. At the end of this thesis we will show that our conclusions are still valid even when this simplification is not made. Additionally, in such a system, the spread of velocities within the satellite galaxy is small compared to the orbital velocity of the satellite around the host. This means that all the material in the satellite galaxy starts out in a relatively small patch of phase space compared to the volume of phase space accessible to all the material in the host galaxy. As the satellite's material is stripped away by tides and spreads out in space, the velocity spread decreases in order to conserve the volume of that phase space patch. The material ends up in one long, thin stream in phase space (Figure 1-1).

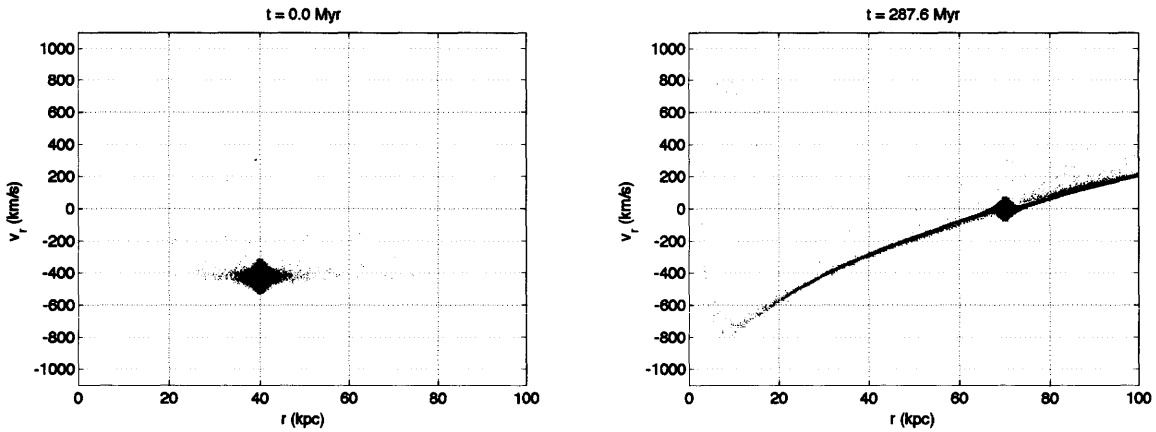


Figure 1-1: Material in a small satellite galaxy (blue points) falling into the potential of a large host galaxy (not shown) initially has a relatively small spread in both velocity and position compared to the entire accessible region (left panel). In order to conserve phase space volume, therefore, the stripped material must reduce its velocity dispersion as it spreads out in position (right panel). In the right panel, the satellite has passed near the center of the galaxy once and some material was stripped while some remains bound to the satellite.

For this work we focus on mergers with low orbital angular momentum, which have two additional advantages. First, the tidal streams created by such mergers are distinctive: from certain viewing angles they resemble fans or shells with a sharp outer edge (Figure 1-2, left panel). Often, as in this example, there are also one or more long tails that point away from the galactic center. As we will show in this section, these features are a reflection of the cooling of the stream thanks to phase space conservation and the particular phase-space path the stream traces out when

very little orbital angular momentum is present. The second advantage is that if the orbital angular momentum is small, the system can effectively be treated in one spatial dimension rather than three, since the amount of orbital precession of the satellite material will be small. This is why in Figure 1-1 and in the right panel of Figure 1-2 the conservation of phase space volume is apparent even though only two of the six phase space dimensions are shown. The combination of a large mass ratio and a small amount of orbital angular momentum gives rise to a particular type of structure known as a caustic, which is responsible for the sharp edges of the shells. We will next discuss the mathematics used to describe them and its significance for the work in this thesis.

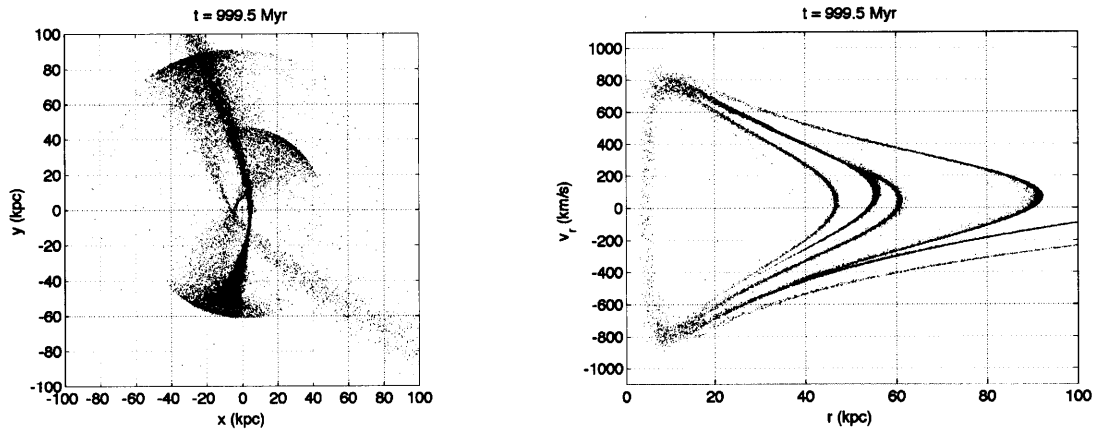


Figure 1-2: Computer simulation of a nearly-radial merger of a small satellite (blue points) with a large host galaxy (not shown, but the center is marked with a green target in the left panel). The merger creates distinctive tidal features (left panel) that reflect the shape of the highly correlated phase-space distribution (right panel). In this example, the satellite’s initial velocity vector was pointed 15 degrees away from the center of the galaxy (located at the origin) and the response of the host galaxy was ignored for the reasons stated in the text. The satellite galaxy experienced two pericenter passages; during the first one about half the material was stripped off by tides and in the second one the remaining material was stripped. The two tidal tails extending from the center correspond to the two epochs of stripping.

1.2.1 Caustics

In mathematics, a caustic is a feature of a mapping between an initial and final state. It denotes the boundary between regions of the map that are one-to-one and regions in which multiple initial states can lead to the same final state, so at the caustic surface the derivative of the map with respect to the initial state is ill-defined. In physics, one instance where caustics can occur is in the map that represents the dynamical evolution of a system over some time period. The states on which the map operates consist of the set of phase-space coordinates of the components of the system at a given time. In our case the map represents the evolution of the positions and velocities of the material in the satellite galaxy under the influence of gravity.

In some cases, material from multiple initial positions can end up at the same location later on. A simple one-dimensional example is given in Figure 1-3, which shows the force-free evolution in time (left to right) of a set of 500 particles with initially random positions. For this example we will assume that the physical size of the particles is zero, so that they may change their positions freely. The top row of the figure shows the phase space distribution (v_x versus x) of the material. The second row plots the initial position q_x versus the current position x and the bottom row is a histogram of the particle positions, a proxy for the density ρ . The position-velocity relation is contrived so that material from one side of the distribution passes material from the other side, heading in the opposite direction. Before this begins to occur the map from initial to final positions is single valued, as shown in the first three plots of the middle row. After some particles start to pass each other, the map becomes double-valued: there are two values of q_x that could result in the same value of x for some region in x . The caustic is the point where the distribution becomes double-valued (where the plots of v_x and q_x versus x are vertical). At the caustic, a density spike develops (rightmost two plots of the bottom row) as a result of the conservation of particle number (or mass):

$$\rho(x, t)dx = \sum_{q_x \rightarrow x} \rho_0 dq_x \quad (1.1)$$

At the caustic location, $dx/dq_x = 0$, so

$$\rho(x, t) = \sum_{q_x \rightarrow x} \frac{\rho_0}{|dx/dq_x|} = \infty \quad (1.2)$$

as long as $\rho(q_x)$ is larger than zero. The formation of the caustic is a coincidence of the dynamics of the system (in this case, the fact that we chose the initial velocities so that particles would pass each other) but its shape and behavior over time are universal. The existence of a caustic can also be inferred from the phase space picture of the particles (top row of plots) by noting that regions where the phase space distribution is vertical have large density when projected onto the x axis. This is because the initial positions and velocities of the material are correlated. As we discussed previously, the same is true in a high-mass-ratio galaxy merger because the satellite galaxy initially occupies such a small patch of phase space. Although the time-evolution of such a system is much more complicated than this simple force-free example, the phase-space diagram will still reveal the existence of caustics.

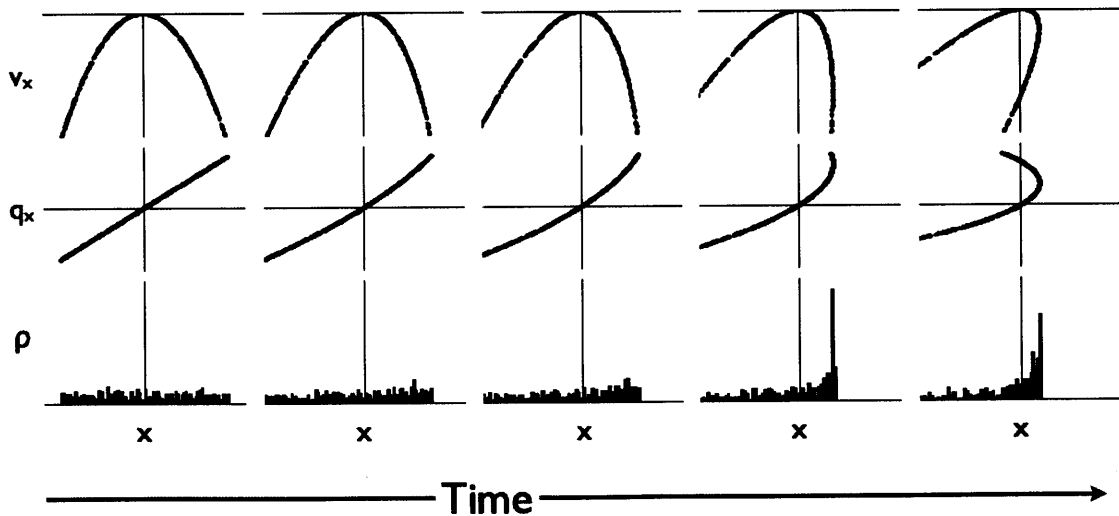


Figure 1-3: An example of the formation of a caustic in a one-dimensional force-free system. The first column shows the initial state of the system, and time increases to the right. See the text for a discussion.

1.2.2 Caustics created by gravity

Caustics in systems governed by the gravitational force law have a few important differences from the simple force-free system used as an example in Figure 1-3. Our simple example resulted in only one caustic, but since bound orbits in a gravitational potential are periodic, multiple caustics can form in a gravitating system. The formation of the successive caustics occurs because of a spread in energy in the orbiting material, which for most potentials corresponds to a spread in orbital periods. The material with the shortest orbital period eventually catches up with and “laps” the material with the longest periods, like runners in a long-distance race on an indoor track. This process is called phase-wrapping, and each time it happens another loop appears in the projected phase space diagram (the equivalent of Figure 1-2, right panel). Each loop corresponds to a caustic, just as in Figure 1-3. The loops are sorted by their orbital phase: outermost loop is made of material on its first orbit after being stripped, the second-outermost of material on its second orbit, and so forth.

In a radial merger the caustics form from material that has been stripped by tides from the small satellite galaxy, so every time more material is stripped, another set of caustics is formed. Tidal stripping in a radial merger occurs each time the satellite reaches pericenter, until all the material is unbound. For extremely low values of orbital angular momentum, the entire satellite can be destroyed after only one or two pericenter passages. The example phase-space distribution in the right panel of Figure 1-2 has four caustics: three from the material stripped off after the satellite’s first pericenter passage and a fourth created from the material unbound after the second pericenter passage, which has been orbiting for a shorter time. In a bound, nearly-radial orbit the caustics occur roughly, but not exactly, at the outer turning points (apocenters) of the orbits of the stripped material, corresponding to the edges of the shells in the left panel of Figure 1-2. This is why we use the term “tidal caustics” to describe such features.

The caustics in Figure 1-2 do not occur exactly at the turning point (on the r

axis). This is because the phase space distribution does not trace out the orbit of one test particle, but instead is a snapshot of many particles at neighboring locations in their orbits, so the point where particles are passing each other is not necessarily at zero velocity. We will show in this thesis that although the force-laws that produced the phase space distributions in Figures 1-2 and 1-3 are completely different, the shape of the distribution in the region of the caustic and the shape of the density spike are the same in each case.

Caustics produced by gravity have a finite maximum density, contrary to the prediction of Equation (1.2). This is a result of the extra spatial dimensions in the three-dimensional case: although dr/dq_r might be zero in Figure 1-2, the full derivative $d\vec{x}/d\vec{q}$ cannot be zero because the phase space volume is incompressible. In fact, the maximum density is determined by the size of the finite phase space volume: the smaller the initial volume, the “colder” the material and the larger the maximum density in the caustic. For a typical dwarf galaxy, this maximum achievable density can be quite high. Although the finite maximum density and width mean the structure is no longer truly a caustic in the mathematical sense, its properties change very little otherwise.

1.3 Goals of this work

The work described in this thesis had two main goals. The first was to assess whether tidal caustics, thanks to their ability to produce an extremely high density at a large distance from the galactic center, could provide a confusion-free scenario for detecting the products of interactions between dark matter particles. For this work we used an N-body model of a real set of caustics discovered around the Andromeda galaxy, a close neighbor of the Milky Way. The Fermi gamma ray observatory, now in operation, can distinguish between high-energy radiation coming from these caustics and that coming from Andromeda’s center, if the signal from the caustics is bright enough to be detected at all. We used the N-body model to estimate the size of the signal and compare it to the sensitivity of Fermi. This work is discussed in Chapter 2.

The second goal was to determine what information, if any, about the host and satellite galaxies in a merger that formed tidal caustics could be gleaned from images of these features. The theoretical motivation for this portion of the work comes from the fact that the very existence of tidal caustics restricts the properties of the merger significantly, and from the fact that the density profile of the caustics depends very simply on a small number of characteristics of the two galaxies (as presented in Chapter 3). The practical motivation is that images of such caustics surrounding many different types of galaxies already exist, and a method that could reveal characteristics of these galaxies would allow us to study mergers in a wide range of environments, whereas the current sample of modeled tidal streams includes only our galaxy and Andromeda. In Chapter 4 we derive the constraints that can be placed on the host and satellite galaxies by fitting only the images of the caustics, and additional constraints that can be gleaned from analyzing multiple shells or from determining line-of-sight velocities from spectra (the two additional pieces of information most likely to be obtained). In Chapter 5 we discuss the crucial step in obtaining these constraints: obtaining the radial density profile by deprojecting the two-dimensional image.

Chapter 2

Calculation of the gamma-ray flux from tidal caustics in the Andromeda galaxy

This chapter was published as the body of [43]. Appendix C of [43] appears as Chapter 3 in this thesis; Appendices A and B of [43] appear as Appendices A and B in this thesis.

2.1 Introduction

The nature of the dark matter is one of the foremost questions in astrophysics. Although most astrophysicists agree that it is probably some kind of particle [8], there are to date no conclusive detections. Countless experiments are attempting to do so, whether by detecting dark matter particles directly [44, 45, 46, 12, 11, 47, 48], creating them in the laboratory [49, 50], or observing the standard-model byproducts of interactions between them [14, 51, 52, 15, 13, 53]. This last method, usually referred to as “indirect detection”, usually assumes that the dark matter particle is its own antiparticle, so that an interacting pair of dark matter particles self-annihilates to produce various kinds of standard model particles. This assumption is motivated by predictions of supersymmetry that the dark matter could be the lightest supersym-

metric partner (LSP) of a standard model particle, and by the cosmological result that such a particle with a mass of between 20 and 500 GeV would have been produced in the early universe in sufficient number to resolve the discrepancy between the energy density of luminous matter observed today and the total energy density of matter required to explain the gravitational history of the universe [8].

Indirect detection, because it relies on observing the products of pairwise annihilation, has a signal strength that varies as the dark matter density squared. It is therefore most effective in regions with the very highest number density of dark matter particles. Because dark matter interacts with luminous matter primarily through gravity, these are often the same regions where the density of luminous matter is highest, such as the centers of galaxies [54]. Although the signal from pairwise dark matter annihilation may be highest in those regions, it also suffers from confusion with high-energy astrophysical sources like pulsars and X-ray binaries, which tend to be concentrated wherever the density of luminous matter is high [55]. However, in some instances the dark matter density can be high while the luminous matter density is low; for example, in the recently discovered ultra-faint dwarf galaxies orbiting the Milky Way [56]. Since the dark matter has a lower kinematic temperature than luminous matter, it may also have more small-scale structure than luminous matter, increasing or “boosting” the production of standard model particles by pairwise annihilation above the level predicted for a smooth distribution [57, 58, 59, 60, 61, 62, 63]. Cases where there is no confusion between particles produced by pairwise dark matter annihilation and those produced by high-energy astrophysical sources offer a high potential for a confirmed indirect detection, provided that the signal is still detectable.

One possible scenario for indirect detection is the debris created by a merger between a larger host galaxy and a smaller progenitor galaxy on a nearly radial orbit. [64, 65] showed that in such a case, the mass from the progenitor accumulates at the turning points of its orbit, producing shells of high-density material at nearly constant radius on opposing sides of the host galaxy. The dynamics governing the formation and shape of the shells can be understood in the context of earlier work on spherically symmetric gravitational collapse. [36] and [35] demonstrated that radial

infall of gravitating, cold, collisionless matter forms a series of infinite-density peaks at successive radii, known as caustics. [37] extended this case to include warm matter with various velocity dispersions: the effect of the velocity dispersion is to make the peaks of finite width and height, so they are no longer caustics in the mathematically rigorous sense but retain many of the same properties, including the possibility of extremely large local density enhancements. There have been multiple attempts to estimate the production of gamma rays through self-annihilation from infall caustics in the dark matter halos of galaxies [e.g. 66, 67, 68, 69, 60, 70, 71, 72, 73] or to otherwise determine ways in which dark matter in caustics could be detected [72, 69, 67]. Most work on the gamma-ray signal has found that caustics enhance the production from a smooth distribution by a factor of between 10 and 100. In the case of so-called “tidal caustics” like those seen around shell galaxies, the infall is not spherically symmetric, but the density is still clearly enhanced in the shells.

Dark matter and luminous matter alike are concentrated in tidal caustics at large distances from the bulk of the luminous matter in the host galaxy. Shell galaxies [see, e.g., 41] are beautiful examples of the extreme case of this phenomenon, where the angular momentum of the progenitor is nearly zero. Unfortunately, all the known shell galaxies are too far away to indirectly detect the dark matter in the shells: the flux of gamma rays is too attenuated, incoming charged particles are deflected by the Galactic magnetic field, and high-energy detectors have insufficient angular resolution to separate the shells from the host. However, M31 also appears to have shells [32], though they are not as symmetric as those in classic shell galaxies, and is close enough that the Fermi LAT can distinguish their position from that of M31’s center [74]. Furthermore, an N-body model of the shells already exists [33] that can be used to estimate whether dark matter in them could be indirectly detected.

N-body models of dark matter distributions have been used to estimate standard-model particle fluxes for indirect detection in the Galactic halo and the factor by which dark matter substructure could increase the rate of pairwise annihilations [59]. Both these quantities are proportional to the volume integral of the square of the dark matter density, which we will call the “rate” for short. The rate is estimated from

an N-body representation of the dark matter distribution by substituting a Riemann sum for the volume integral and inferring the density in each Riemann volume from the N-body representation by one of many well-studied methods. However, neither the estimation of the square of the density rather than the density itself nor the choice of a suitable set of Riemann volumes has been tested. Likewise, the ability to recover the correct rate when the density distribution is sharply peaked has not been explored, although this is the scenario that would most likely lead to an observable signal and the reason that the M31 shells are of interest.

This paper describes tests of a number of well-known algorithms for calculating the rate and discusses the best algorithm to use in situations where the density gradient is large (Section 2.2). We then present estimates, using the optimal algorithm, of both the boost factor from the M31 shells over the smooth distribution of dark matter in M31's halo and the rate at which gamma rays from pairwise annihilation would be seen by Fermi given likely parameters of a supersymmetric dark matter candidate (Section 2.3).

We find that the best way to estimate the rate from an N-body representation, whether the density is nearly uniform or has a large gradient, is with the simplest possible method: a nearest-neighbor estimator with a relatively small smoothing number to find the density, and fairly small constant Riemann cubes to perform the integral. This result is surprising, given that so many more sophisticated density estimators exist. We further find, using this result, that the largest boost factors from tidal debris in M31 are 2.5 percent in the most concentrated regions of the shells, and that the gamma rays from the debris are too few to be detected by Fermi for likely supersymmetric dark matter candidates: the total additional flux in gamma rays, which is model-dependent, is less than $7.4 \times 10^{-11} \gamma \text{ cm}^{-2}\text{s}^{-1}$ for likely dark matter models.

An ancillary result from our analysis of the tidal caustics, and a consequence of radial infall, is that the density profile of each shell and the radial spacing of the shells depend on the radial derivative of the gravitational potential at each shell and on the mass and size of the dwarf galaxy before infall. This means that information

about the initial qualities of the dwarf galaxy can be inferred from the shells without requiring a detailed model of the potential of the host galaxy. Assuming that the stars in the dwarf galaxy are initially virialized (with a Maxwellian velocity distribution), the density profile can be fit with an analytic function whose width depends on these properties. We discuss further implications of this result for recently discovered shells around other nearby galaxies.

2.2 The Optimal Estimator for High Density Contrast

The key to the calculation presented in this paper is the estimation of the integrated squared density,

$$\Gamma = \int \rho^2 dV, \quad (2.1)$$

from an N-body representation of the dark matter mass distribution. The particles making up the N-body representation are independent observations of the mass density function ρ . Generally, probability density functions are defined as those that are everywhere positive and normalized to one [as in 75, Chapter 4]. The mass density function sampled by the particles in the N-body representation satisfies the first of these two conditions, and dividing by the total mass to get a scaled number density will satisfy the second. So the analysis of estimators for the probability density and its functionals applies equally to the problem at hand. The development and characterization of estimators for this quantity is a well-studied problem in statistics, in the context of estimators for probability density distributions [76, 77, 78, 79, 80, 81, 82, 83, and many others].

Density estimators studied in the literature are divided into two classes: parametric (in which a particular functional form for ρ is assumed) and nonparametric (in which assumptions about the form of ρ are kept to a minimum). Nonparametric estimators are commonly used with data sets like N-body realizations, where the goal is usually to discover the form of ρ and/or calculate other quantities from it [75]. Among

the wide variety of nonparametric estimators available, nearest-neighbor estimators [84] are one of the oldest and most well-studied varieties. The nearest-neighbor estimator uses an adaptive local smoothing length equal to the distance to the N_s^{th} nearest particle to the location where the density is being estimated. The density at that point is then taken to be $N_s/V_d(N_s)$, where V_d is the volume in d dimensions, centered on the target location, that encloses N_s particles. [85] and [86] showed that nearest-neighbor density estimators converge to the underlying distribution at every point as the number of particles in the realization, N_p , goes to infinity, provided that $N_s/N_p \rightarrow 0$ in the same limit. They can also be considered as part of the larger class of adaptive kernel estimators [87] and are even more closely related when V_d is replaced by a weighted sum over the N_s particles [88]. However, because the function they return may not be normalizable, nearest-neighbors is more suited to individual density estimates at a point than to recovery of the entire function [89]. All the estimators we test in this work are based on either the simple nearest neighbors method or one using a weighted sum, although the shape of V_d varies. We describe them in detail in Appendix A.

The usual measure of the quality of a nearest-neighbors estimator is its root-mean-squared (RMS) error,

$$\text{r.m.s.e.} \equiv \frac{1}{\Gamma_{\text{true}}} \sqrt{E \left[\left(\hat{\Gamma} - \Gamma_{\text{true}} \right)^2 \right]}, \quad (2.2)$$

The RMS error compares the expectation value of the estimator, in this case the rate estimator $\hat{\Gamma}$, with the true value of the rate, Γ_{true} . For the tests in this work, we used density distributions for which Γ_{true} may be calculated analytically. [77] demonstrated that, given some constraints on the maximum slope of the underlying density distribution, the error of a one-dimensional integrated squared density estimator with a kernel of a constant size can converge as $N_p^{-1/2}$; [82] recently showed that a simple estimator of this type can be made adaptive using a particular rule to calculate the kernel size from the data and still converge at the same rate. Most of the estimators we test in this work use adaptive kernels with a simpler rule than the one

suggested by [82] for two reasons. The first is simply conceptual and computational simplicity: rules for choosing an optimal kernel size tend to require minimizing the cross-validation function (a proxy for the RMS error) of the data, which requires an optimization program, and the resolution convergence even with the optimal kernel chosen in this way can still be slower than $N_p^{-1/2}$. The second is that extending the result of [82] to several orthogonal dimensions is not trivial [90].

As is common in the literature, we consider the RMS error in two parts: the bias and standard deviation [91], where

$$(\text{r.m.s.e.})^2 = b^2 + \left(\text{std}(\hat{\Gamma})\right)^2. \quad (2.3)$$

The bias, b , is the difference between the expectation value of the estimator and the true value of the parameter it is estimating. An unbiased estimator has $b = 0$, one for which $E(\hat{\Gamma}) > \Gamma_{\text{true}}$ has a positive bias, and one for which $E(\hat{\Gamma}) < \Gamma_{\text{true}}$ has a negative bias. The standard deviation indicates the size of the spread of individual estimates around the expectation value. For this work we scale the bias, standard deviation and RMS error by a factor of Γ_{true} , so

$$b \equiv E\left(\frac{\hat{\Gamma} - \Gamma_{\text{true}}}{\Gamma_{\text{true}}}\right) = \frac{E(\hat{\Gamma})}{\Gamma_{\text{true}}} - 1 \quad (2.4)$$

and

$$\text{std}(\hat{\Gamma}) \equiv \frac{1}{\Gamma_{\text{true}}} \sqrt{E\left[\left(\hat{\Gamma} - E(\hat{\Gamma})\right)^2\right]} \quad (2.5)$$

are consistent with Equations (2.2) and (2.3).

We used numerical experiments to assess the bias, standard deviation, and RMS error of the various estimators, so we must be clear about how these values are calculated numerically. For each experiment, 10^4 random realizations of the density distribution of interest comprise one sample. The expectation value of a quantity is then defined as the mean of that quantity over the set of all random realizations. The random realizations are subject to Poisson fluctuations, so this number of realizations corresponds to sampling error of about one percent. We take 20 samples of the

expectation value, so the relative error on the mean from these 20 samples is about 0.2 percent.

The number of particles in each random realization (shown as points in an example in Figure 2-1) is drawn from a Poisson distribution with a specified mean, denoted in the following sections as N_p . This precaution keeps the number of particles in the density distribution, and in the subset of that distribution used for the volume integral (the shaded box in Figure 2-1), purely Poisson; the error associated with using a fixed number of particles depends on N_p , so we must eliminate it if we wish to establish how the estimators behave as N_p varies.

The method for estimating the rate has two distinct parts: how to determine the number and placement of the Riemann volumes making up the sum, and how to estimate the density in each Riemann volume. We tested five different rate estimators that together use three different well-known density estimation methods and two different ways of assigning Riemann volumes (adaptive and constant). The rate estimators are described in detail in Appendix A and briefly summarized in Table 2.1.

We first evaluate and, if possible, eliminate the bias. We constructed rate estimators using density estimators that have a very small or zero bias when used to estimate the density, but we demonstrate in this section that they do not always produce unbiased estimates of the rate without further correction. We wish to reduce the bias of the estimators when it is possible to do so without increasing their standard deviations. Bias resulting from the statistics of Poisson point processes, here referred to as “Poisson bias,” can be eliminated this way—analytically for some of our estimators, and numerically for the others—once it has been measured using random realizations of the uniform density distribution (Section 2.2.1). We examined the RMS error of rate estimates for the uniform density distribution, after correcting for the Poisson bias, to separate the contribution of Poisson processes to the overall RMS error of each estimator from additional error that arises when the density is not uniform (Section 2.2.2).

The second step in determining the best estimator is to determine the RMS error

in the case where the density distribution has sharp features with high contrast (like shells). Discreteness effects will then introduce additional bias that depends on the resolution, the smoothing number, and the scale of the high-contrast features (Section 2.2.3). To understand when and how this bias contributes, we tested each estimator using random realizations of a simple caustic density distribution that has an analytic expression for Γ_{true} (Section 2.2.4). Finally, we compared the RMS errors of the estimators for the high-contrast density distribution to determine which one to use in the calculation of the gamma-ray flux (Section 2.2.5).

2.2.1 Eliminating Poisson Bias

Using Poisson statistics, it is possible to construct a rate estimator $\hat{\Gamma}_u$ that is unbiased for a uniform density distribution; that is, one for which $E(\hat{\Gamma}_u) = \Gamma_{\text{true}}$ (Appendix A, Equation (A.5)). It uses a constant Riemann volume dV for the integral and the distance to the N_s^{th} nearest neighboring particle, denoted r_{Ns} , to estimate the density. However, this estimator is not necessarily unbiased for non-uniform distributions. If the density is location-dependent, the integration volume dV must be small enough to accurately sample the density gradient everywhere in the distribution. This choice of dV can be impractically small for distributions with high density contrast. One solution is to choose dV adaptively based on the local density of particles (Figure 2-2), using smaller boxes in higher-density regions, but this method introduces bias because the box dimensions, like r_{Ns} , are then subject to Poisson statistics. We chose $\hat{\Gamma}_n$ (Appendix A, Equation (A.6)) to isolate the contribution from choosing dV adaptively.

Furthermore, the simple nearest-neighbor estimator is not the only option. Many other density estimation algorithms exist, often optimized for much better performance. However, even for these algorithms it is usually the case that $E(\hat{n})^2 \neq E(\hat{n}^2)$, especially since the bias and standard deviation are usually optimized for the first moment (the density) at the expense of the higher moments (including density-squared). We have chosen two other methods besides nearest-neighbor, both well-studied as density estimators, to see whether a good density estimator also makes a good rate

estimator ($\hat{\Gamma}_f$, Equation (A.8), and $\hat{\Gamma}_s$, Equation (A.14)). We also include a rate estimator based on one of these density estimators that has been used in the literature to calculate the rate ($\hat{\Gamma}_d$, Equation (A.15)). All the rate estimators are summarized in Table 2.1.

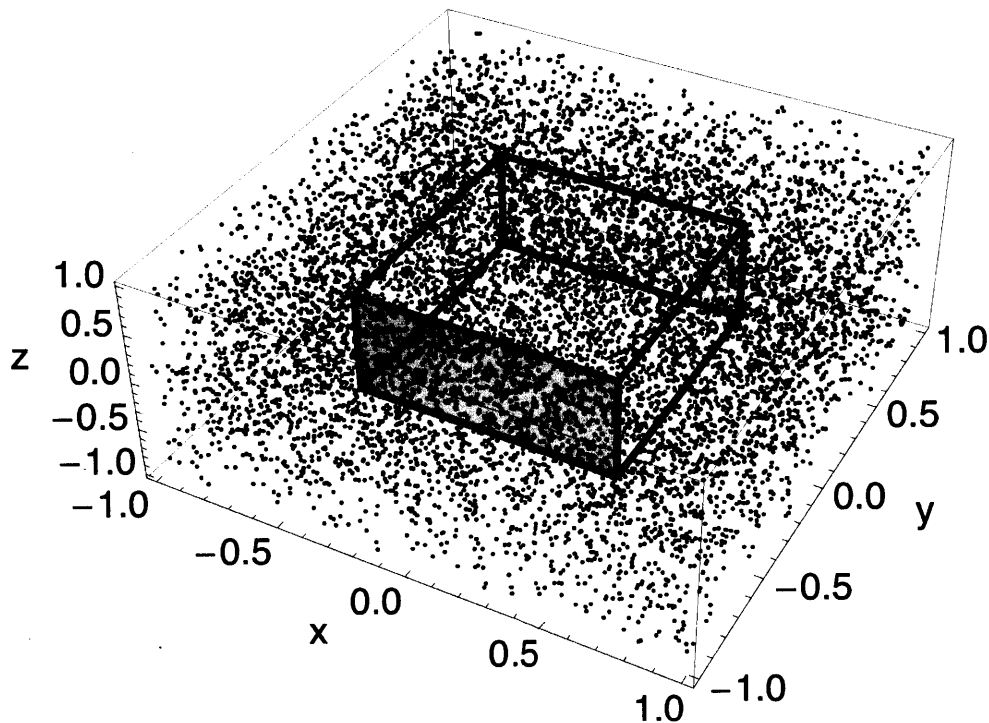


Figure 2-1: A three-dimensional random N-body realization of the uniform distribution. All the points (particles in the realization) in the full volume are used to estimate the density; to avoid edge effects, the volume tessellation and Riemann sum are performed over the volume indicated by the yellow box.

We calculated the expectation value and standard deviation of each rate estimator using random realizations of a uniform distribution, as described above. Equations (2.3) and (2.4) then give the RMS error and bias in terms of these quantities. As expected, the analytically unbiased estimator $\hat{\Gamma}_u$ has a numerically confirmed bias of zero (orange squares in Figure 2-3). Furthermore, using an adaptive Riemann

Estimator	Definition	Long name	See equation(s)
$\hat{\Gamma}_u$	$= dV \sum_{i=1}^{N_V} \hat{n}_{u,i}^2$	Unbiased nearest-neighbors estimator with uniform box size	A.4, A.5
$\hat{\Gamma}_n$	$\propto \sum_{i=1}^{N_p} \hat{n}_{u,i}^2 dV_i$	Unbiased nearest-neighbors estimator with adaptive box size	A.4, A.6
$\hat{\Gamma}_f$	$\propto \sum_{i=1}^{N_p} (\hat{n}_{f,i})^2 dV_i$	FiEstAS estimator [92] with adaptive box size	A.7, A.8
$\hat{\Gamma}_s$	$\propto \sum_{i=1}^{N_p} (\hat{n}_s(r_i))^2 dV_i$	Epanechnikov kernel density estimator with adaptive box size	A.9, A.13, A.14
$\hat{\Gamma}_d$	$\propto \sum_{i=1}^{N_p} \hat{n}_s(r_i)$	Method from [58]	A.15

Table 2.1: Various estimators of the gamma-ray emissivity Γ . See Appendix A for longer descriptions of the various estimators.

volume does change the bias (green diamonds in Figure 2-3). The more complicated estimators require even more substantial correction (cyan triangles, blue pentagrams, and purple hexagrams in Figure 2-3). Interestingly, all the bias curves have the same overall shape. We use the multiplicative factor that transforms the biased density-squared estimator into the unbiased one,

$$\frac{E[(\hat{n}_b)^2]}{E(\hat{n}_u^2)} = \frac{N_s^2}{(N_s - 1)(N_s - 2)}, \quad (2.6)$$

generalized to the form

$$b(N_s) + 1 = \frac{b_1 N_s^2}{(N_s - b_2)(N_s - b_3)}, \quad (2.7)$$

as an ansatz for the shape of each bias curve. This parameterization evokes the introduction of an effective N_s , but can also adjust for the statistical effect of choosing adaptive Riemann volumes, and is a good fit to all the bias curves (Figure 2-3).

Fitting $E(\hat{\Gamma}_n)/\Gamma_{\text{true}}$ to a function of the form in Equation (2.7) determined b_n , the bias from using adaptive Riemann volumes. The results for $\hat{\Gamma}_f$, $\hat{\Gamma}_s$, and $\hat{\Gamma}_d$ were fit using the same function to determine b_f , b_s , and b_d respectively, to identify bias from using a particular density estimation method. The fitted parameters for each estimator are summarized in Table 2.2. We used the fits to correct for the Poisson

Estimator	b_1	b_2	b_3
u	1	2	1
n	1.0001 ± 0.0004	2.81 ± 0.01	-2.99 ± 0.03
f	0.991 ± 0.004	1.41 ± 0.03	1.41 ± 0.03
s	0.971 ± 0.003	-0.3 ± 0.2	4.1 ± 0.1
d	1.00024 ± 0.00005	0.60 ± 0.01	1.89 ± 0.01

Table 2.2: Best-fit bias corrections. Error ranges indicate the 95-percent confidence level. The form of the bias correction is given in Equation (2.7). See Appendix A and Table 2.1 for descriptions of the various estimators.

bias in all the remaining work discussed in this paper.

2.2.2 Performance of Estimators on the Uniform Distribution

All the estimators can be corrected to measure Γ_{true} for a uniform density with an accuracy of better than one percent, although the f and s estimators have higher-order uncorrected behavior (Figure 2-4). The typical standard deviation of any of the estimators is larger than this residual bias by about an order of magnitude, so it will dominate the RMS error (compare Figure 2-5 to Figure 2-4). We want the estimator with the best combination of small RMS error and small N_s : although the value of N_s is not as important in the uniform-density case, it limits the sensitivity of the estimator to small-scale density fluctuations if the density is not uniform. We will discuss this further in Section 2.2.3.

It is not clear that reducing the bias will automatically reduce the RMS error, since correcting for the bias can change the standard deviation of an estimator. In our case, for a given N_s , the bias correction simply multiplies the uncorrected estimator by a constant factor, so the expectation values are related by

$$E(\hat{\Gamma}_{\text{corr}}) = E(\hat{\Gamma}_{\text{uncorr}})/(b(N_s) + 1) \quad (2.8)$$

and the standard deviations are related by the same factor:

$$\text{std}(\hat{\Gamma}_{\text{corr}}) = \text{std}(\hat{\Gamma}_{\text{uncorr}}/(b(N_s) + 1)) = \text{std}(\hat{\Gamma}_{\text{uncorr}})/(b(N_s) + 1) \quad (2.9)$$

From this last expression, we see that if the uncorrected estimator *underestimates* Γ_{true} , correcting it will increase the standard deviation. However, all the uncorrected estimators overestimate Γ_{true} (Figure 2-3), so correcting the bias will also reduce the standard deviation. $b(N_s) + 1$ is always close to unity and never larger than 2, so the change to the standard deviation is slight.

For a given N_s , the RMS error of the nearest-neighbor-style estimators is generally smaller than that of the kernel-based estimators and decreases faster with N_s (Figure 2-5). In a kernel-based estimator, because each particle within the smoothing radius is individually weighted, the estimator must know the location of every one of the N_s particles used in the density estimate, not just the N_s th one, and each weight is less than 1. N_s in the kernel-based estimators must therefore be much larger to get the same RMS error as in a nearest-neighbors estimator with a given N_s , as can be clearly seen in Figure 2-5. To achieve the same RMS error as $\hat{\Gamma}_u$ at $N_s = 10$, $\hat{\Gamma}_s$ must use $N_s \sim 25$. The RMS error of $\hat{\Gamma}_d$ converges so slowly that although it starts out with slightly better performance than $\hat{\Gamma}_n$, it only begins to compete with $\hat{\Gamma}_u$ at $N_s > 30$. Using such a large N_s is a built-in disadvantage for these estimators if one hopes to retain sensitivity to small-scale density fluctuations, and also significantly increases the computational load. For this reason we decided not to test the kernel-based estimators on systems with high density contrast, since the nearest-neighbors estimators are better suited to our needs.

Using an adaptive Riemann volume with the spherical nearest-neighbor estimator, as is done in $\hat{\Gamma}_n$, increases the RMS error at low N_s . At the same time, $\hat{\Gamma}_f$ uses the same adaptive Riemann volume and achieves a lower RMS error. This unusual behavior may be caused by the different shapes of the density estimation volume (spherical) and Riemann volume (orthohedral) used in $\hat{\Gamma}_n$; we discuss this possibility in Section 2.2.4. The RMS error of $\hat{\Gamma}_n$ may be smaller relative to that of the other

estimators in the case of high density contrast because the adaptive Riemann volumes can resolve the density gradient better, so we tested it, along with $\hat{\Gamma}_u$ and $\hat{\Gamma}_f$, on samples with high density contrast. These tests are described in the next few sections.

2.2.3 Additional Bias for Systems with High Density Contrast

In regions of high density contrast, like caustics, the maximum resolvable density is limited by the minimum nearest-neighbor distance expected for a given smoothing number N_s and number of particles N_p . In the limit that N_s and N_p are both large, the expectation value of the minimum nearest-neighbor distance scales as

$$E(\hat{r}_{N_s, \min}) \propto \left(\frac{N_s}{N_p} \right)^{1/3}, \quad N_s, N_p \gg 1. \quad (2.10)$$

This scaling is derived by calculating the first order statistic of the probability distribution of nearest-neighbor distances for a three-dimensional Poisson point process (for a longer explanation, please see Appendix B). The scaling with N_p is valid for $N_p \gtrsim 10^2$, but the scaling with N_s only approaches the asymptotic limit for values much too large to be practical ($N_s \gtrsim 10^3$). For reasonable values of N_s , the power-law index must be determined numerically as discussed in Appendix B:

$$E(\hat{r}_{N_s, \min}) \propto \frac{N_s^\gamma}{N_p^{1/3}}, \quad \gamma = 0.51 \pm 0.06 \quad N_p > 10^2, \quad 10 \lesssim N_s \lesssim 45 \quad (2.11)$$

The corresponding maximum density then scales as

$$n_{\max} \propto \frac{N_s}{r_{N_s, \min}^3} \propto \frac{N_p}{N_s^{3\gamma-1}}. \quad (2.12)$$

For N_s in the range of interest, $3\gamma - 1 \approx 1/2$. As expected, using more particles or a smaller smoothing number increases the sensitivity to small-scale fluctuations and the maximum density. The upper limit on the density introduces bias into the density estimation that also depends on N_p and N_s in the combination given by Equation

(2.12). The estimator will perform normally as long as the local density is less than n_{\max} , but returns n_{\max} for densities larger than n_{\max} . Given a density estimator \hat{n} , the undersampling-limited density estimator \hat{n}_{ul} that incorporates this effect can be written

$$\hat{n}_{\text{ul}} = \begin{cases} \hat{n} & E(\hat{n}) < n_{\max} \\ n_{\max} & E(\hat{n}) > n_{\max} \end{cases} \quad (2.13)$$

The upper limit on the density changes the way the corresponding rate estimator works, since now the piecewise function \hat{n}_{ul} separates regions of the Riemann sum where the density is less than the upper limit from regions where the density is too high to be resolved. So given a bias-free uniform-density estimator $\hat{\Gamma}$, the corresponding density-limited estimator $\hat{\Gamma}_{\text{dl}}$ is

$$\hat{\Gamma}_{\text{dl}} = \hat{\Gamma}_i \Big|_{n < n_{\max}} + \frac{n_{\max}^2 V_{n > n_{\max}}}{b + 1} \quad (2.14)$$

where b is a factor with the form of Equation (2.7) and the appropriate fitted constants from Table 2.2. There is therefore an undersampling bias, b_{us} , in the rate estimator that depends on both N_p and N_s . The N_p - and N_s -dependence enter two ways: in the criterion for separating the Riemann sum and directly in the rate calculation for one of the terms:

$$\begin{aligned} b_{us} &\equiv \frac{\hat{\Gamma}_{\text{dl}}}{\Gamma_{\text{true}}} - 1 \\ &= \frac{E(\hat{\Gamma}) \Big|_{n < n_{\max}}}{\Gamma_{\text{true}}} + \frac{n_{\max}^2 V_{n > n_{\max}}}{\Gamma_{\text{true}}(b + 1)} - 1, \end{aligned} \quad (2.15)$$

Each of the first two terms is less than or equal to 1 because they are both evaluated over subsets of the full integration volume. Additionally, their sum must be less than or equal to 1 because the limiting density is less than or equal to the density in all the Riemann volumes in that sum. So $b_{us} \leq 0$. In the limit where the realization is fully resolved, the bias should be zero since $V_{n > n_{\max}} = 0$, independent of N_s and N_p .

In general, the N_p - and N_s -dependence in Equation (2.15) is complicated since

some unknown fraction of the total volume is under-resolved. To determine the undersampling bias, we tested $\hat{\Gamma}_u$, $\hat{\Gamma}_n$, and $\hat{\Gamma}_f$, including their respective corrections for Poisson bias, on N-body realizations of a one-dimensional caustic for which Γ_{true} can be calculated analytically. The caustic has an adjustable sharpness represented in terms of a velocity dispersion σ : the smaller σ is, the narrower and taller the peak in the density. The caustic density as a function of position and time may be expressed in terms of Bessel functions:

$$\begin{aligned} \rho(x, t) &= \frac{\rho_0}{\sqrt{2\pi\sigma^2 t^2}} \sqrt{\frac{|x - x_c|}{t}} e^{-(x-x_c)^2/4\sigma^2 t^2} \\ &\times \mathcal{B} \left[\frac{(x - x_c)^2}{4\sigma^2 t^2} \right] \end{aligned} \quad (2.16)$$

with

$$\mathcal{B}(u) = \begin{cases} \frac{\pi}{\sqrt{2}} [\mathcal{I}_{-1/4}(u) + \mathcal{I}_{1/4}(u)] & x \leq x_c \\ \frac{\pi}{\sqrt{2}} [\mathcal{I}_{-1/4}(u) - \mathcal{I}_{1/4}(u)] & x > x_c \end{cases} \quad (2.17)$$

where $x_c = 1/4\alpha t$ is the position of the caustic and \mathcal{I} is a modified Bessel function of the first kind. Table 2.3 briefly explains the parameters ρ_0 , α , and σ and the dimensions for the integration volume, and gives the values used in our tests where applicable. We derive this result and explain the parameters more fully in Chapter 3. Equation (2.16) is given in terms of the mass density ρ , which is easily related to the number density n . When we construct random realizations of the caustic, we hold the normalization of the mass density ρ_0 and the size of the integration volume V constant as N_p changes by setting the particle mass $m_p = \rho_0 V / N_p$, so that realizations with different N_p will have the same Γ_{true} .

Given the values in Table 2.3, Γ_{true} can be calculated by performing a single numerical integral (Chapter 3). We can use this simple model to vary the density contrast and the scale of the density variations, simply by generating N-body representations of the caustic described by Equation (2.16) for different σ .

To make a random N-body realization of the caustic, N_p particles are initially distributed uniformly in their initial three-dimensional positions \vec{q} . Then a displacement function $x(q_x)$ is applied to the q_x coordinate to generate the sample (the set of blue

points in Figure 2-6). The general form of $x(q_x)$ is given in Equation (3.6). The form for a particular sample is determined by setting the parameters α and t , which also set the location of the peak density of the caustic, x_c , and by setting σ , the width of the normal distribution from which the random components of the particles' initial velocities are drawn. The values of α , t , and σ we used are summarized in Table 2.3. The locations of the particles in y and z remain uniform.

It is important to choose the initial range of q_x so that the corresponding range of x values defined by the mapping $x(q_x)$ is larger than the integration range in x , since otherwise the system will be incompletely sampled in x . In practice, we determined the range in q_x by choosing a range in x that is larger than the integration range and then inversely mapping it to q_x under the assumption that the random velocity contribution is zero (otherwise the mapping is not invertible). This method does not work for random velocities comparable in magnitude to the bulk velocity, and we adjusted our method of calculating Γ_{true} to account for the incomplete sampling in these cases, as noted in Chapter 3.

To test the estimators, we used them to calculate the rate for a set of samples with a given density contrast and resolution by integrating density-squared over the integration volume (shown as a yellow box in Figure 2-6). The integration volume is smaller than the dimensions of the realization to avoid unwanted edge effects, but extends well beyond the edge of the caustic, which is the feature of interest. Since the undersampling bias depends on both the resolution and smoothing number, we varied both these parameters for each set of samples with a given σ . Table 2.4 summarizes the ranges and step sizes we used to explore this parameter space. Because the parameter space for these tests was so much larger than in the uniform-density case, we used 5000 random realizations at each combination of contrast and resolution, so the expected level of sampling fluctuations is about 1.5 percent.

2.2.4 Understanding the Undersampling Bias

We draw conclusions about the behavior of the undersampling bias with various N_s and N_p using the results of our tests at various levels of density contrast, represented

Parameter	Value	Dimension	Notes
ρ_0	625	$[M]/[L]^3$	Mass density of the sample at $t = 0$. Particle mass is adjusted to keep the same mass density at varying resolution.
σ	varies	$[L]/[T]$	Sharpness parameter. A smaller σ makes a sharper caustic.
α	1/2	$1/[L][T]$	Describes the displacement function used to generate the caustic. See Chapter 3, Equations (3.6) and (3.20).
t	1	$[T]$	Corresponds to $x_c = 0.5$. See Chapter 3, Equation (3.25).
L_y, L_z	0.5	$[L]$	The rate is integrated from $-L_{y,z}$ to $L_{y,z}$, where y and z are the dimensions parallel to the caustic. Avoids edge effects, as illustrated in Figure 2-6.
$[x_-, x_+]$	[-0.5, 2]	$[L]$	Limits of the rate integration in the direction perpendicular to the caustic, chosen so that the rate is integrated across the caustic. Shown in Figure 2-6.

Table 2.3: Parameters for the analytic one-dimensional caustic. These quantities are defined in more detail in Chapter 3. The units are given as dimensions only since they may be scaled as needed.

Parameter	Values Tested
Estimator	{u,n,f}, corrected for Poisson bias where appropriate using Equation (2.7) and the appropriate coefficients in Table 2.2.
N_s	10...30, in steps of 2
$\log_{10} N_p$	3...4.5, in steps of 0.25
$\log_{10} \sigma$	-3...0.5, in steps of 0.25

Table 2.4: Parameter space for testing undersampling bias.

in our model caustic by the parameter σ .

As is expected, using higher resolution (a larger N_p) leads to better rate estimates of sharper features (Figure 2-7). Even the highest-resolution realizations we tested could only resolve moderately sharp features. The estimators using adaptive Riemann volumes (for example, $\hat{\Gamma}_f$ shown in the right panel of Figure 2-7) appear to require more particles to obtain the same bias as the constant-Riemann-volume case; this is partly due to the fact that in the adaptive scheme there is exactly one volume per particle, but the same number of constant Riemann volumes (about 10^4) is used regardless of resolution.

We found that the undersampling bias depends strongly on N_s when the features of interest are marginally or under-resolved, and only weakly when they are fully resolved (Figure 2-8, left panel). As expected, using a smaller N_s leads to a lower bias because the size of the volume used for density estimation scales with N_s as shown in Equation (2.11), so that a larger N_s blurs out more small-scale features. A lower N_s also leads to a larger standard deviation of the density estimates, which increases the total RMS error, but this effect is small compared to the improvement in the bias in the under-resolved regime and only very slightly affects the fully-resolved case (Figure 2-8, right panel).

We found that the algorithm used to determine the Riemann volumes adaptively was sensitive to the way in which the boundaries of the integration volume were treated. Boundaries closest to the sharp edge of the caustic (which in our test cases is parallel to one face of the integration volume) must be trimmed as described in Section 2.2 of [92] to avoid artificially overestimating the rate: without trimming, the Riemann volumes on the face of the caustic next to the boundary are artificially elongated into the region ahead of the sharp edge where the density is effectively zero. The density estimate in those volumes will then be artificially high because the density is assumed to be constant over the entire Riemann volume, leading to rate estimates with positive bias when the distribution is marginally resolved (Figure 2-9, solid [blue] lines). Such "bleed-over" still occurs with a constant Riemann volume but is much less significant because the box size does not depend on the local density.

However, applying the trimming algorithm uniformly to all the Riemann volumes with one or more faces on a boundary of the integration volume artificially *underestimates* the rate by significantly reducing the total integration volume on the trailing edge of the caustic, where the density is small but still nonzero (Figure 2-9, dashed [cyan] lines). In this system, restricting the trimming only to the face nearest the caustic edge results in the correct bias behavior (Figure 2-9, dot-dashed [green] lines). Because the choice of how to treat the boundaries appears to depend on the particular geometry of the system in question, it may be difficult to extrapolate the performance of estimators that use the adaptive Riemann volumes from our test case to systems with arbitrary geometry. In particular, it is not immediately clear how this method would extend to the shells in the dynamical model of M31, which have spherical edges near several boundaries of the integration volume.

2.2.5 Performance of Estimators on the Non-uniform Distribution

The best estimator; that is, the one with the smallest RMS error, has the best combination of undersampling bias near zero and small standard deviation for the smallest σ . In the resolution-limited regime the RMS error is dominated by the bias. If the caustic is fully resolved, the bias is zero and the standard deviation, which is constant for a given N_p , dominates the RMS error. We have already established that a smaller N_s improves performance in the under-resolved regime without substantially increasing the RMS error for resolved distributions. Figure 2-10 shows that all three estimators achieve the convergence rate of $N_p^{-1/2}$ predicted by [77] and [82] for sharpnesses that are fully resolved at the highest resolution.

Comparing the RMS errors for sharper and sharper caustics shows that all three tested estimators have very similar performance (Figure 2-11). The nearest-neighbors estimator with constant Riemann volume ([red] circles in Figure 2-11) converges slightly faster than the other two but once the distribution is resolved they are nearly indistinguishable from one another (Figure 2-11, inset). The FiEstAS method, thanks

to the space-filling tree organizing the particles, is faster than the nearest-neighbors estimators, so if the distribution is known to be completely resolved (the regime shown in the inset of Figure 2-11), then this method can be used without loss of performance to take advantage of its greater speed. However, in situations where parts of the distribution may be under-resolved and creating a higher-resolution realization is not possible, $\hat{\Gamma}_u$ should be used to take advantage of its ability to resolve slightly sharper features with fewer particles.

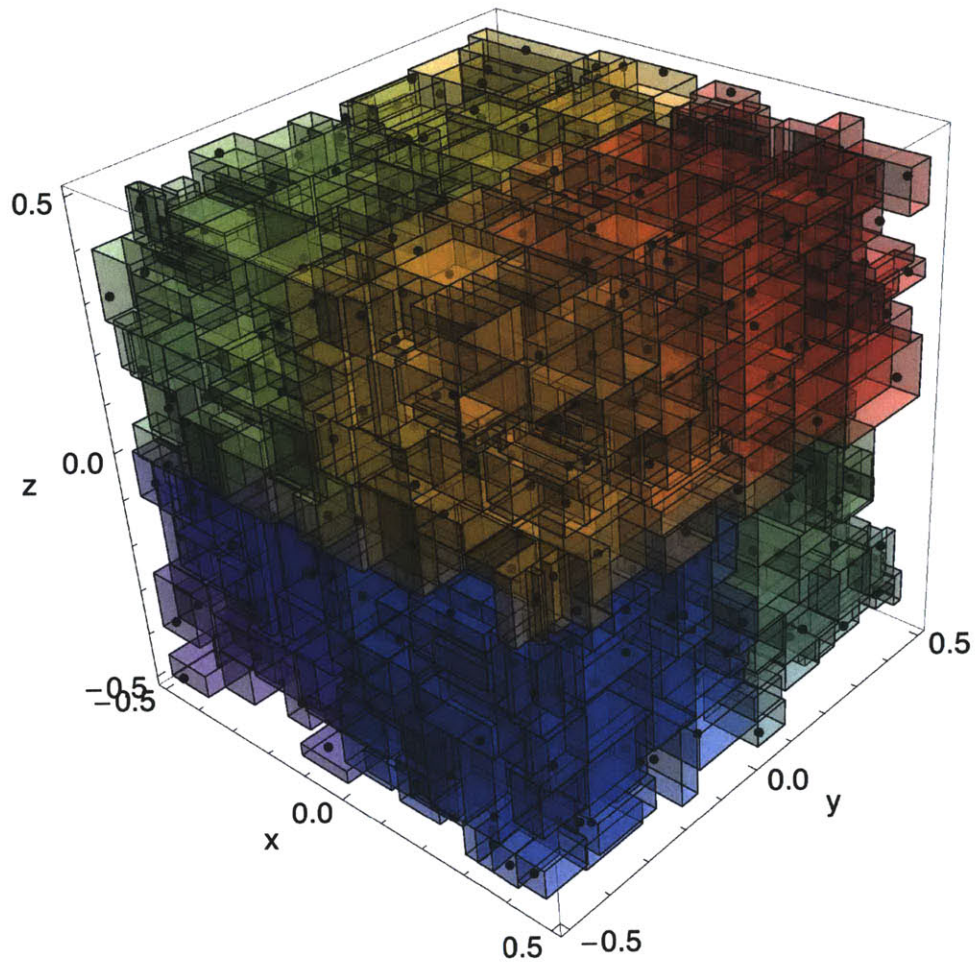


Figure 2-2: Example of adaptive volume tessellation referred to in the text, produced by the space-filling tree in the FiEstAS estimator. The colors, from red to blue, indicate the way in which the tree divides the space (the spatial distribution of the inorder traversal). For clarity, not all the boxes with edges on the boundaries are shown.

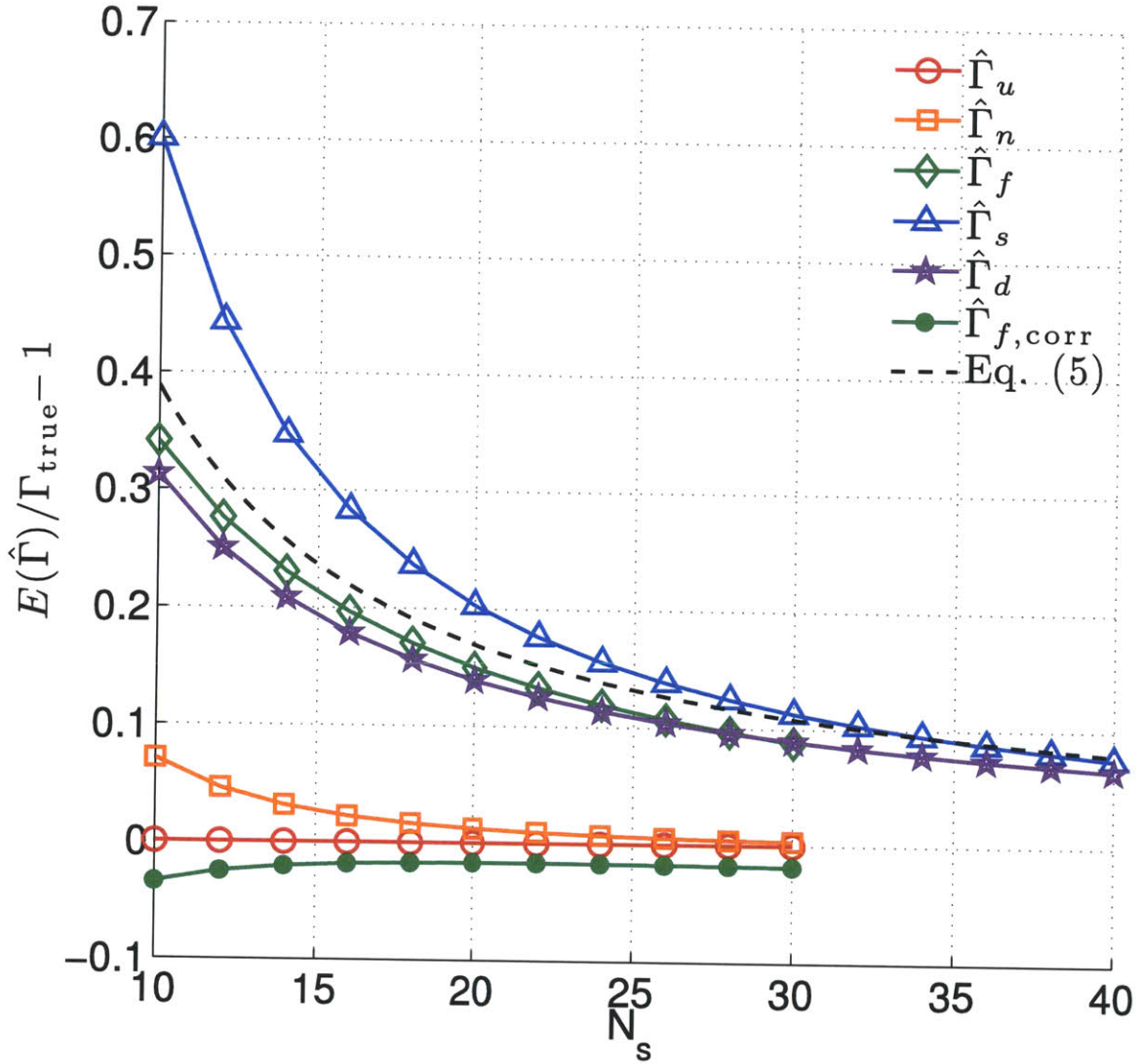


Figure 2-3: All the rate estimators except for $\hat{\Gamma}_u$ (red circles) require slightly different but similarly shaped bias corrections. Equation (2.7) is an adequate ansatz for the N_s -dependence of the correction: an example corresponding to Equation (2.6) is shown as the black dashed line and the green filled points show the effect of using Equation (2.6) to correct $\hat{\Gamma}_f$. Fit parameters are listed in Table 2.2; Table 2.1 explains the label abbreviations.

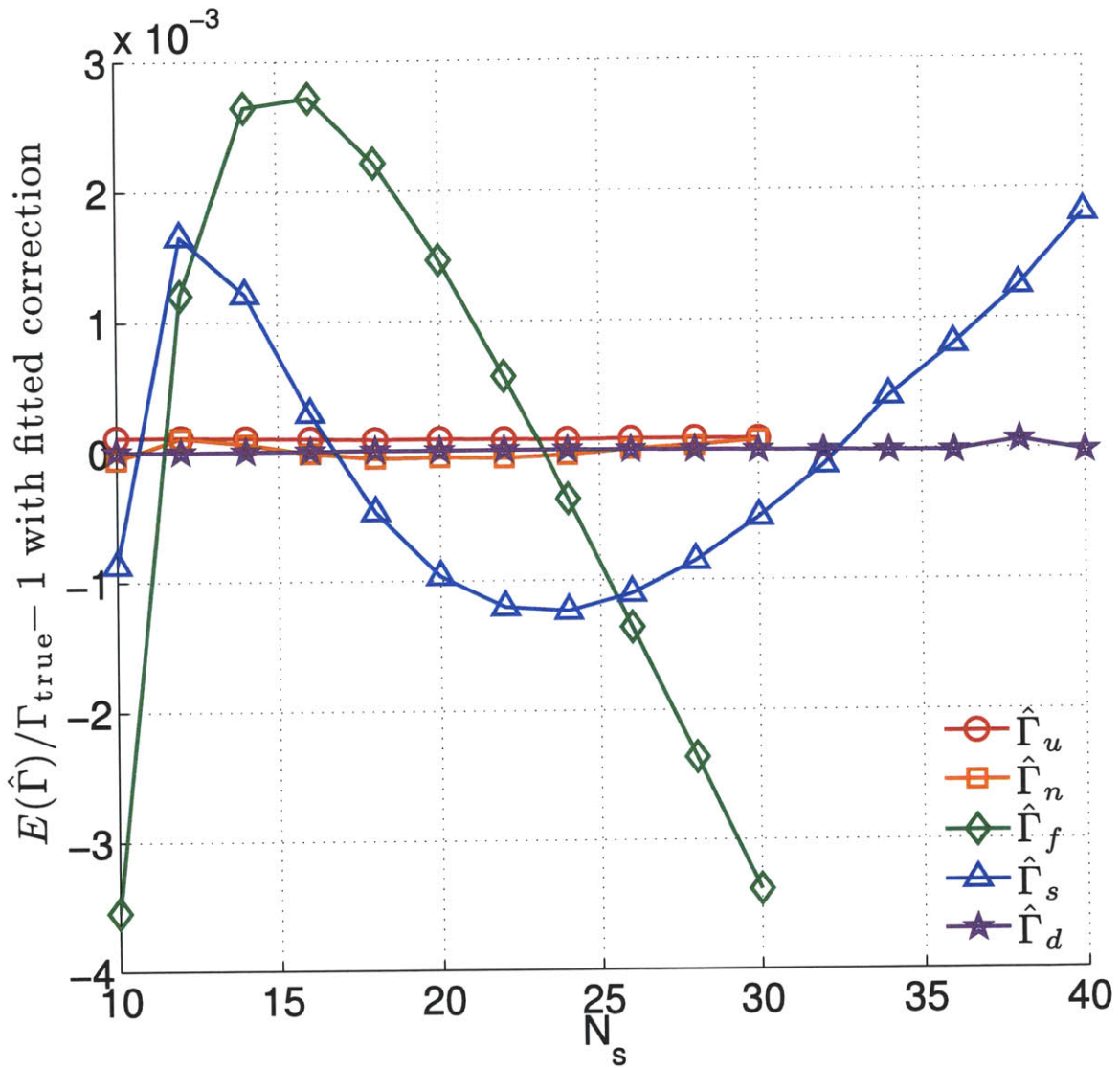


Figure 2-4: Once estimators are corrected for Poisson bias using the fitting formula in Equation (2.7), the corrected versions all have a bias much smaller than the typical standard deviation of a few percent. The residual bias is comparable to that of $\hat{\Gamma}_u$. Table 2.1 explains the label abbreviations.

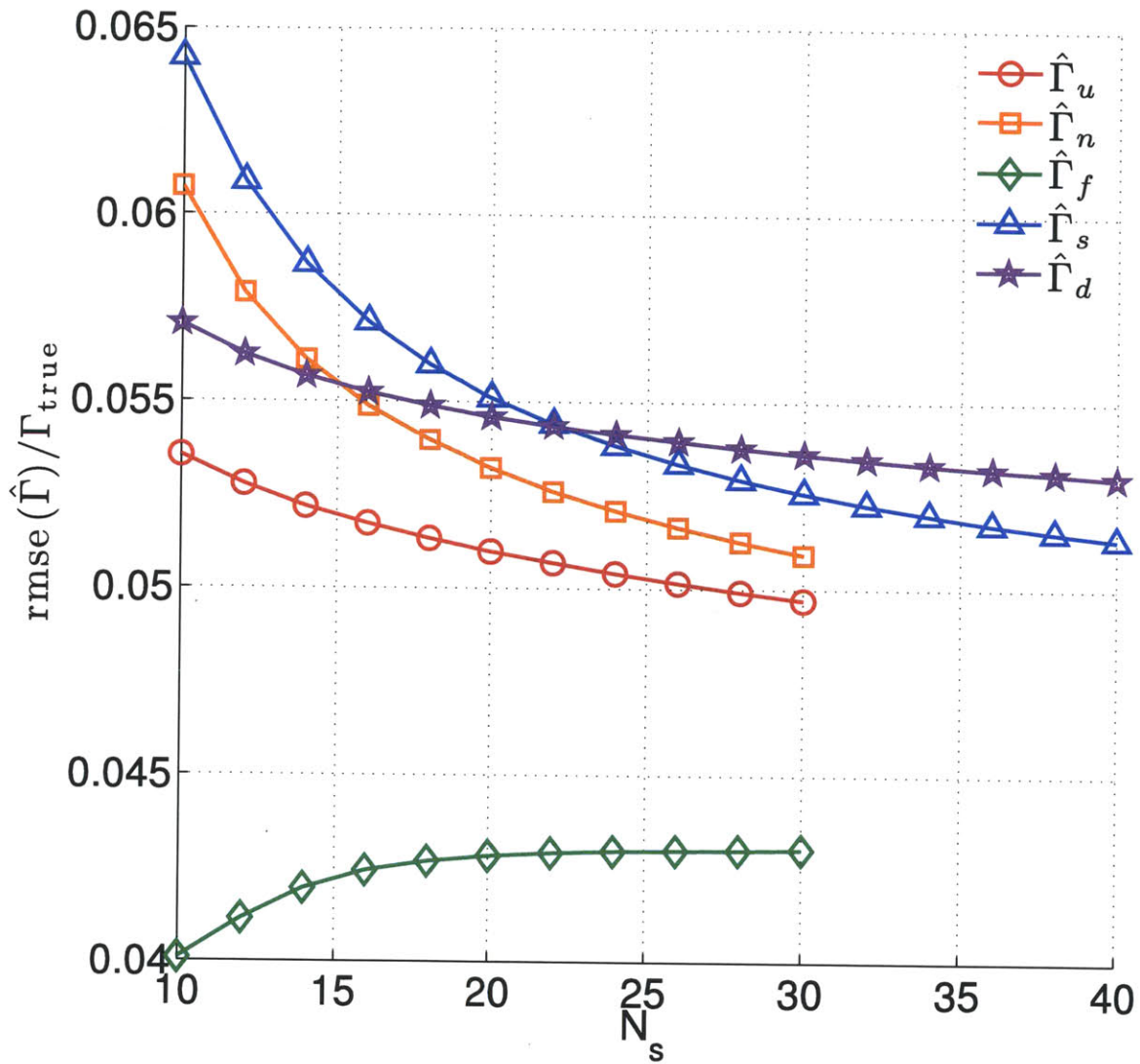


Figure 2-5: The RMS error of the corrected estimators is around five percent, and declines with increasing N_s . $\hat{\Gamma}_f$ (green diamonds) and $\hat{\Gamma}_u$ (red circles) have the smallest RMS error at low N_s . Introducing an adaptive Riemann volume (orange squares) increases the RMS error significantly at low N_s . The RMS error starts at a higher value and scales more slowly with N_s for the kernel-based estimators (blue triangles and purple stars) than for the nearest-neighbors estimators, an effect of the weighting function used to estimate the density. Table 2.1 explains the label abbreviations.

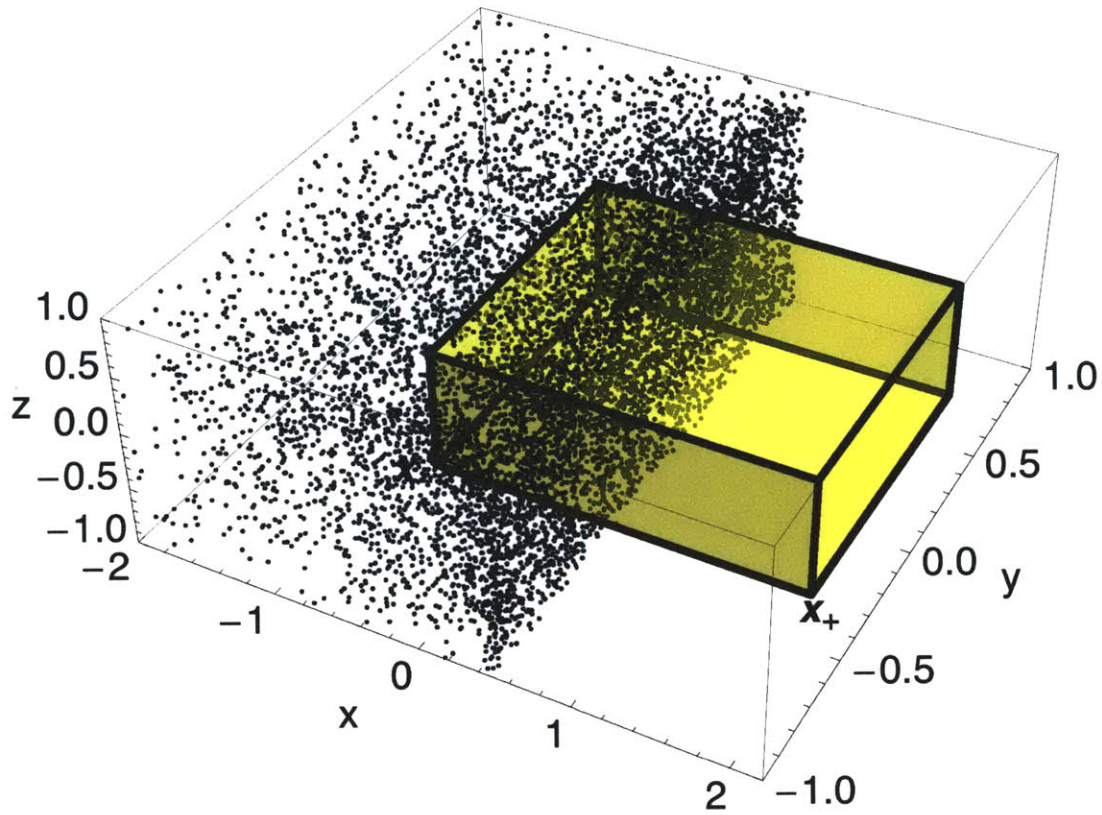


Figure 2-6: An example of a random N-body realization of the analytic one-dimensional caustic used to test estimators for undersampling bias. The points are the locations of the particles in the realization. The yellow box indicates the integration volume, positioned to include the complete caustic shape and avoid edge effects. For this sample, $N_p = 10^4$, $\sigma = 0.01$, and the dimensions of the box are given in Table 2.3.

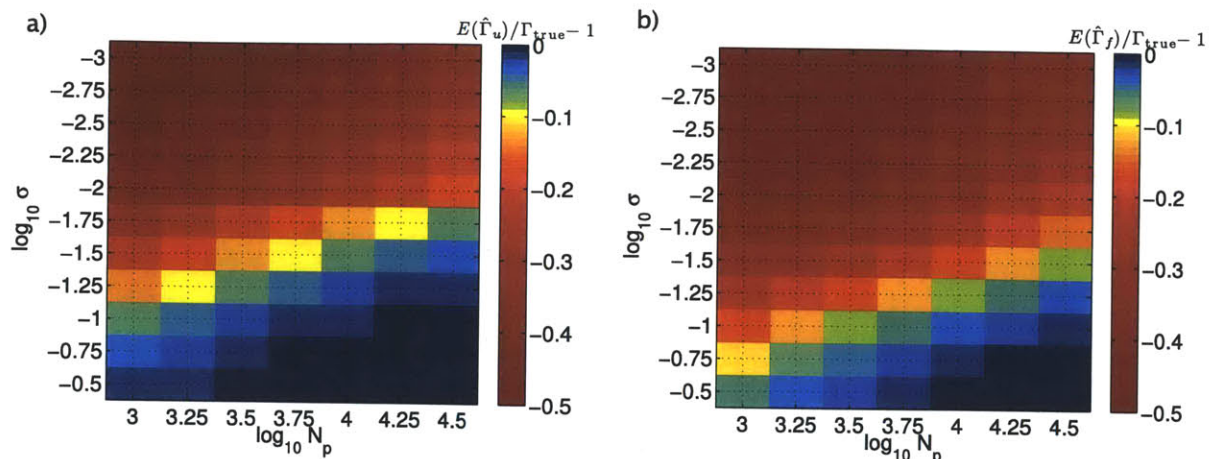


Figure 2-7: Using a larger number of particles resolves sharper caustics in all three cases: $\hat{\Gamma}_u$ (a), $\hat{\Gamma}_f$ (b), and $\hat{\Gamma}_n$ (not shown, but very similar to the two others). Regions colored dark or bright blue are considered fully resolved because the bias is less than the typical standard deviation of about 2 percent. The different colors indicate the magnitude of the undersampling bias b_{us} : blue indicates $|b_{us}| < 0.02$, aqua $|b_{us}| \sim 0.05$, yellow $|b_{us}| \sim 0.1$, and red $|b_{us}| > 0.2$.

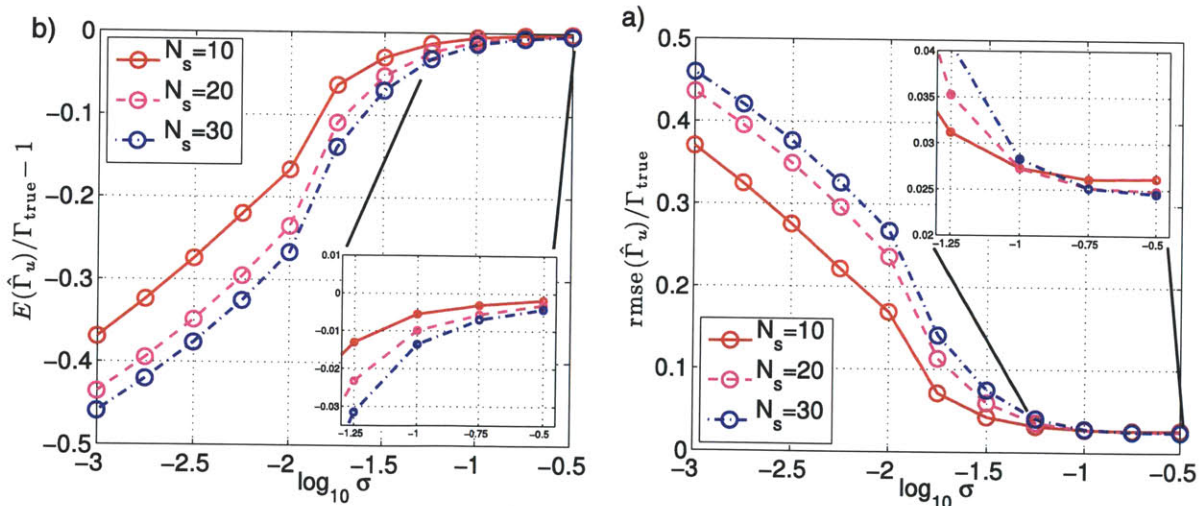


Figure 2-8: Using a lower value of N_s leads to lower bias whether or not the realization fully resolves all the features in the underlying distribution (a, main figure and inset). The increase in the standard deviation, and therefore RMS error, from using a smaller N_s is small compared to the improvement in the bias when there are under-resolved features (b). If the realization is fully resolved, the increase in RMS error is detectable but extremely small (b, inset).

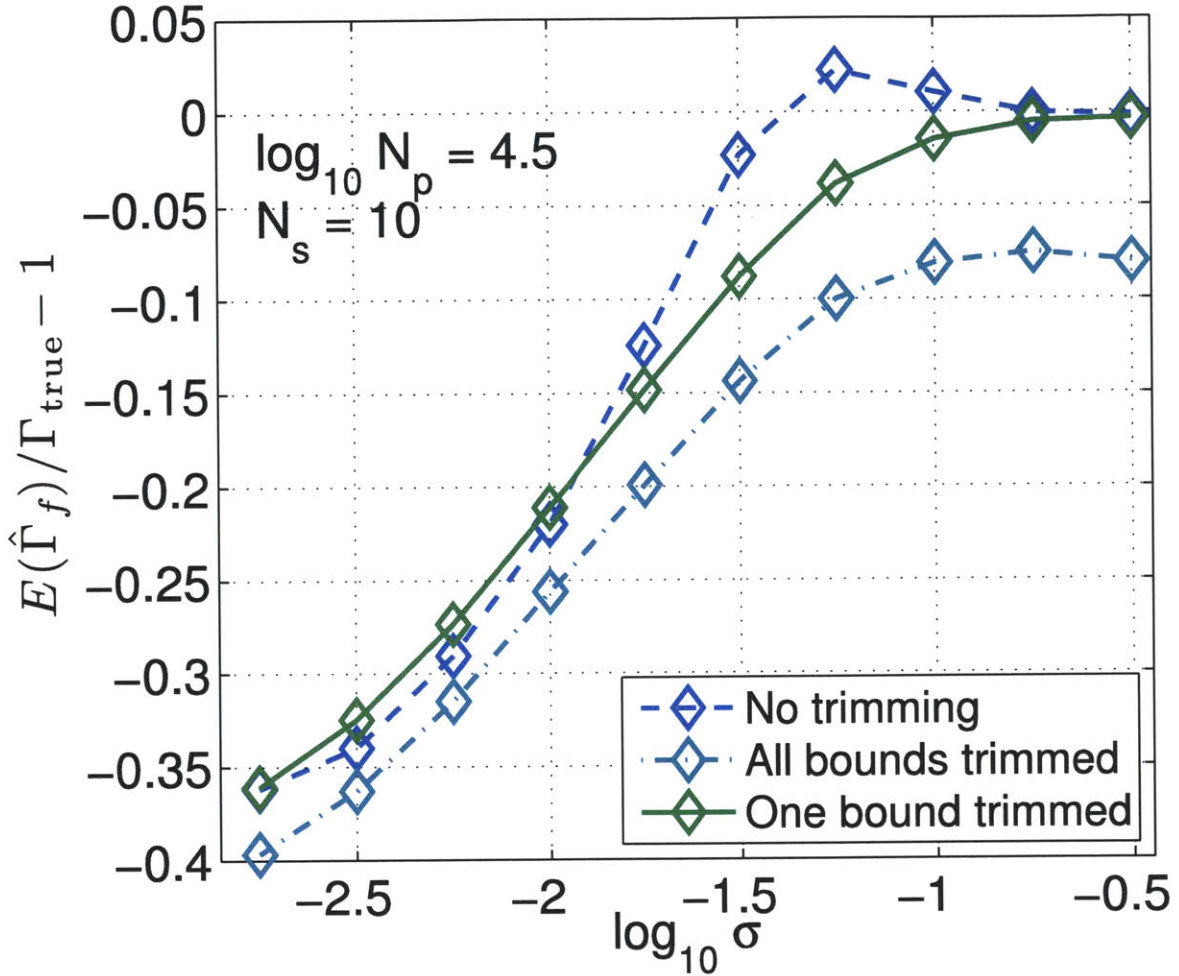


Figure 2-9: The adaptive Riemann volume algorithm is sensitive to the treatment of the boundaries of the integration volume relative to the geometry of the density distribution. Thanks to the location of the caustic parallel to one boundary of the integration volume, the rate is overestimated (solid blue lines) at marginal resolutions unless the Riemann volumes on that boundary are trimmed to fit the face of the caustic (dot-dashed green lines), but trimming all the boundaries in the same manner underestimates the rate (dashed cyan lines).

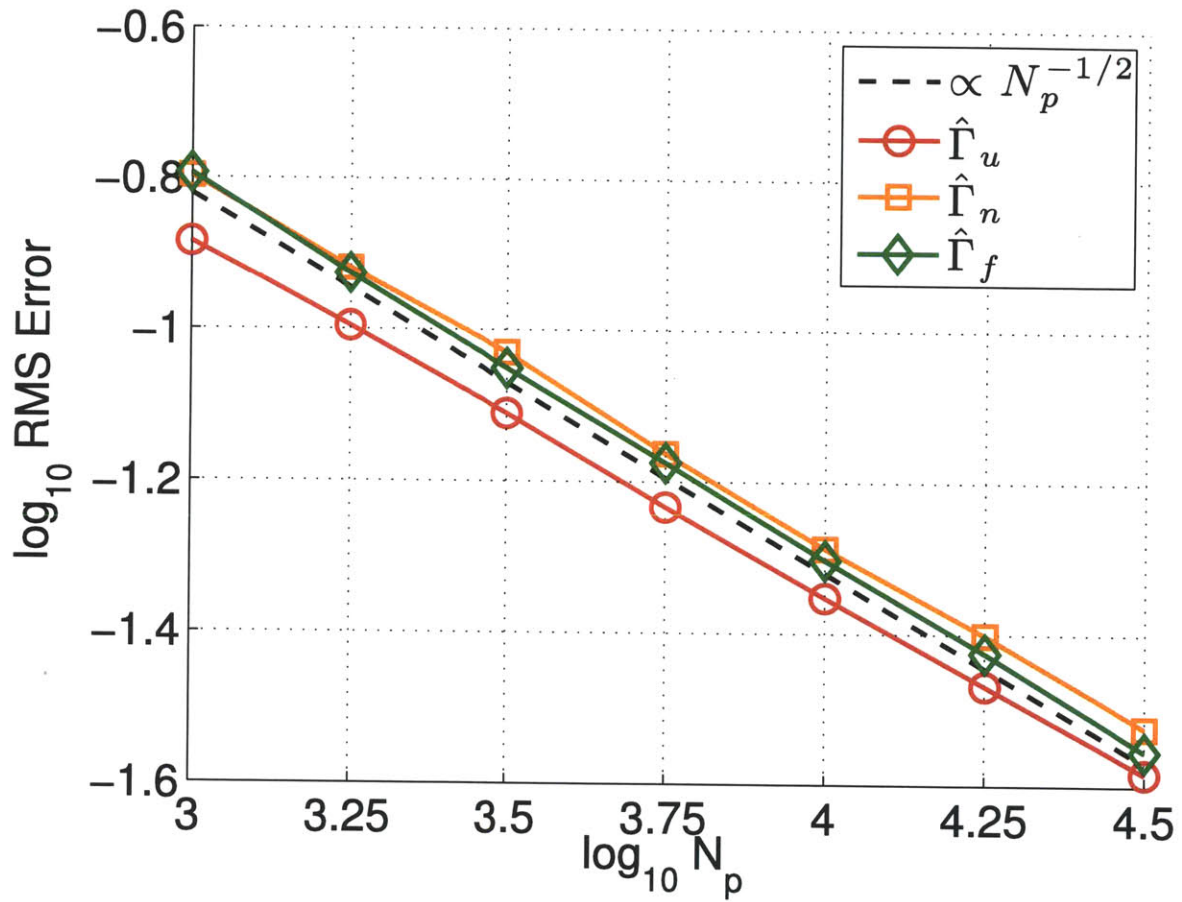


Figure 2-10: All three estimators tested on non-uniform distributions achieve a convergence rate of $N_p^{-1/2}$ (dashed line) as predicted by prior analytical work. Tests with $N_s = 10$ and $\log_{10} \sigma = -0.75$ are shown.

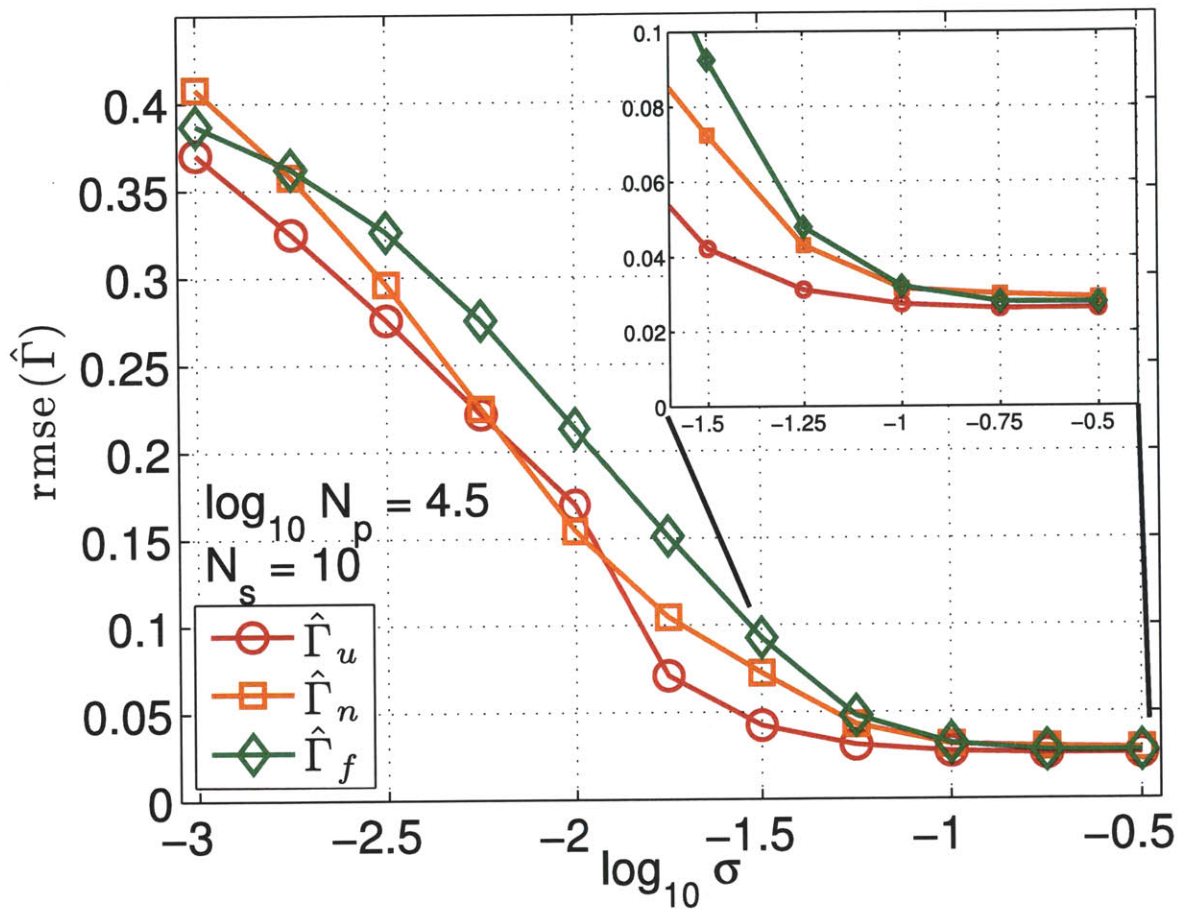


Figure 2-11: All the estimators have very similar performance. $\hat{\Gamma}_u$ (red circles) converges faster in the marginally-resolved regime than the estimators that use adaptive Riemann volumes (orange squares and green diamonds). In the fully resolved regime the RMS errors of the three are nearly identical (inset).

2.3 Calculation of the Boost Factor and Gamma-ray Flux

In this section we describe the N-body model of the M31 tidal debris (Section 2.3.1) and present derivations of formulae for the boost factor and gamma-ray flux (Section 2.3.2) in terms of the numerically estimated rate. We then describe how we used the results of the bias calibrations described in Section 2.2 to correct our numerical estimates of the boost factor and flux (Section 2.3.3), and present the results (Sections 2.3.4 and 2.3.6).

2.3.1 N-body Model

For this work we use the N-body model of the tidal shell system in [32]. The model uses a Plummer sphere as the progenitor of the tidal debris, orbiting in a static, 3-component representation of M31's potential: a spherical halo and bulge, and an axisymmetric exponential disk. To construct a model of the potential, the parameters of the halo, bulge, and disk were first fit to a rotation curve of M31 from several combined sources, not including the tidal shells and the associated tidal stream [see 93, for details]. Then the orbit of the center of mass of the progenitor satellite was fit to the three-dimensional position data and radial velocity measurements available for the stream [33]. Finally, the mass and size of the progenitor were constrained with an N-body model of the stream using the previously determined orbit [32]. Dynamical friction and the response of M31 to the merger are ignored since the mass ratio of the progenitor to M31 is approximately 1/500. [32] emphasize that the N-body model is not the result of a full exploration of this many-dimensional parameter space, but for this work we are only interested in the end configuration of the debris, which acceptably matches the available observations.

We further make the assumption that there is a comparable mass of dark matter associated with the stellar tidal debris. This assumption is probably generous, since the dark matter in galaxies is thought to be much more diffuse than the stellar matter

and much of it will have been stripped away by tides before the progenitor of this tidal debris even reaches the starting point of the simulation. However, for a first estimate we consider this assumption sufficient, though it is not a strict upper limit for reasons we will discuss at the end of this paper.

2.3.2 Formulae

In this section we derive expressions for the boost factor and gamma-ray flux in terms of the estimated rate from the N-body representation.

Boost Factor

To calculate the boost factor, we must assume a halo model because dark matter in the shell can interact with dark matter in the halo. Denoting the halo dark matter with h and the shell dark matter with s , there are three terms in the total rate:

$$\begin{aligned}
\Gamma &= \int n_{\text{tot}}^2 dV \\
&= \int (n_h + n_s)^2 dV \\
&= \int (n_h^2 + 2n_h n_s + n_s^2) dV, \\
&\equiv \Gamma_{hh} + \Gamma_{sh} + \Gamma_{ss}
\end{aligned} \tag{2.18}$$

Equation (2.18) shows that the boost factor β depends on the halo dark matter density:

$$\beta \equiv \frac{\Gamma_{\text{tot}} - \Gamma_{hh}}{\Gamma_{hh}} = \frac{1}{\Gamma_{hh}} \int n_s (2n_h + n_s) dV. \tag{2.19}$$

If $n_h \gtrsim n_s$, the additional emissivity from the tidal debris is dominated by the first term in equation (2.19) and the boost factor scales linearly with the dark matter density in both the halo and the shell.

For the halo, we use the same density distribution that was used in the dynamical

model of the shells [93] with the addition of a small core with size r_{core} of half the size of one Riemann volume. The core eliminates the infinite-density cusp at $r = 0$. This halo is spherically symmetric and of Navarro-Frenk-White form [94, 95]:

$$n_h(r) = \frac{n_{h,0}}{[(r + r_{\text{core}})/r_h][1 + (r + r_{\text{core}})/r_h]^2} \quad (2.20)$$

where $n_{h,0} \equiv \rho_{h,0}/m_p = 3.67 \times 10^3 \text{ kpc}^{-3}$ and $r_h = 7.63 \text{ kpc}$ are determined by [93] by fitting a mass model to a set of measurements of dynamical tracers of M31’s halo. We divide the fitted mass density $\rho_{h,0}$ by m_p , the mass of the simulation particles, to get consistent number densities n_h and n_s .

Gamma-ray Flux

We use the notation of [96, hereafter FPS] to present the calculation of the gamma-ray flux. The differential flux of photons in an infinitesimal band of photon energy E_γ , $d\Phi_\gamma/dE_\gamma$, can be factored into a contribution from “particle physics” that specifies the spectrum of the radiation and an achromatic contribution Φ^{cosmo} from “cosmology”—the shape, size, density, and distance of the dark matter—that sets the normalization, as in eq. 1 of FPS:

$$\frac{d\Phi_\gamma}{dE_\gamma} = \frac{d\Phi^{\text{SUSY}}}{dE_\gamma} \Phi^{\text{cosmo}} \quad (2.21)$$

The rate at which gamma rays would be detected by the Fermi LAT is

$$\frac{dN_\gamma}{dt} = \Phi^{\text{cosmo}} \int_{E_{\text{th}}}^{m_\chi} \frac{d\Phi^{\text{SUSY}}(E_\gamma)}{dE_\gamma} A_{\text{eff}}(E_\gamma) dE_\gamma \quad (2.22)$$

where $A_{\text{eff}}(E_\gamma)$ is the effective area of the detector, E_{th} is the lowest detectable energy, and m_χ is the dark matter mass. For the Fermi LAT, E_{th} is 30 MeV, and above 1 GeV the effective area for diffuse events is roughly independent of energy to within about 10 percent of the mean value over the energy range of integration [74]. Furthermore, the supersymmetric calculations of the particle physics contribution in [10], which we use for this work, take $E_{\text{th}} = 1 \text{ GeV}$. So for the remainder of this work we will use $E_{\text{th}} = 1 \text{ GeV}$ and assume A_{eff} is independent of energy. With this simplification we

can work with the total flux for now and later multiply it by A_{eff} to get the detection rate in the LAT.

Also following FPS, the particle-physics contribution is

$$\frac{d\Phi^{\text{SUSY}}}{dE_\gamma} = \frac{\langle\sigma v\rangle}{2m_\chi^2} \frac{dN_\gamma}{dE_\gamma} \quad (2.23)$$

where $\langle\sigma v\rangle$ is the velocity-averaged cross section and N_γ is the yield, called Q_γ in [57] and FPS. We have moved the constant factor $1/4\pi$ to be part of Φ^{cosmo} because it is most easily understood as part of the attenuation of the gamma-ray flux over distance.

To evaluate N_γ , we use a subset of the benchmark models of [9] to span the space of supersymmetric WIMP candidates. [10] have calculated the gamma-ray yields above 1 GeV for 10 of the 12 models in [9]: models A' through L' with the exception of models E' and F' . We use the values given in the table for the number of photons in the continuum emission times the cross section, $N_{\gamma,\text{cont.}} \langle\sigma v\rangle$, listed in Table 1 of [10]. The continuum emission is not as diagnostic as the line emission at $E_\gamma = m_\chi/2$, but the branching ratio for line emission is smaller by a factor of 10^3 .

Compared to Equations (2.22) and (2.23), we find that

$$\Phi^{\text{SUSY}} = \int_{E_{\text{th}}}^{m_\chi} \frac{d\Phi^{\text{SUSY}}}{dE_\gamma} dE_\gamma = \frac{N_{\gamma,\text{cont.}} \langle\sigma v\rangle}{2m_\chi^2} \quad (2.24)$$

For the models in [10], $N_{\gamma,\text{cont.}} \langle\sigma v\rangle$ is in the range 10^{-29} – 10^{-24} $\text{cm}^3 \text{s}^{-1}$, and m_χ is generally given in GeV. So a useful scaling of this formula in typical units is

$$\begin{aligned} \Phi^{\text{SUSY}} &= 1.54 \times 10^{-8} \text{cm}^4 \text{kpc}^{-1} \text{s}^{-1} \text{GeV}^{-2} \\ &\times \left(\frac{N_{\gamma,\text{cont.}} \langle\sigma v\rangle}{10^{-29} \text{cm}^{-3} \text{s}^{-1}} \right) \left(\frac{m_\chi}{1 \text{GeV}} \right)^{-2}. \end{aligned} \quad (2.25)$$

In addition to the benchmarks, we also use the most optimistic value of Φ^{SUSY} from FPS. According to their Figure 8, $\Phi^{\text{SUSY}} \lesssim 10^{-8}$ for all the models explored,

and the maximum occurs for $m_\chi \approx 40$ GeV or $N_{\gamma,\text{cont.}} \langle \sigma v \rangle \approx 10^{-25}$ cm³ s⁻¹. We use these values to represent the most optimistic estimate of the flux.

The astrophysical factor Φ^{cosmo} depends on the square of the dark matter mass density ρ_χ :

$$\Phi^{\text{cosmo}} = \frac{1}{4\pi d^2} \int_{\text{obj}} dV \rho_\chi^2(x, y, z) \quad (2.26)$$

where x, y, z indicate physical distances in a coordinate system centered on the object, and the integral is over the total volume of the object. The factor $1/4\pi d^2$ accounts for the attenuation in flux over the distance d from the object to the observer. The key element in calculating the astrophysical contribution to the flux is thus determining the integral-density-squared $\int \rho^2 dV$. We calculate this quantity from the simulation results, which consist of the locations of the N_p simulation particles, each with mass m_p . A numerical density estimator gives the number density n_p of the simulation particles as a function of position, which is related to the mass density ρ_χ by mass conservation:

$$\rho_\chi = m_\chi n_\chi = m_p n_p \quad (2.27)$$

so that

$$\int \rho_\chi^2 dV = m_p^2 \int n_p^2 dV \quad (2.28)$$

and

$$\Phi^{\text{cosmo}} = \frac{m_p^2}{4\pi d^2} \int n_p^2 dV \equiv \frac{m_p^2}{4\pi d^2} E(\hat{\Gamma}_u). \quad (2.29)$$

It is useful to rewrite the expression (2.29) with the units and values used in the N-body representation. For the simulation of [32], $m_p = 1.68 \times 10^4 M_\odot$ and $d = 785$ kpc. $E(\hat{\Gamma})$ is calculated in units of kpc⁻³. So a useful version of (2.29) for this work

is

$$\begin{aligned}
\Phi^{\text{cosmo}} &= (1.87 \times 10^{-14} \text{ GeV}^2 \text{ kpc cm}^{-6}) \\
&\times \left(\frac{m_p}{10^4 M_\odot} \right)^2 \left(\frac{d}{785 \text{ kpc}} \right)^{-2} \\
&\times \left(\frac{E(\hat{\Gamma})}{\text{kpc}^{-3}} \right)
\end{aligned} \tag{2.30}$$

The total flux of gamma rays for a given model of dark matter and density distribution, scaled to typical values in the problem, is obtained by combining equations (2.25) and (2.30):

$$\begin{aligned}
\Phi_\gamma &= \Phi_{\gamma,0} \left(\frac{N_\gamma(\sigma v)}{10^{-29} \text{ cm}^3 \text{ s}^{-1}} \right) \\
&\times \left(\frac{m_\chi}{1 \text{ GeV}} \right)^{-2} \left(\frac{m_p}{10^4 M_\odot} \right)^2 \\
&\times \left(\frac{d}{785 \text{ kpc}} \right)^{-2} \left(\frac{E(\hat{\Gamma})}{\text{kpc}^{-3}} \right),
\end{aligned} \tag{2.31}$$

where $\Phi_{\gamma,0} = 2.88 \times 10^{-22} \text{ cm}^{-2} \text{ s}^{-1} = 9.09 \times 10^{-11} \text{ m}^{-2} \text{ yr}^{-1}$. The effective area of the Fermi LAT varies between 0.7–0.85 square meter above 1 GeV [74].

2.3.3 Calibration of the Rate Estimate

To calibrate the result from the density-squared calculation, we must estimate the number of particles in each shell (and thereby N_p) and the σ that gives the best approximation to each caustic's shape. Since the shells are located at the apocenters of the particles' orbits, we can unambiguously determine which particles are in which shell by counting the number of pericenter passages, N_{peri} , that each particle has experienced (Figure 2-12). There are three shells and one tidal stream in the system, each composed of particles with a different N_{peri} . Examining the system in a projection of phase space in the r - v_r plane, shown in Figure 2-13, shows the order in which the shells were formed. The first caustic to form corresponds to the first (outermost)

winding in phase space with the largest apocenter. Particles making up this caustic at a given moment have the lowest N_{peri} because this caustic marks the location of the first turnaround point for bound particles. The most recently formed caustic has the highest number of pericenters and smallest apocenter distance, and has not yet been fully filled: the innermost winding is not yet complete because only a few particles have had time to complete three full orbits. We chose the two oldest shells for analysis because they contain most of the mass. We will refer to the oldest caustic, whose constituent particles are shown in green and have undergone two pericenter passages, as caustic 1 or shell 1. The second caustic, whose constituent particles are shown in red and have undergone three pericenter passages, will be referred to as caustic 2 or shell 2.

Since the satellite galaxy's orbit is nearly radial, the caustics are nearly spherical. However, slight systematic deviations from spherical symmetry, in projection along r , can slightly increase the perceived width of the caustic and cause σ to be overestimated. In order to use our analytical caustic to estimate σ for each shell in M31 and determine the undersampling correction to the result from the density estimator, we had to correct for the slight asphericity in each shell. To do this we determined the caustic radius of each shell in projection along the ϕ direction by binning the particles in r and ϕ and finding the r -bin with the highest number of counts for each slice in ϕ . The bins were chosen as small as possible for resolution while still being able to identify the peak r bin for each ϕ . Once the peak r was obtained as a function of ϕ , we fit this set of points with a polynomial $r_c(\phi)$ and calculated $dr = r - r_c(\phi)$ for each particle to correct for the asphericity of the shell in ϕ . The polynomial was of the lowest order possible, since higher orders introduce more spurious spread at the edges of the fitted region. In all but one case a linear fit was sufficient. The process was repeated with the corrected particle radii, dr , in the θ direction to find and subtract $dr_c(\theta)$. Figure 2-14 illustrates this process.

Correcting for the asphericity in this manner determines the radius $r_c(\theta, \phi)$ of the

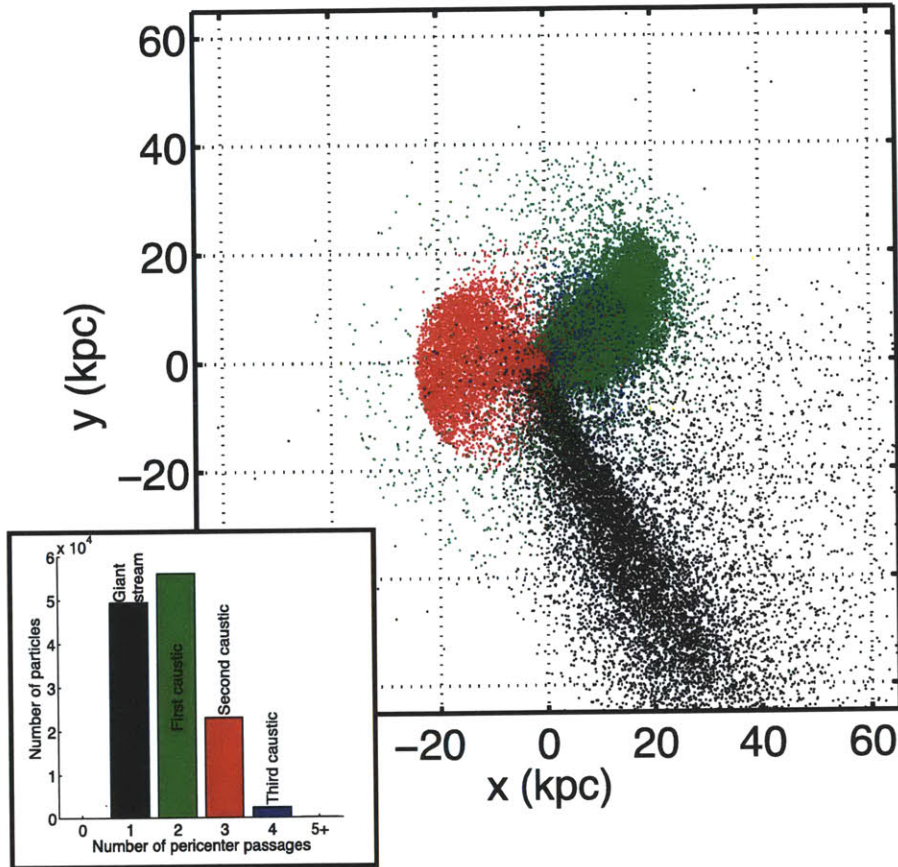


Figure 2-12: Sorting particles by the number of pericenter passages, N_{peri} (the inset gives the color scheme used here), easily identifies the main dynamical structures in the tidal debris. This x - y projection shows the N-body model as it would be seen from Earth, with x and y measured relative to M31's center and aligned with the east and north directions on the sky, respectively. In this projection the edge of the younger (red) shell is sharper than that of the older (green) shell, but both are nearly spherical relative to M31's center.

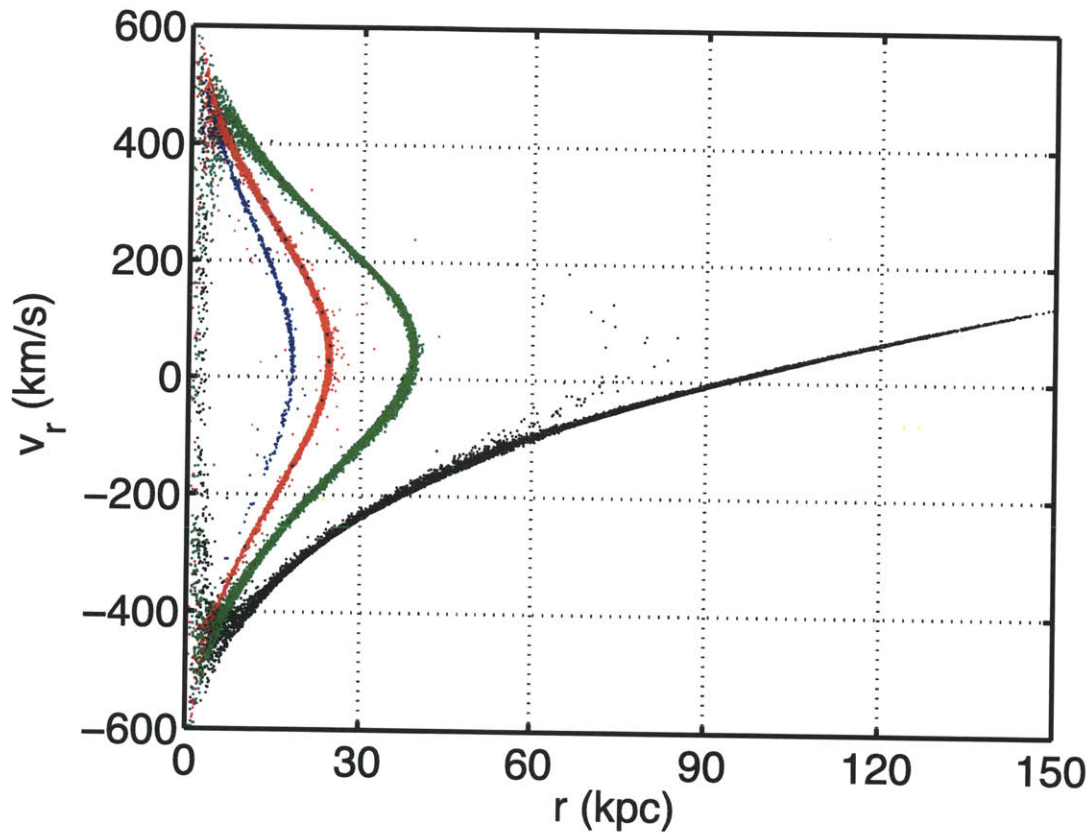


Figure 2-13: Looking at the projected phase space of r and v_r , measured relative to M31's center, lets us determine the order in which the shells formed. In the self-similar model, the outermost caustic forms first: we see here that this corresponds to the bound particles with lowest N_{peri} (colors are the same as in Figure 2-12). The youngest caustic contains the particles that have undergone the largest number of orbits.

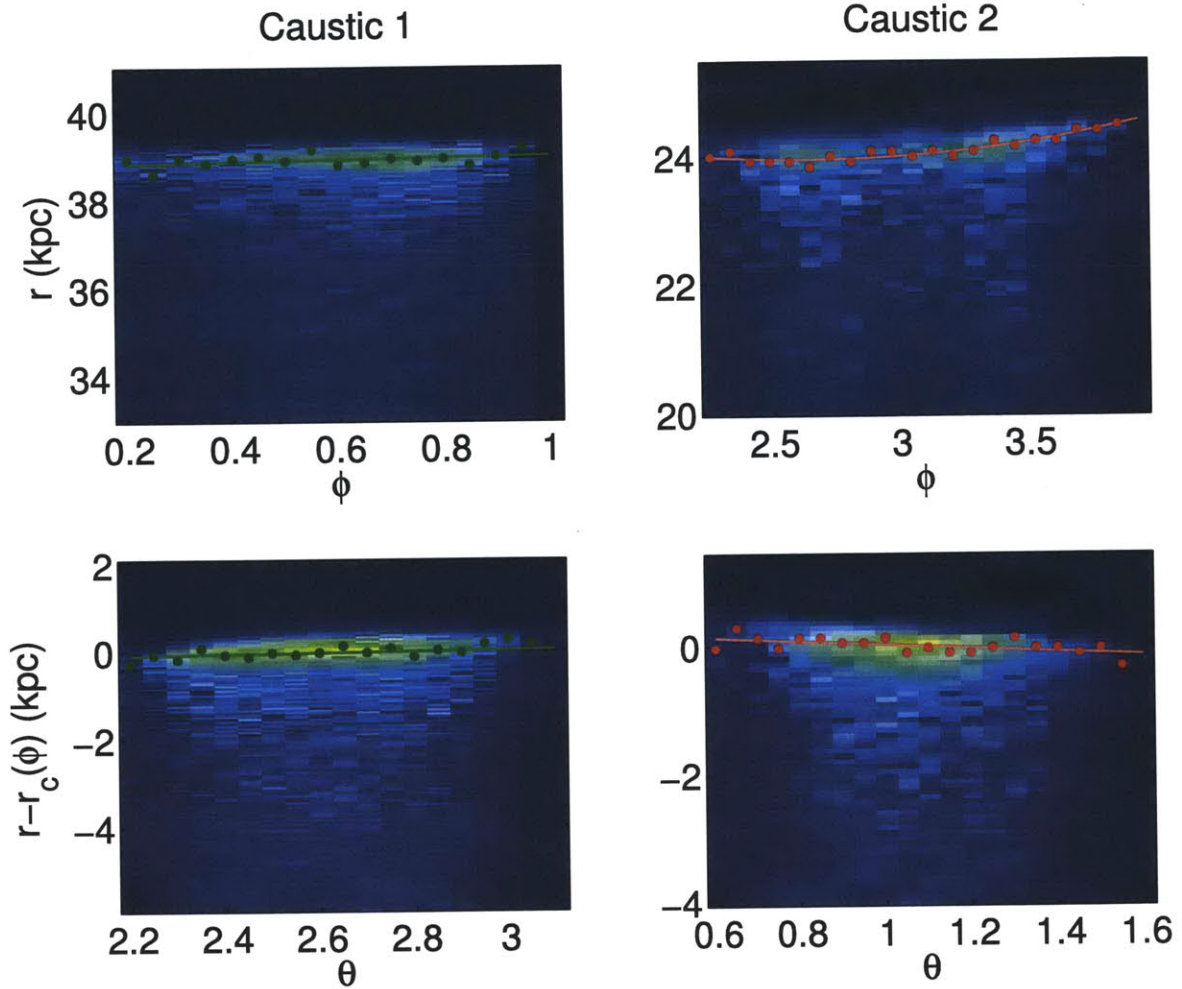


Figure 2-14: Each caustic surface was corrected for asphericity by fitting the position of the peak bin in r as a function of angle, first in the ϕ direction (top row) and then in the θ direction (bottom row). The points show the peak bin location and the lines indicate the best fit for the position of the caustic surface. In the background is the two-dimensional binned density map of the caustic. A linear fit was sufficient for all but the ϕ -dependence of Caustic 2 (top right panel) which used a quadratic fit.

caustic surface in two steps, so that

$$r_c(\theta, \phi) \equiv r_c(\phi) + dr_c(\theta). \quad (2.32)$$

The bin widths in r used for caustics 1 and 2 in this procedure limit the accuracy of $r_c(\theta, \phi)$ to 0.05 and 0.08 kpc, respectively. However, fitting the radial profile with the analytical caustic determines the caustic radius more accurately, assuming that $r_c(\theta, \phi)$ has sufficiently corrected the asphericity.

After correcting the caustics for asphericity, we binned the particles in $x \equiv r - r_c(\theta, \phi)$ to construct a radial density profile. We used the Wand rule [97] to choose a starting bin width and fit Equation (3.37) to the density profile to get baseline parameters, then decreased the bin width until the fit parameters converged. The final bin width was 0.1 kpc for both caustics, larger than the bins used to correct for the asphericity, indicating that $r_c(\theta, \phi)$ adequately represents the caustic surface. Using smaller bins produced no appreciable change in the fitted parameters.

Having determined the ideal bin width, we fit the model caustic of Equation (3.37) to each density profile. The model has a total of four parameters: the caustic width δr , x_c , the initial phase space density f_0 , and the phase-space curvature κ , that are discussed in detail in Chapter 3. As noted there, this density profile is universal for caustics created by quasi-radial infall of initially virialized material. We allowed all the parameters except κ to vary in the fit, but checked the fitted values of x_c , δr and f_0 by comparing them to estimates. For x_c this is trivial; it should be close to zero after the asphericity correction. We used mass conservation to estimate f_0 , since the number of particles in the caustic and its geometry are both known, and energy conservation to estimate δr , using Equation (3.36) and about 10-20 percent of the particles in each shell just around the caustic peak. For reference, the covariance term in Equation (3.35) was two orders of magnitude smaller than the other terms.

For each caustic, we determined κ by fitting a parabola to the phase-space profile near the peak. The portion of data to fit was determined by the range over which the parabola was a good fit to the phase-space profile. As a check, we also estimated

κ using the local gravity at each caustic, using Equation (3.31). Both are shown in the table and agree with each other. Keeping the value of κ constant, we then fit the density profile of the caustic, comparing the fitted values of f_0 and x_c with the estimated or expected values. The results of the fitting are shown in Table 2.5.

As expected, the model fits the data well all the way through the peak in both cases (Figure 2-15). The abrupt drop in density behind the caustic (at negative x) is an artifact of selecting the particles via their phase-space profile. In both cases x_c , the correction to r_c from the fit (dashed line), is within a few bins of the value from the asphericity correction (solid line at zero). x_c is not positioned exactly at the peak of the caustic because the satellite is not totally cold (see Chapter 3 for a more detailed explanation). The widths of the two caustics are close, but not identical, reflecting the slight difference in the initial velocity dispersions of the particles creating them. The estimate of δr for caustic 1 is much closer than that for caustic 2; this could be because the second caustic is at about half the distance of the first and is thus more affected by the non-spherical portions of the potential. The spread of energies with r in caustic 2 is certainly both larger and less symmetric than in caustic 1. Both estimates are slightly high thanks to the simplifying assumptions described in the appendix.

To finish calibrating the rate calculation, we compared the characteristic widths δr and N_p of the two caustics with those used in the numerical experiments to determine whether the rate estimate suffered from significant undersampling bias. The numerical experiments used $t = 1$ Myr (we can choose the units of time and length freely). Caustic 1 has $\delta r = 0.20$ kpc and Caustic 2 has $\delta r = 0.23$ kpc, which correspond to test distributions with $\log_{10} \sigma = -0.70$ and $\log_{10} \sigma = -0.63$, respectively. According to Table 2.5, Caustic 1 has $\log_{10} N_p = 4.3$ and Caustic 2 has $\log_{10} N_p = 3.8$. From the results of the tests in Section 2.2.5, we find that for $\log_{10} N_p = 4.25$, the RMS error at $\log_{10} \sigma = -0.75$ is 4.0 percent using $N_s = 10$ with the FiEstAS estimator and 3.4 percent with the uniform estimator. For $\log_{10} N_p = 3.75$, the RMS error at $\log_{10} \sigma = -0.75$ is 5.8 percent using $N_s = 10$ and the uniform estimator. All these errors are dominated by the standard deviation, not the undersampling bias.

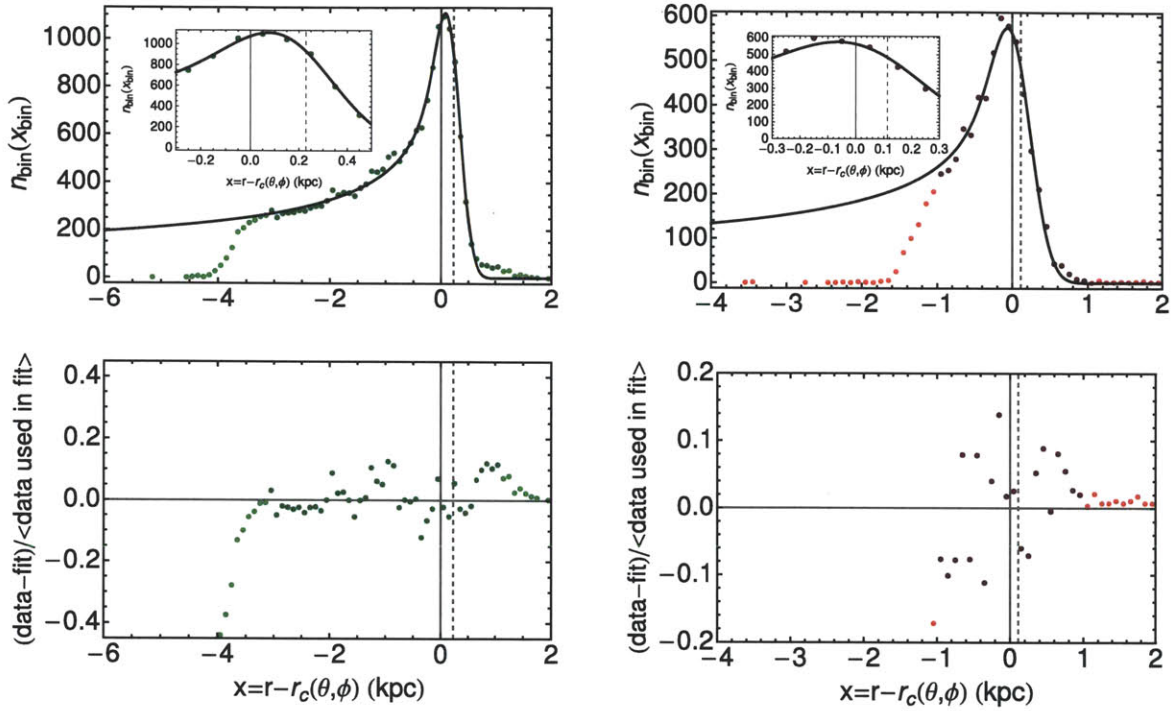


Figure 2-15: The radial density profiles of caustics 1 (left upper panel; green) and 2 (right upper panel; red) can be fit surprisingly well with the functional form in Equation (2.16) after they have been corrected for asphericity with the process illustrated in Figure 2-14. The insets in the upper panels show how well the function fits the region right around the peak of each caustic, which is the most important region for determining the true width. The residuals (lower panels) from the data used in the fit (shown in a darker shade) are evenly scattered around zero in each case, indicating that the peak's relative height is accurately determined. The fit parameters are given in Table 2.5.

Parameter	Value for caustic 1	Value for caustic 2	Notes
N_p	19779	6524	number of particles in part of caustic used in fit
$r_c(\theta, \phi)$ (kpc)	$38.4 + 0.21\theta + 0.14\phi$	$26.0 - 1.8\theta + 0.34\theta^2 + 0.3\phi$	as defined in Equation (2.32)
x_c (kpc)	0.228 ± 0.005	0.113 ± 0.009	from fit to density profile after asphericity corrections using Equation (3.36)
δr , estimated (kpc)	0.23	0.39	from fit to density profile
δr , fitted (kpc)	0.201 ± 0.005	0.232 ± 0.009	estimated using Equation (3.31)
κ , estimated, kpc $(\text{km s}^{-1})^{-2}$	4.6×10^{-4}	2.4×10^{-4}	from fit to phase-space profile
κ , fitted, kpc $(\text{km s}^{-1})^{-2}$	$(4.2 \pm 0.2) \times 10^{-4}$	$(2.31 \pm 0.04) \times 10^{-4}$	estimated using mass conservation
f_0 , estimated, kpc $^{-3}$ $(\text{km s}^{-1})^{-1} \text{sr}^{-1}$	0.37	0.096	from fit to density profile
f_0 , fitted, kpc $^{-3}$ $(\text{km s}^{-1})^{-1} \text{sr}^{-1}$	(0.33 ± 0.01)	(0.086 ± 0.003)	

Table 2.5: Density profile parameters for caustics 1 and 2. Error ranges on fitted parameters indicate the 95 percent confidence interval.

2.3.4 Boost factor

Based on the fits in the previous section, we expect the shells to be fully resolved in the simulation. So we are free to choose any of the three estimators we tested to calculate the contribution to the total rate from interactions between shell particles, Γ_{ss} . We chose to use $\hat{\Gamma}_f$ for convenience, since choosing a constant Riemann volume whose size relative to the shells' thickness is consistent with our tests would require at least 10^7 Riemann volumes to fill the simulation volume, whereas with adaptive Riemann volumes the large low-density portions require much less computation time. Γ_{sh} , which represents interactions between dark matter from the shell and dark matter in the halo, was calculated using the density estimator \hat{n}_f (Equation (A.7)) to estimate n_s , and evaluating Equation (2.20) for n_h at the same points where n_s is estimated. The core radius was set to half the size of the Riemann volume enclosing the origin; in practice about 0.1 kpc. Γ_{hh} was calculated analytically by integrating Equation (2.20) over the simulation volume. We find that the boost factor $\beta = 2.4 \times 10^{-3}$ for Γ integrated over the entire simulation volume, independent of the particle physics model.

By far the largest contribution to Γ_{hh} comes from the very center of the halo. To get a more realistic estimate of the boost factor we recalculated it with this region excised, which would certainly be done for a real observation given that astrophysical gamma rays appear to come from the disk. We excluded a square region 1.35 kpc (0.2 degrees) on a side, centered on M31's center. This choice of exclusion region corresponds to about twice the resolution limit of the Fermi LAT, and is intended to be equivalent to excising the central 4 pixels about M31's center. This technique increases the boost factor by only a small amount, to $\beta = 0.0027$ or 0.27 percent.

Finally, we mapped the spatial variation of the boost factor by using the space-filling tree in the FiEstAS algorithm. As expected, the largest boost factors come from the edges of the two shells containing most of the mass, as shown in Figure 2-16; the maximum is 2.5 percent, independent of the particle physics model.

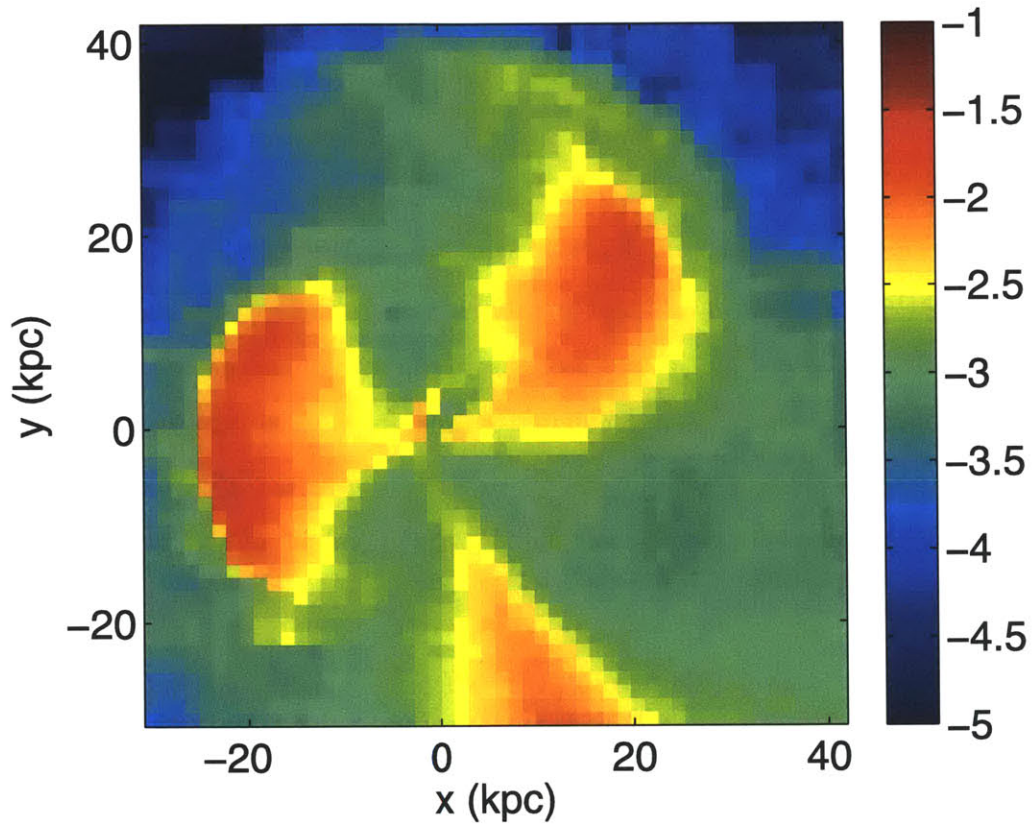


Figure 2-16: The largest boost factors, shown in red, are from the edges of the two shells. The contrast in this figure is independent of the parameters for the particle physics model (summarized as Φ^{SUSY}).

2.3.5 Astrophysical factor

The “astrophysical factor” is the quantity Φ^{cosmo} defined by Equations (2.26), (2.29), and (2.30). For a given model of dark matter, comparing values of Φ^{cosmo} gives the relative strength of different mass distributions as sources of high-energy particles via self-annihilation. As described in Section 2.3.4, we used $\hat{\Gamma}_f$ to calculate Φ^{cosmo} for the shells, both for interactions between two dark matter particles in the shell material and interactions between dark matter in the shell and dark matter in the halo (Table 2.6). For comparison we also calculated Φ^{cosmo} for the dwarf galaxy used in the N-body model. The dwarf is represented as a Plummer sphere with mass $M_P = 2.2 \times 10^9 M_\odot$ and scale radius $b = 1.03$ kpc. Integrating the density-squared of the dwarf over volume shows that

$$\Gamma_{\text{dwarf}} = \frac{1}{b^3} \left(\frac{3M_P}{8m_p} \right)^2, \quad (2.33)$$

to be used with Equation 2.30. We also compared these values with one calculated by [56] from measurements of the mass and mass profile of the Ursa Minor dwarf galaxy in the Milky Way, scaled as if it were located in M31. Ursa Minor has approximately the same mass as M31’s faint satellites AndIX and AndXII [98]. We find that the shell signal is comparable in magnitude to the signal from this type of dwarf galaxy at the same distance, and two orders of magnitude less than the signal from the original Plummer sphere (whose mass is about ten times that estimated for Ursa Minor and its analogues in M31).

Object	$\Phi^{\text{cosmo}}, \text{Gev}^2 \text{kpc cm}^{-6}$
shell-shell (using Γ_{ss})	9.9×10^{-9}
shell-halo (using Γ_{sh})	8.5×10^{-7}
Plummer dwarf (using Γ_{dwarf})	4.1×10^{-5}
“Ursa Minor” [based on 56]	6.9×10^{-7}

Table 2.6: Values of the astrophysical factor Φ^{cosmo} for various configurations of the tidal debris, calculated using Equation (2.30).

2.3.6 Gamma-ray Signal

Beyond calculating Γ and β , we used Equation (2.31) to estimate the flux of gamma rays in the Fermi LAT for the various benchmarks in [10]. The results are shown in Table 2.7 along with the parameters used to calculate Φ^{SUSY} in each case.

By using the three-dimensional representation of the density-squared field from the shell constructed with the FiEstAS estimator as a piecewise definition of the rate, we can integrate the rate along the line of sight and across pixels of arbitrary size such that the sum of the flux in all the pixels equals the total flux in Table 2.7. With this technique we made two-dimensional maps of the expected gamma-ray emission using the most optimistic value of Φ^{SUSY} (labeled as “Upper Limit” in Table 2.7). Using this value, even the center of M31’s halo is only barely detectable by Fermi, if the halo is shaped as we assumed for the dynamical model and the amount of dark matter in the dwarf galaxy is comparable to the amount of luminous matter, although the emission from the dark halo completely dominates over that from the shell (Figure 2-17, left panel). The tidal structure is at least an order of magnitude too faint to be detected even if the halo component is fitted and removed (Figure 2-17, right panel). Because of interactions between halo and tidal dark matter, the signal from the tidal debris scales linearly with the mass of its progenitor as long as $\rho_h > \rho_s$ (Equation (2.19)) so the ratio of dark matter to luminous matter in the dwarf galaxy would need to be several orders of magnitude larger, even after tidal stripping, for the tidal debris to be detectable with Fermi or for the scaling of the boost factor to become quadratic in the shell dark matter density. For a smaller progenitor such a ratio might be plausible, but dwarf galaxies with masses of $10^9 M_\odot$ or higher tend to have comparable masses of luminous and dark matter in their centers based on our understanding of the Tully-Fisher relation at those masses [25].

2.4 Conclusions

We find that unless all the features in a given density distribution are known to be fully resolved, the best way to estimate the volume integral of the square of the

Table 2.7: Contribution to the Flux of gamma-rays above 1 GeV for various MSSM benchmarks.

Table 7
Contributions to the Flux of Gamma Rays Above 1 GeV from WIMP Self-annihilation, for Various MSSM Benchmarks

...	A'	B'	C'	D'	G'	H'	I'	J'	K'	L'	Upper Limit
m_{χ} , GeV ^a	242.8	94.9	158.1	212.4	148.0	388.4	138.1	309.1	554.2	181.0	40
$N_{\gamma}(\sigma v)^a$	120	782	195	63.6	1032	86.5	6303	930	7.08×10^4	1.87×10^4	1.30×10^4
Φ_{SUSYb}	3.14	134	12.0	2.18	72.7	0.885	510	15.0	356	882	1.26×10^4
$\Phi_{\gamma,\text{hh,all}}^{c,d}$	3.38	144	13.0	2.34	78.3	0.953	550	16.2	383	950	13530
$\Phi_{\gamma,\text{sh,all}}^e$	0.00780	0.341	0.0307	0.00554	0.185	0.00225	1.30	0.0382	0.906	2.25	32.0
$\Phi_{\gamma,\text{ss,all}}^f$	$< 10^{-3}$	0.00424	$< 10^{-3}$	$< 10^{-3}$	0.0023	$< 10^{-3}$	0.0161	$< 10^{-3}$	0.0112	0.0279	0.398
$\Phi_{\gamma,\text{addl,all}}^g$	0.00810	0.345	0.0310	0.00561	0.187	0.00228	1.32	0.0387	0.917	2.28	32.4
$\Phi_{\gamma,\text{total,all}}^h$	3.39	145	13.0	2.35	78.5	0.955	551	16.2	384	953	13560
$\Phi_{\gamma,\text{hh,nc}}^i$	1.92	81.8	7.35	1.33	44.4	0.5402	311	9.17	217	539	7672
$\Phi_{\gamma,\text{sh,nc}}$	0.00509	0.217	0.0195	0.00353	0.118	0.00144	0.827	0.0244	0.577	1.43	20.4
$\Phi_{\gamma,\text{ss,nc}}$	$< 10^{-3}$	0.00362	$< 10^{-3}$	$< 10^{-3}$	0.00196	$< 10^{-3}$	0.0138	$< 10^{-3}$	0.00961	0.0238	0.339
$\Phi_{\gamma,\text{addl,nc}}$	0.00518	0.221	0.0198	0.00359	0.120	0.00146	0.841	0.0248	0.586	1.46	20.7
$\Phi_{\gamma,\text{total,nc}}$	1.92	82.0	7.37	1.33	44.5	0.542	312	9.19	218	540	7693

Notes. The subscripts hh, sh, and ss refer to the various terms in Equation (18). $N_{\gamma}(\sigma v)$ has units $10^{-29} \text{ cm}^3 \text{ s}^{-1}$. Φ_{SUSY} has units $10^{-11} \text{ cm}^4 \text{ kpc}^{-1} \text{ s}^{-1} \text{ GeV}^{-2}$. All Φ_{γ} have units $10^{-14} \text{ } \gamma \text{ cm}^{-2} \text{ s}^{-1}$. For reference, the *Fermi point source* sensitivity for photons with $E > 100 \text{ MeV}$ is on the order of $10^{-9} \text{ } \gamma \text{ cm}^{-2} \text{ s}^{-1}$ (Rando et al. 2009).

^a Taken from Table 1 of Gondolo et al. (2004) except for the rightmost column, which is based on Figure 8 of FPS as described in the text.

^b Calculated analytically using Equation (25).

^c Calculated analytically using the same NFW halo as for the dynamical model.

^d *all*: integration is over entire line of sight and covers the region shown in Figure 17 in x and y .

^e Calculated by constructing a numerical estimate for n_{shell} in each integration volume element, then evaluating n_{halo} analytically at the center of that volume element and assuming its value is constant over the entire element.

^f Calculated numerically as described in the text.

^g $\Phi_{\gamma,\text{addl}} \equiv \Phi_{\gamma,\text{hs}} + \Phi_{\gamma,\text{ss}}$.

^h $\Phi_{\gamma,\text{total}} \equiv \Phi_{\gamma,\text{hh}} + \Phi_{\gamma,\text{hs}} + \Phi_{\gamma,\text{ss}}$.

ⁱ *nc*: a central region is excluded from the calculation, as described in the text.

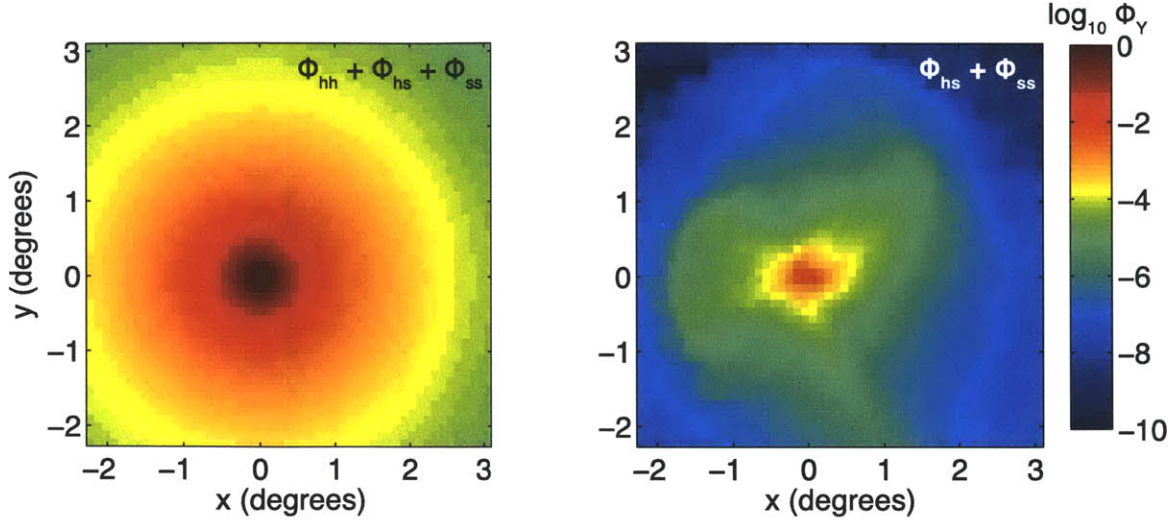


Figure 2-17: Gamma-ray emission from the dark halo, though faint, dominates over emission from the shell even at large radius (left panel). If $\Phi_{\gamma, hh}$ is removed, the remaining emission from the shell is too faint to detect with Fermi even for the most optimistic parameters in the set of benchmarks (right panel). In these images the pixels are 0.1 degree on a side to imitate the approximate degree resolution of Fermi at the energy scale of interest, and the zero of the degree scale is centered on M31’s center.

density (the “rate”) from an N-body realization is to use the simple nearest-neighbors estimator with a constant Riemann volume. If the realization completely resolves even the sharpest features, all three estimators we tested should agree on the result. The simplest method for estimating the rate works best for this problem because the other, more complicated algorithms are optimized for density estimation, not rate estimation, and because estimators using adaptive Riemann volumes appear to require slightly more particles to resolve features of a given sharpness. In any case the estimator should be calibrated for Poisson bias as we describe in Section 2.2.1. We also find that the improvement in the standard deviation achieved by increasing the smoothing number is smaller than the increased bias from blurring more small-scale structure for $N_s > 10$. The correct calibration of the estimator for Poisson bias can change the estimated result by up to 10 percent for reasonable values of N_s . The correct calibration for undersampling bias can change the result by a factor of 2 or more if the simulation is under-resolved; rather than attempt to correct for

it, it is better to ensure that the N-body realization has sufficient resolution for the small-scale features to be resolved.

Using a calibrated estimator and a sufficiently resolved N-body realization, we calculated the boost factor and signal in gamma rays from tidal debris in M31 that displays high-contrast features. Although we find as expected that the largest boosts come from the shell edges, they only increase the total signal by at most 2.5 percent over the signal from a self-consistent smooth halo. Likewise, the total gamma-ray flux from the shells is three orders of magnitude lower than emission from the dark halo, and too low to be detected by Fermi for likely dark matter parameters (Table 2.7). The total signal is comparable to that predicted for an ultra-faint satellite of M31.

2.5 Future Work

The existence of shell features around M31 provides many avenues other than indirect detection for learning about the nature, dynamics, and distribution of dark matter. The very existence of the shells demands that the dwarf galaxy that created them must have had very low angular momentum relative to M31 because the pericenter distance is so small. Whereas high-angular-momentum systems like that of the Sagittarius dwarf galaxy in the Milky Way are useful for constraining the shape of dark halos because the tidal debris explores a large range in angle, low-angular-momentum systems like the M31 shells and giant stream probe M31's potential over a large range in radius, and are best suited for constraining the degeneracy between the different mass components of the host galaxy. They also act as a sensitive probe of the mass profile of the progenitor, since the combination of relatively cold initial conditions in the dwarf and a small pericenter distance acts as a kind of spectrometer, spreading the mass of the satellite galaxy out in space according to its total energy. Because the shells' relative orientations are a good limit on the projected angular momentum of the progenitor, variations in the initial position and velocity of the center of mass are not likely to be degenerate with variations in the shape and phase space distri-

bution of the debris, although this is still being tested. This makes the shapes and phase space distributions of the shells extremely sensitive to the initial phase space distribution of the progenitor satellite, and can place limits on the cuspieness of the mass profile of the dwarf.

The analytical caustic used in this work can also be used as a model for caustics that form under the much more complicated equations of motion responsible for quasi-radial gravitational infall. As shown in Appendix 3, the resulting form is identical to that obtained by [37] in their analysis of those caustics with intuitive identifications of the normalization, caustic location and distance from the caustic surface, and requires no numerical integration to obtain the complete profile so it may be easily used for fitting. Although our model is less general (it does not predict the relative locations of caustics) it is consistent with the more general case, and more tractable if only the universal density profile is desired. The height and width of each caustic are sensitive to the initial phase space distribution of material in the caustic, while the profile depends on the potential of the host galaxy only through the gravitational force at the location each caustic—a complete mass model is not necessary. In light of recent discoveries of shells around many more nearby galaxies besides Andromeda [29], this technique may provide a way to constrain the properties of luminous matter in dwarf galaxies by examining the tidal debris they produce. This topic is discussed in the next two chapters of this thesis.

Chapter 3

Derivation of an analytic, one-dimensional caustic

In this chapter we derive the density profile stated in Equation (2.16) from a simple, one-dimensional force-free model (Section 3.1). As discussed in Chapter 1, this profile can also be used to describe a radial caustic created by a tidally disrupted satellite galaxy. For this case, we relate the parameters of the density profile to the properties of the satellite and host galaxies (Section 3.2). In Chapter 2, we used the integrated density-squared (the “rate”) calculated from this density profile to test the performance of density estimators on systems with high density contrast. In Section 3.3, we derive an expression for this quantity in three Cartesian dimensions of which one contains a caustic.

3.1 Density Profile

Following the notation of [38], we define the Eulerian coordinates $\vec{x} \equiv (x, y, z)$ in terms of the Lagrangian coordinates or “initial conditions” $\vec{q} \equiv (q_x, q_y, q_z)$ for a collisionless nongravitating ensemble of particles:

$$\vec{x}(\vec{q}, t) = \vec{q} + \vec{v}_0(\vec{q})t \tag{3.1}$$

where \vec{v}_0 is the initial velocity vector of the particles.

Conservation of mass states that:

$$\rho(\vec{x}, t)d^3\vec{x} = \sum_{\vec{q} \rightarrow \vec{x}} \rho_0 d^3\vec{q} \quad (3.2)$$

where $\vec{x}(\vec{q}, t)$. The sum is taken over all values of \vec{q} for which particles end up at a given \vec{x} . Assuming a uniform starting density ρ_0 , the local density ρ of the distribution then evolves in time as

$$\rho(\vec{x}, t) = \sum_{\vec{q} \rightarrow \vec{x}} \frac{\rho_0}{|d\vec{x}/d\vec{q}|} = \sum_{\vec{q} \rightarrow \vec{x}} \frac{\rho_0}{\left| \delta_{ik} + t \frac{\partial v_{0,i}}{\partial q_k} \right|} \quad (3.3)$$

Caustics arise when the determinant in the denominator is zero.

To randomly sample a simple one-dimensional caustic along x in our 3D distribution, we can choose a uniform distribution in y and z :

$$y = q_y \quad (3.4)$$

$$z = q_z \quad (3.5)$$

and an extremely simple displacement rule for x that includes a constant initial velocity $v_{0,x}(q_x)$ with a small random component $v_{\text{th}}(q_x)$ along the dimension of the caustic and no further interactions:

$$x = q_x + t(v_{0,x}(q_x) + v_{\text{th}}(q_x)) \quad (3.6)$$

$v_{\text{th}}(q_x)$ is a random variable drawn from a distribution $f(v_{\text{th}})$ for each particle in the system labeled by a unique q_x . This ‘‘thermal velocity’’ is included so that the density will not diverge at the caustic surface. Equation (3.6) is not invertible since the v_{th} are random. Furthermore the particle ordering at $t > 0$ is unknown because of these random velocities, so the density at some point may now include many streams, since the initial velocity determines when the particles cross the caustic. To make sure we

include all the streams, we express the sum over streams as an integral over a delta function (dropping the subscripts on q for clarity):

$$\rho[v_{\text{th}}(q)](x, t)dx = \rho_0 \int_{-\infty}^{\infty} dq \delta[x - q - v(q)t] dx \quad (3.7)$$

where $v(q) \equiv v_0(q) + v_{\text{th}}(q)$ includes both the uniform and random parts of the initial particle velocity. The delta function and integral over q form a sum over all particles that arrive at the location x from any q at a time t .

We choose to represent the thermal velocities with a Gaussian (Maxwell) distribution with one-dimensional dispersion σ :

$$f(v_{\text{th}})dv_{\text{th}} = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-v_{\text{th}}^2/2\sigma^2} dv_{\text{th}} \quad (3.8)$$

$v_{\text{th}}(q)$ is thus a function of q in the sense that for each particle, labeled by its q , a different random velocity is assigned. The dispersion σ could easily be time-dependent, but again we take the simplest case and assume it is a constant. We refer to the v_{th} as a thermal component because systems in thermal equilibrium have velocities described by the same distribution, so it is often associated with temperature since $\sigma^2 \sim T$ in that case.

To make the integral in the expression for $\rho(x)$ easier to evaluate, we can also replace the delta function with its limit definition in terms of a Gaussian:

$$\delta[x - q - v(q)t] = \lim_{\epsilon \rightarrow 0} \frac{1}{\sqrt{2\pi\epsilon^2}} e^{-[x - q - v(q)t]^2/2\epsilon^2} \quad (3.9)$$

Combining, we find

$$\rho[v_{\text{th}}(q)](x, t) = \rho_0 \int_{-\infty}^{\infty} dq \lim_{\epsilon \rightarrow 0} \frac{1}{\sqrt{2\pi\epsilon^2}} e^{-[x - q - v(q)t]^2/2\epsilon^2} \quad (3.10)$$

To get the distribution $\rho(x, t)$ we must take the ensemble average over v_{th} :

$$\rho(x, t) = \langle \rho[v_{\text{th}}(q)](x, t) \rangle = \int_{-\infty}^{\infty} dv_{\text{th}} \rho[v_{\text{th}}(q)](x, t) f(v_{\text{th}}) \quad (3.11)$$

We reintroduce $v(q) \equiv v_0(q) + v_{\text{th}}(q)$ and pull out the parts of the exponent containing v_{th} :

$$\rho(x, t) = \rho_0 \int_{-\infty}^{\infty} dq \lim_{\epsilon \rightarrow 0} \frac{1}{\sqrt{2\pi\epsilon^2}} e^{-[x-q-v_0(q)t]^2/2\epsilon^2} \int_{-\infty}^{\infty} dv_{\text{th}} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-v_{\text{th}}^2/2\sigma^2} e^{-\{t^2 v_{\text{th}}^2 - 2v_{\text{th}}t[x-q-v_0(q)t]\}/2\epsilon^2} \quad (3.12)$$

For notational simplicity, we temporarily define $[x - q - v_0(q)t] \equiv \Delta$. Regrouping terms, we find a quadratic expression in the exponent in the inner integral:

$$\rho(x, t) = \rho_0 \int_{-\infty}^{\infty} dq \lim_{\epsilon \rightarrow 0} \frac{1}{2\pi\sigma\epsilon} e^{-\Delta^2/2\epsilon^2} \int_{-\infty}^{\infty} dv_{\text{th}} \exp \left\{ - \left[v_{\text{th}}^2 \left(\frac{1}{2\sigma^2} + \frac{t^2}{2\epsilon^2} \right) - \frac{t\Delta}{\epsilon^2} v_{\text{th}} \right] \right\} \quad (3.13)$$

The integral over v_{th} can be evaluated by completing the square on the quantity in curly brackets. Setting

$$a^2 \equiv \left(\frac{1}{2\sigma^2} + \frac{t^2}{2\epsilon^2} \right) \quad (3.14)$$

we find that the inner integral becomes

$$I \equiv \int_{-\infty}^{\infty} dv_{\text{th}} \exp \left[-a^2 \left(v_{\text{th}} - \frac{t\Delta}{2\epsilon^2} \right)^2 + \frac{t^2\Delta^2}{4\epsilon^4 a^2} \right] \quad (3.15)$$

Using the substitution $u = a(v_{\text{th}} - t\Delta/2\epsilon^2)$, we can immediately evaluate the integral to be:

$$I(x, t) = \exp \left(\frac{t^2\Delta^2}{4\epsilon^4 a^2} \right) \sqrt{\frac{\pi}{a^2}} \quad (3.16)$$

Replacing this into the equation for the density, we find that

$$\rho(x, t) = \rho_0 \int_{-\infty}^{\infty} dq \lim_{\epsilon \rightarrow 0} \frac{1}{\sqrt{4\pi\sigma^2 a^2 \epsilon^2}} \exp \left[-\Delta^2 \left(\frac{1}{2\epsilon^2} - \frac{t^2}{4\epsilon^4 a^2} \right) \right] \quad (3.17)$$

Replacing a^2 with its definition and simplifying makes the limit easy to take:

$$\rho(x, t) = \rho_0 \int_{-\infty}^{\infty} dq \lim_{\epsilon \rightarrow 0} \frac{1}{\sqrt{2\pi(\epsilon^2 + \sigma^2 t^2)}} e^{-\Delta^2/2(\epsilon^2 + \sigma^2 t^2)} \quad (3.18)$$

and finally we are down to the last integral:

$$\rho(x, t) = \frac{\rho_0}{\sqrt{2\pi\sigma^2 t^2}} \int_{-\infty}^{\infty} dq e^{-[x-q-v_0(q)t]^2/2\sigma^2 t^2} \quad (3.19)$$

As σ or t approaches zero, this expression approaches $\delta[x - q - v_0(q)t]$, and we recover the density in the case of zero dispersion.

To perform the integral over q (that is, to sum over streams) we must choose a form for $v_0(q)$. The simplest form that creates a localized caustic surface is $v_0(q) = -\alpha q^2$:

$$\rho(x, t) = \frac{\rho_0}{\sqrt{2\pi\sigma^2 t^2}} \int_{-\infty}^{\infty} dq e^{-[x-q+\alpha q^2 t]^2/2\sigma^2 t^2} \quad (3.20)$$

For this form of v_0 we can evaluate the integral

$$I(x, t) = \int_{-\infty}^{\infty} dq e^{-[x-q+\alpha q^2 t]^2/2\sigma^2 t^2} \quad (3.21)$$

in terms of Bessel functions if we make a slightly suspect change of variables. First we complete the square in q inside the brackets:

$$I(x, t) = \int_{-\infty}^{\infty} dq \exp \left\{ -\frac{\alpha^2}{2\sigma^2} \left[\frac{x}{\alpha t} - \frac{1}{4\alpha^2 t^2} + \left(q - \frac{1}{2\alpha t} \right)^2 \right]^2 \right\} \quad (3.22)$$

Then we make the substitution

$$u \equiv q - \frac{1}{2\alpha t} \quad (3.23)$$

This change of variables is certainly suspect as $t \rightarrow 0$, but we make it anyway, hoping that since we know the answer at $t = 0$ we can check the result and verify that it gives ρ_0 everywhere. With this substitution

$$I(x, t) = \int_{-\infty}^{\infty} du \exp \left[-\frac{\alpha^2}{2\sigma^2} \left(\frac{x}{\alpha t} - \frac{1}{4\alpha^2 t^2} + u^2 \right)^2 \right] \quad (3.24)$$

We now identify the caustic position x_c . In the case of perfectly cold initial conditions, the location of the caustic is determined by setting the denominator of Equation (3.3)

to zero and solving for x_c , with the help of Equation (3.1). We use this same quantity to describe the “position” of the caustic in the warm case, since although the caustic now has a width, the peak density will still occur near x_c . For our choice of $v_0(q)$ we find that this corresponds to

$$\frac{x_c}{\alpha t} - \frac{1}{4\alpha^2 t^2} = 0 \quad (3.25)$$

so that $x_c = 1/(4\alpha t)$ and

$$\frac{x}{\alpha t} - \frac{1}{4\alpha^2 t^2} = \frac{1}{\alpha t}(x - x_c) \equiv \bar{x} \quad (3.26)$$

for notational simplicity when performing the integral. So now we have a factored quartic in the exponent:

$$I(x, t) = \int_{-\infty}^{\infty} du \exp \left[-\frac{\alpha^2}{2\sigma^2} (\bar{x} + u^2)^2 \right] \quad (3.27)$$

Expanding this expression and pulling out a constant term leaves us with an integral that can be evaluated in terms of Bessel functions:

$$I(\bar{x}, t) = e^{-\alpha^2 \bar{x}^2 / 2\sigma^2} \int_{-\infty}^{\infty} du \exp \left(-\frac{\alpha^2}{2\sigma^2} u^4 - \frac{\bar{x}\alpha^2}{\sigma^2} u^2 \right) \quad (3.28)$$

Performing the integral, replacing \bar{x} with its definition, and reintroducing the prefactor, we find that

$$\rho(x, t) = \frac{\rho_0}{\sqrt{2\pi\sigma^2}} \sqrt{\frac{|x - x_c|}{\alpha t^3}} e^{-(x-x_c)^2/4\sigma^2 t^2} \mathcal{B} \left[\frac{(x - x_c)^2}{4\sigma^2 t^2} \right] \quad (3.29)$$

with

$$\mathcal{B}(u) = \begin{cases} \frac{\pi}{2} [\mathcal{I}_{-1/4}(u) + \mathcal{I}_{1/4}(u)] & x \leq x_c \\ \frac{\pi}{2} [\mathcal{I}_{-1/4}(u) - \mathcal{I}_{1/4}(u)] & x > x_c \end{cases} \quad (3.30)$$

where \mathcal{I} is a modified Bessel function of the first kind. Taking the limit $t \rightarrow 0$ recovers ρ_0 everywhere, and validates our change of variables. This form is the same as that obtained by [37] for the density profile of a caustic with Gaussian velocity dispersion, with the substitutions $\alpha_k \sigma_v \rightarrow \sigma t$, $\Delta x \rightarrow x - x_c$, and $A_k = \rho_0 / \sqrt{\alpha t}$. That is, it

describes the same shape as a caustic formed by secondary self-similar infall from an initial population with Gaussian velocity dispersion, but whose normalization varies as $1/\sqrt{t}$ and whose position varies linearly with time.

3.2 Gravitational caustic model

The density profile 3.29, with a few slight changes, is universal for any initial population with a Gaussian velocity dispersion and small initial phase space volume, regardless of the equation of motion of those particles. This is because the shape of the phase space distribution of particles near the caustic can always be approximated by a tilted quadratic with curvature κ . For the equation of motion used to derive Equation (3.29), $\kappa = \alpha t^3$; for motion in a gravitational potential, κ depends on the radial gravitational force at the caustic:

$$\kappa = -\frac{1}{2} \left. \frac{d^2 r}{dv_r^2} \right|_{r_c} = -\frac{1}{2g(r_c)} \quad (3.31)$$

where $g(r)$ is the gravitational field or radial derivative of the potential, $\partial V/\partial r$.

Inspection of Equation (3.29) shows that the curvature κ influences the height of the caustic relative to the initial density, while the product σt determines its sharpness. In a more general case, the spreading of the particles in phase space is not linear in time and the initial velocity dispersion, and σt can be replaced everywhere in the density profile by a generic parameter representing the width of the caustic, δr . The quantity $\rho_0/\sigma \equiv f_0$ in the normalization of the caustic indicates that the physical density at time t depends on the initial phase space density, with the caveat that in cases where a caustic forms in one dimension but motion occurs in more than one dimension, σ remains the one-dimensional velocity dispersion while ρ_0 must be determined via mass conservation, since the density profile does not account for the behavior of the population in the neglected dimensions or for the Jacobian associated with the volume element in non-Cartesian metrics (for example, spherical coordinates).

The caustic width δr can be estimated using energy conservation. Although each

particle turns around at apocenter where $v_r = 0$, the caustic itself can have a net positive velocity because its location at a given time depends on the relative periods of particles on neighboring orbits. Thus the energy at the caustic surface, assuming spherical symmetry, is:

$$E = V(r) + \frac{1}{2}v_r^2 \quad (3.32)$$

For two particles on neighboring orbits, their energy difference ΔE is found by subtracting their energies, so

$$\Delta E = \frac{1}{2}(v_2^2 - v_1^2) + V(r_2) - V(r_1) \quad (3.33)$$

Particles at the very face of the caustic are close in r and far from $r = 0$, so we can write $r_2 = r_1 + \Delta r$, taking $\Delta r/r_1$ to be much less than 1. We cannot also write $v_2 = v_1 + \Delta v$ and $\Delta v/v_1 \ll 1$, since the average velocity of particles in the caustic may not be much larger than zero. Using the condition on r_1 , the energy difference can be rewritten in terms of the local gravity (defined earlier):

$$\Delta E = \frac{1}{2}(v_2^2 - v_1^2) - g(r_1)\Delta r \quad (3.34)$$

In reality many particles will comprise the caustic surface; one can use this expression by taking its variance, realizing that r and v are correlated and assuming $g(r)/\Delta r \gg dg/dr$:

$$\sigma_E^2 = \frac{1}{2}\sigma_{v^2}^2 - g(r)\text{cov}(r, v^2) + g(r)^2\sigma_r^2 \quad (3.35)$$

The size of the covariant term is related to the amount of curvature over the region used to calculate the variances. It can be minimized by taking a region around the caustic small enough that the curvature is smaller than the radial thickness, or $\delta v \leq \delta r/\sqrt{\kappa}$. Then the covariance term can be ignored and the equation solved to give an estimate of the width σ_r :

$$\sigma_r = \frac{\sqrt{\sigma_E^2 - \frac{1}{2}\sigma_{v^2}^2}}{g(r_c)} \quad (3.36)$$

With these physically motivated generalizations, the density profile

$$\rho(x, t) = \frac{f_0}{\sqrt{2\pi}} \sqrt{\frac{|x - x_c|}{\kappa}} e^{-(x-x_c)^2/4\delta r^2} \mathcal{B} \left[\frac{(x - x_c)^2}{4\delta r^2} \right] \quad (3.37)$$

will universally fit the projected radial density profile of any shell near its peak. Using this relation, we can find the location of the peak by solving $d\rho/dr = 0$. The equation reduces to a linear combination of various Bessel functions, which can be solved numerically to find r_{\max} . Here we give the result to 10 decimal places, but more are easily available if necessary simply by increasing the number of iterations in the root finder.

$$r_{\max} = r_c - 0.7649508674\delta r \quad (3.38)$$

Substitution into the density profile gives the maximum density in terms of the physical parameters. Within a few percent the numerical coefficient is unity:

$$\rho_{\max} = 1.0211365847 f_0 \sqrt{\frac{\delta r}{\kappa}} \quad (3.39)$$

The peak density occurs about one characteristic width behind the nominal caustic surface r_c (Figure 3-1).

3.3 Integrated density-squared

Armed with an analytic expression for the density, we can now calculate the integrated squared density or “rate” Γ for comparison with estimates from random realizations. We must integrate the density with finite limits on q and x , since we used the density estimator on a finite-sized sample of the distribution for the tests described in Chapter 2. Using Equation (3.29) for the density profile of the caustic in the x direction and assuming a constant density in y and z , the integrated squared density is

$$\Gamma(t) = 4L_y L_z \int_{x_-}^{x_+} \rho^2(x, t) dx \quad (3.40)$$

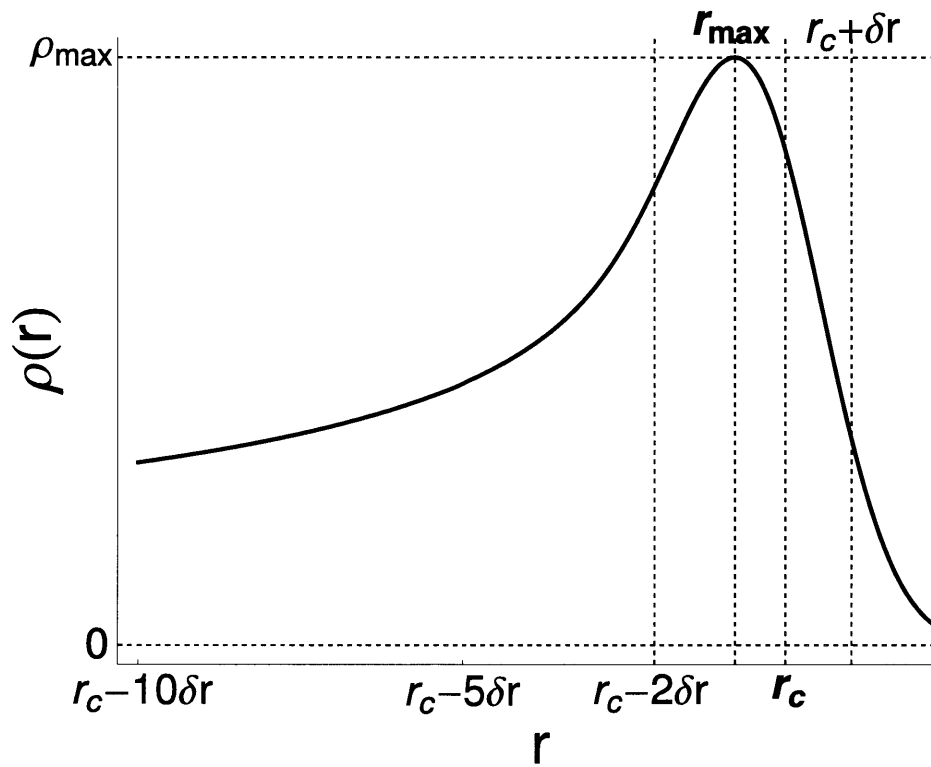


Figure 3-1: Universal density profile for a one-dimensional caustic from with initial Gaussian velocity dispersion. The parameters are defined in the text.

with x_{\pm} the lower and upper edges of the sample in the x direction, and $\pm L_y$ and $\pm L_z$ the sample dimensions in the y and z directions, respectively. Furthermore, we are now considering only particles that originated in some range $[q_-, q_+]$, so we must also recalculate the density function integral, Equation (3.27), over a finite range:

$$\rho(\bar{x}, t) = \frac{\rho_0}{\sqrt{2\pi\sigma^2 t^2}} \int_{u_-}^{u_+} du \exp \left[-\frac{\alpha^2}{2\sigma^2} (\bar{x} + u^2)^2 \right] \quad (3.41)$$

where

$$u_{\pm} \equiv q_{\pm} - \frac{1}{2\alpha t} \quad (3.42)$$

Technically, we should use this expression for ρ^2 in the rate integral, so that

$$\Gamma(t) = \frac{2L_y L_z \alpha \rho_0^2}{\pi \sigma^2 t} \int_{\bar{x}_-}^{\bar{x}_+} \left[\int_{u_-}^{u_+} du \exp \left[-\frac{\alpha^2}{2\sigma^2} (\bar{x} + u^2)^2 \right] \right]^2 d\bar{x}, \quad (3.43)$$

with

$$\bar{x}_{\pm} \equiv \frac{x_{\pm} - x_c}{\alpha t} \quad (3.44)$$

the endpoints of the rate integral about the caustic.

Both these integrals must be evaluated numerically, so we would like to borrow our analytic solution, Equation (3.29), instead if possible. We can see that this is sufficient by examining the form of Equation (3.29). The exponential and Bessel function terms are functions of the quantity

$$\frac{(x - x_c)}{2\sigma t} \equiv \bar{x}_d \quad (3.45)$$

which compares the distance from the caustic to the characteristic length scale for dispersion, σt . For distances much larger than σt , the exponential piece approaches zero. The modified Bessel function $\mathcal{K}_{1/4} \equiv \frac{\pi}{\sqrt{2}} [\mathcal{I}_{-1/4}(u) - \mathcal{I}_{1/4}(u)]$ also has an exponential cutoff for $(x - x_c) \gg \sigma t$, or $u \rightarrow \infty$, as seen in its asymptotic series:

$$K_{1/4}(u) = \sqrt{\frac{\pi}{2}} e^{-u^2} \left[\frac{1}{u} - \frac{3}{32} \frac{1}{u^3} + \mathcal{O}\left(\frac{1}{u^5}\right) \right] \quad (3.46)$$

For $x < x_c$, the leading term in the asymptotic series of ρ declines as a relatively slow power-law:

$$\rho(x, t) = \rho_0 \sqrt{\frac{t\sigma^2}{\alpha|x - x_c|}} + \mathcal{O}(|x - x_c|^{-5/2}). \quad (3.47)$$

From examining these limits, we find that for $x > x_c$ the density is completely negligible for distances more than a few times σt from the caustic, and for $x < x_c$ it becomes negligible like a power law. Thus it is unsurprising that substituting the infinite-range expression for the finite-range one gives the right answer.

Thus, as long as $x - x_c \gg \sigma t$ at the endpoints, we can safely write

$$\Gamma(t) = 4L_y L_z \frac{\rho_0^2}{2\pi\sigma^2 t^2} \frac{1}{t} \int_{x_-}^{x_+} |x - x_c| e^{-2(x-x_c)^2/4\sigma^2 t^2} \mathcal{B}^2 \left[\frac{(x - x_c)^2}{4\sigma^2 t^2} \right] dx \quad (3.48)$$

Practically speaking, the approximation is sufficient for our needs as long as $\bar{x} \gtrsim 500$ at the endpoints. If this condition isn't satisfied we must do the two numerical integrals in Equation (3.43).

We change variables to \bar{x}_d . Inserting the form for \mathcal{B} breaks the integral up into two parts:

$$\begin{aligned} \Gamma(t) = & \frac{4L_y L_z \rho_0^2}{t} \left\{ \int_{\bar{x}_{d-}}^0 |\bar{x}_d| e^{-2\bar{x}_d^2} [\mathcal{I}_{-1/4}(\bar{x}_d^2) + \mathcal{I}_{1/4}(\bar{x}_d^2)]^2 d\bar{x}_d + \right. \\ & \left. + \int_0^{\bar{x}_{d+}} |\bar{x}_d| e^{-2\bar{x}_d^2} [\mathcal{I}_{-1/4}(\bar{x}_d^2) - \mathcal{I}_{1/4}(\bar{x}_d^2)]^2 d\bar{x}_d \right\} \quad (3.49) \end{aligned}$$

where we have assumed that we are integrating across the caustic so that the new endpoints,

$$\bar{x}_{d\pm} \equiv \frac{x_{\pm} - x_c}{2\sigma t}, \quad (3.50)$$

are such that $\bar{x}_{d-} < 0$ and $\bar{x}_{d+} > 0$. This final integral must be performed numerically.

Chapter 4

Constraining properties of dark matter halos with tidal caustics

4.1 Introduction

Evidence that galaxies interact with one another has been seen for a long time, and the creation of streams of stars through tidal interactions is an established field of study. It has also long been known that tidal debris comes in many shapes and sizes, from the huge tidal arms created by interacting, roughly equal-mass galaxies to the fainter features observed around nearby galaxies that are thought to come from interactions with much smaller satellites, called minor mergers. Evidence that minor mergers occur in nature is an important link to our understanding of cosmological history, since cosmological simulations of dark matter predict that many more small galaxies exist than larger ones, and that minor mergers should therefore occur frequently [for example, 31]. Minor mergers are also useful as a way to constrain the shape and mass of the large galaxy, since the material unbound from the satellite galaxy behaves as test particles in the relatively undisturbed potential of the large galaxy. Observing and interpreting minor mergers helps relate the dark components of galaxies, predicted with great accuracy by cosmological models and simulations, to the luminous matter in them, which is more difficult to predict.

Some minor mergers create patterns of tidal debris that look like shells or umbrel-

las. The first such debris was identified around elliptical galaxies by [41]; these galaxies were called “shell galaxies” because of these distinctive features. [99, 100, 64, 65] showed that the shells were probably created by a minor merger on a nearly radial orbit. This explains the alternate spacing of the shells on either side of the host galaxy since they are formed as material from the satellite piles up approximately at turning points: the satellite material initially had a distribution of energies that is reflected by the different radii of the shells. More recently, similar features have been discovered around nearby disk galaxies [101, 30, 29]. The vast improvements in imaging since shell galaxies were first identified, and the relative proximity of these objects, has revealed more of their structure than had previously been observed. In some cases the objects are even close or bright enough that velocity information could be obtained. Additionally, since the first numerical experiments to model shell galaxies were performed, computational capacity has increased exponentially to allow a much broader exploration of parameter space with more sophisticated models. Finally, at the time that shell galaxies were being studied, the concordance cosmological model (Λ CDM) had not yet been proposed, let alone experimentally verified, and dark matter was not necessarily thought to form a significant part of the mass distribution of galaxies. All these advances in theoretical understanding and technical capability suggest that shell galaxies are worth another look.

Shells from nearly radial mergers are particularly special because there is a direct correlation between the kinematic properties of debris and its location relative to the host galaxy. Thanks to the near-symmetry of the encounter, the system can be considered in a two-dimensional projection (r, v_r) of the full six-dimensional phase space without much loss of information. In this two-dimensional projection, the initially cold satellite material, once unbound, forms a thin stream that winds through phase space, so that for any spatial location r there are a small, finite number of possible speeds v_r . The dynamics governing the formation and shape of this stream are closely related to earlier work on spherically symmetric secondary infall of matter accreting onto dark matter halos. As shown in [36] and [35] radial accretion of gravitating, cold, collisionless matter forms a series of infinite-density peaks at successive radii, known

as caustics. [37] further showed that for warm matter with a finite velocity dispersion, the peaks take on a finite width and height, but although they are no longer caustics in the mathematically rigorous sense they retain many of the same properties. This prior work was motivated by cosmological simulations and so considered the radial collapse of spherically distributed matter, but [43] show that the results apply equally well to a small, initially self-bound satellite falling radially into the static potential of a larger host galaxy. For this reason, in this work we will refer to shells as “tidal caustics”, a term which emphasizes the high degree of symmetry and the correlation of positions and velocities.

This chapter describes a method to estimate both the local gravitational field near a tidal caustic and the original velocity dispersion of the dwarf galaxy that created it. *These quantities can be constrained using only images of the host galaxy and merger debris, and estimated independently if kinematic information and a rotation curve of the host galaxy are available.* Construction of a full N-body model of the merger is not necessary. The constraints or estimates are made possible by exploiting the universality of the one-dimensional density profile of the material in the caustic. In Section 4.2 we present the universal density profile and explain its parameters. In Section 4.3 we describe a set of numerical experiments to determine qualitatively how radial a merger must be to use our techniques, and what such tidal debris will look like. In Section 4.4 we discuss how the model may be used to set limits on the host and satellite galaxies in a radial merger if only photometric data in projection is available, and in Section 4.5 we discuss a simple geometric model for determining the projection angles. In Section 4.6 we extend to cases where line-of-sight velocities are available. Section 4.7 discusses additional information that is given in systems with multiple visible shells. In Section 4.8 we demonstrate that the conclusions drawn in this chapter still apply when the gravitational response of the host halo is accounted for. In Section 4.9 we discuss the next step in applying this model to real data, and in Section 4.10 we summarize.

4.2 A universal density profile

In Chapter 3, we derived a simple analytical form for a one-dimensional caustic formed from a system with initial random velocities drawn from a Maxwellian distribution (Figure 4-1, left panel). In the case where this system is a small, gravitationally self-bound satellite galaxy falling radially into the center of a static host galaxy, the one-dimensional density profile of the caustic, as a function of the galactocentric radius r , can be written in terms of four physical parameters δr , r_c , κ , and f_0 :

$$\rho(r) = \frac{f_0}{\sqrt{2\pi\kappa}} \sqrt{|r - r_c|} e^{-(r-r_c)^2/4\delta r^2} \mathcal{B} \left[\frac{(r - r_c)^2}{4\delta r^2} \right] \quad (4.1)$$

where \mathcal{B} is a piecewise combination of modified Bessel functions of the first kind:

$$\mathcal{B}(u) = \begin{cases} \frac{\pi}{2} [\mathcal{I}_{-1/4}(u) + \mathcal{I}_{1/4}(u)] & r \leq r_c \\ \frac{\pi}{2} [\mathcal{I}_{-1/4}(u) - \mathcal{I}_{1/4}(u)] & r > r_c \end{cases} \quad (4.2)$$

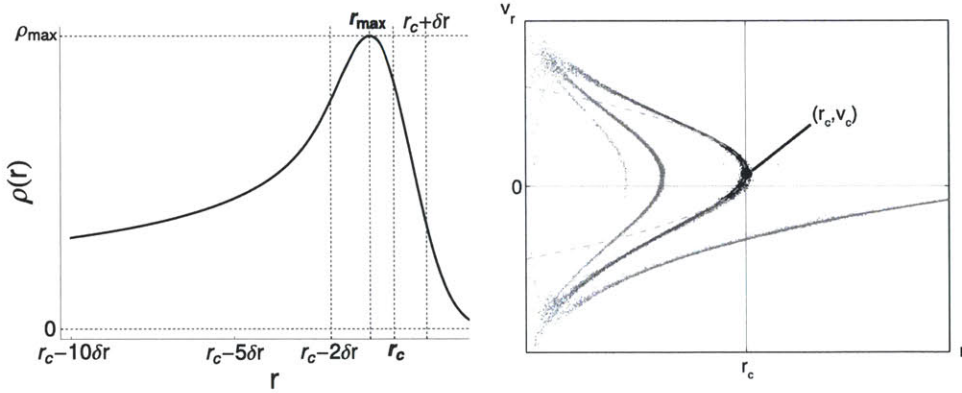


Figure 4-1: Left: Universal caustic form for radial infall from a population with initial Gaussian velocity dispersion. r_c and δr are defined and formulae for the peak radius and density are given in the text. Right: An example phase space distribution taken from a simulation of a minor merger. Near the peak of one of the caustics (black), the phase space distribution follows the curve $r = r_c - \kappa(v_r - v_c)^2$ (dashed line), where v_c is the radial velocity of particles at the caustic surface. κ is defined in the text.

In Equation (4.1), δr is the characteristic width of the caustic surface, which depends on the initial velocity dispersion of the dwarf galaxy that created the caustic. A perfectly “cold” distribution, in which all particles had the same velocity, would

give rise to a true caustic, with zero width and infinite density. We refer to caustics in a less rigorous sense: they have finite density and width but still exhibit a large local density enhancement thanks to the same dynamical processes. The width δr is related, by conservation of energy, to the variance in energy per unit mass, σ_E^2 , and in velocity-squared, σ_{v^2} , of the matter making up the caustic. It also depends on the local gravity of the host galaxy at the caustic, $g(r) \equiv -d\Phi/dr$:

$$\delta r = \frac{\sqrt{\sigma_E^2 - \frac{1}{2}\sigma_{v^2}^2}}{-g(r_c)} \quad (4.3)$$

Although all the particles forming the caustic are near apocenter, the net radial velocity at the caustic surface is not necessarily zero because the spatial overlap of their orbits depends on their orbital periods. The velocity of the caustic surface can be determined by requiring $dr/dE = 0$ for some given time t after the matter becomes unbound, where $r(E)$ is determined implicitly by the equation

$$t = \int_0^r \frac{dr'}{\sqrt{-2[E - \Phi(r)]}} \quad (4.4)$$

This expression depends on the entire mass profile of the host galaxy, not just the local gravity. The assumption of spherical symmetry only holds at large radius, but this is where the particles spend most of their time, so this relationship, even in realistic potentials, is an integral constraint on the mass profile.

The radius of the caustic surface is denoted r_c . It is close to, but not equal to, the radius of peak density, r_{\max} , because of the thickness of the caustic, induced by the nonzero velocity dispersion of the material. For a perfectly cold initial population $r_c = r_{\max}$ and the peak density is infinite, but for a system with nonzero velocity dispersion the peak density is finite,

$$\rho_{\max} = 1.021 f_0 \sqrt{\frac{\delta r}{\kappa}}, \quad (4.5)$$

and r_c is related to r_{\max} by

$$r_{\max} = r_c - 0.765 \delta r. \quad (4.6)$$

The numerical factors come from solving $d\rho/dr = 0$ numerically (necessary because of the Bessel functions) and can be determined to arbitrary precision.

Although in general galaxies are not thought to have a spherical mass distribution, especially near the disk, we make this simplification at the large radii where the caustics are usually detected (several times the scale length of the disk or more). A nearly spherical mass distribution will give rise to a caustic surface that is well fit as a segment of a sphere spanning some solid angle Ω . The appearance of the caustic on the sky is then modified by projection along the viewing angle; deprojection is necessary to determine r_c correctly and avoid overestimating δr . We discuss deprojection methods further in the next section.

The local curvature of the phase space distribution of the particles at the caustic surface is denoted κ . Because by definition the surface is located where the phase space distribution is vertical, near the surface the its shape may always be approximated by a quadratic function. The curvature κ is related to the local gravity at the caustic surface by

$$\kappa = -\frac{1}{2} \left. \frac{d^2 r}{dv_r^2} \right|_{r_c} = -\frac{1}{2g(r_c)} \left[1 - \frac{v_r^2}{g(r)^2} \left. \frac{dg}{dr} \right|_{r_c} \right], \quad (4.7)$$

where the derivative $d^2 r/dv_r^2$ is taken along the projected (r, v_r) phase space curve traced out by the debris. The second term in the square brackets is usually smaller than 1 at large radius because the gravity gradient of the host galaxy and the radial velocity at the caustic surface are both small.

The one-dimensional average phase space density of the material in the caustic, once its self-gravity can be neglected, is f_0 . In a typical collision with low angular momentum, the dwarf galaxy remains self-bound until its first pericenter passage, after which the potential of the host galaxy dominates the evolution and the self-gravity of the dwarf can be ignored. Even if this is not true for all the material, it is so for anything that ends up in caustics. To use f_0 to estimate of the true

average phase space density, the motion in r must dominate over the angular motion so that the projection of the full phase space into the r - v_r plane captures most of the spreading in phase space. We discuss this further in Section 4.3.

In a spherically symmetric system, f_0 can be estimated by using mass conservation:

$$M_{\text{shell}} = \Omega \int_{r_c - \Delta r}^{r_c + \Delta r} r^2 dr \rho(r) \quad (4.8)$$

where Ω is the solid angle spanned by the caustic, $\pm\Delta r$ describes a symmetric slice in r centered on the caustic surface, M_{shell} is the mass in that region, and the density profile $\rho(r)$ is given by Equation (4.1). Changing variables to $u \equiv r - r_c$ and substituting the function ρ ,

$$M_{\text{shell}} = \frac{\Omega f_0}{\sqrt{2\pi\kappa}} \mathcal{I}(\Delta r), \quad (4.9)$$

where

$$\mathcal{I}(\Delta r) \equiv \int_{-\Delta r}^{\Delta r} du (u + r_c)^2 \sqrt{|u|} e^{-u^2/4\delta r^2} \mathcal{B}\left(\frac{u^2}{4\delta r^2}\right) \quad (4.10)$$

Given an estimate for M_{shell} , the pseudo-phase space density is

$$f_0 = \frac{M_{\text{shell}} \sqrt{2\pi\kappa}}{\Omega \mathcal{I}(\Delta r)}. \quad (4.11)$$

4.3 Initial conditions

In order to use the universal density profile, two assumptions must be valid: the potential in which the debris material is moving must be dominated by the potential of the host galaxy, and the orbits of the material must be sufficiently radial that the phase space can safely be collapsed into the r - v_r plane. We conducted a brief series of numerical experiments to illustrate the type of merger that can give rise to a good candidate set of caustics. The host galaxy in these experiments is represented by a spherical isochrone potential with scale mass $M_h = 2.71 \times 10^{12}$ solar masses and scale radius $b_h = 8.02$ kpc. The satellite is represented as a Plummer sphere with scale radius $r_s = 1$ kpc and mass ratio $m_s/m_h = 5.27 \times 10^{-4}$. We start the satellite at a distance of 40 kpc from the host's center with a total speed equal to the circular

speed at that radius, 438 km/s. The direction of the satellite's velocity vector (and therefore its angular momentum) is parameterized by the angle ϑ between the velocity and radius vectors of the satellite's center of mass. ϑ is zero for a perfectly radial orbit and 90 degrees for a circular orbit.

From examining a projection of the spatial distribution of the debris (Figure 4-2), we see that caustics are distinguishable in the debris for ϑ up to 45 degrees. However, examining the r - v_r projection of phase space shows that ignoring the angular motion of the debris begins to introduce significant spread into the projection at much lower angles.

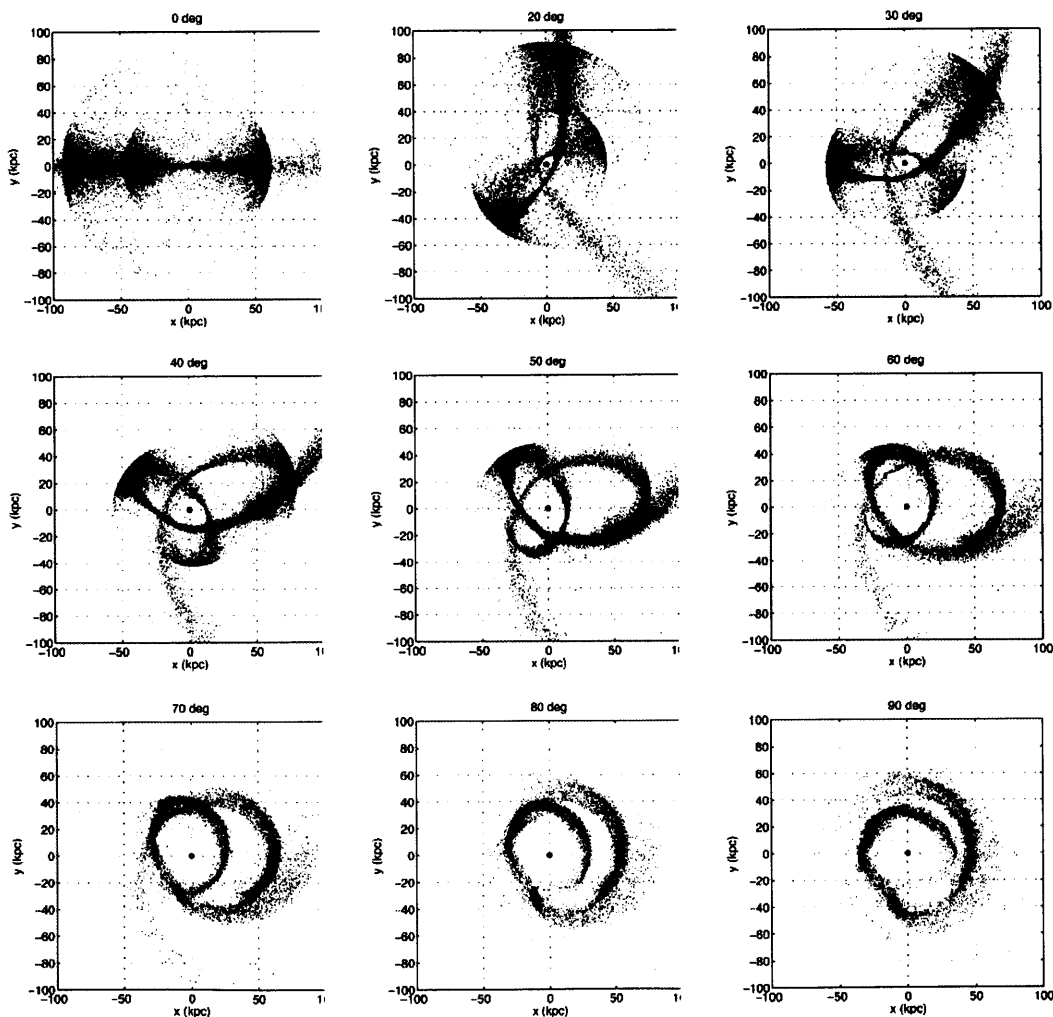


Figure 4-2: The appearance of tidal debris depends on the angle ϑ between the initial velocity and radius vectors of the satellite galaxy's center of mass. Caustics (sharp edges) are prominent for $\vartheta \lesssim 35^\circ$.

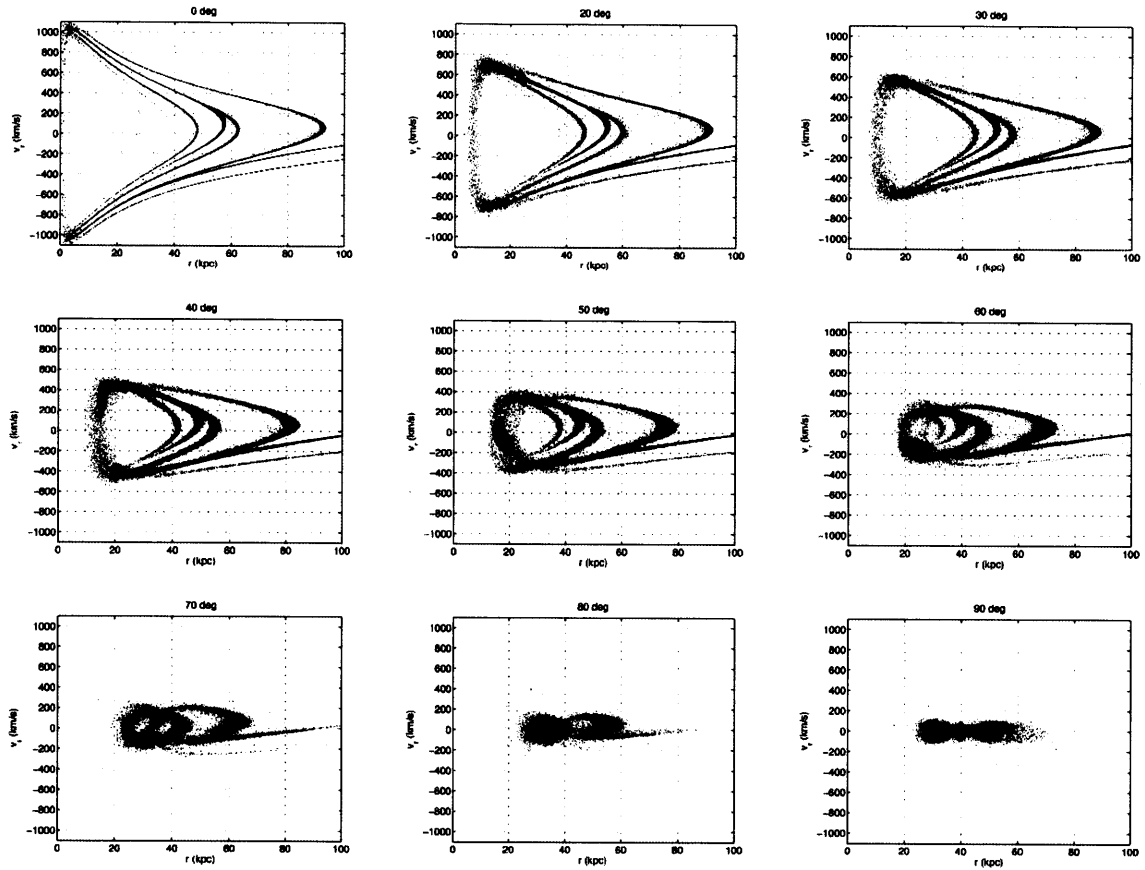


Figure 4-3: Examining the projected phase space shows that when ϑ is larger about 30 degrees, the angular momentum is large enough to significantly deform the shape of the phase space distribution in this projection.

4.4 Spatial distribution

Given only the density profile and deprojected image of a caustic, one can use the relations between the four parameters determining the density profile to constrain the structure of the satellite galaxy. Fitting the density profile measures δr , r_c , and $f_0/\sqrt{2\pi\kappa}$, from which $\mathcal{I}(\Delta r)$ can be calculated by integrating over the observed region Δr around r_c . M_{shell} and Ω for the same region can be estimated from the deprojected image with an appropriate assumption about the mass-to-light ratio. Then Equation (4.11) relates f_0 , the initial spread in the projected phase space, to the potential of the host at r_c :

$$\frac{M_{\text{shell}}}{\Omega\mathcal{I}(\Delta r)} = f_0 \left[-\frac{2\pi}{g(r_c)} \left(1 - \frac{v_r^2}{g(r)^2} \frac{dg}{dr} \Big|_{r_c} \right) \right]^{-1/2} \quad (4.12)$$

The term including the gravity gradient dg/dr is always positive. Since the potential of the host galaxy is not known, we can obtain an upper limit instead:

$$\frac{M_{\text{shell}}}{\Omega\mathcal{I}(\Delta r)} \geq f_0 \sqrt{-\frac{g(r_c)}{2\pi}} \quad (4.13)$$

Using Equation (4.3) allows us to eliminate $g(r_c)$ to obtain a constraint on the initial energy spread and phase space density, and the final radial velocity spread, of the material in the satellite:

$$2\pi\delta r \left(\frac{M_{\text{shell}}}{\Omega\mathcal{I}(\Delta r)} \right)^2 \geq f_0^2 \sqrt{\sigma_E^2 - \frac{1}{2}\sigma_{v_r^2}} \quad (4.14)$$

4.5 Accounting for projection effects

In order to obtain limits on the satellite galaxy's structure from images alone, it is necessary to account for projection effects in the image of the caustic. Fortunately the high symmetry of the system again simplifies the analysis. The edge of the caustic can be approximated as a spherical segment that spans some solid angle, and the angular extent of the debris can be modeled as a cone. The system thus has three angles that describe the shape and orientation of the debris: the angles θ_c and ϕ_c of

the cone relative to the line of sight and the opening angle α of the cone (Figure 4-4). The solid angle enclosed by the cone is $\Omega_c = 2\pi(1 - \cos \alpha)$. The angles are defined in the standard way for spherical coordinates, so that $(\theta_c, \phi_c) = (0, 0)$ corresponds to a cone opening directly away from the observer along the line of sight. If the caustic has a sharp edge in projection, this means that θ_c must be close to $\pi/2$, in which case initial guesses for ϕ_c and α can be estimated directly from the projected image.

The equations for the edges of the cone in Cartesian coordinates can be obtained by rotating the standard cone equations. We defer the full discussion of these equations to Section 5.2.1 of the next chapter, where the deprojection procedure is discussed in detail. Then the projected surface brightness, assuming a constant mass-to-light ratio Υ , is

$$\Sigma(x, y) = \Upsilon \int_{z_{\min}(x, y)}^{z_{\max}(x, y)} \rho \left(\sqrt{x^2 + y^2 + z^2} \right) dz \quad (4.15)$$

where ρ is given by Equation (4.1) and z_{\min} and z_{\max} are the limits of the cone for a given x and y . The method for determining z_{\min} and z_{\max} is also described in Section 5.2.1.

Given a particular pair of orientation angles (θ_c, ϕ_c) and opening angle α , the projected surface brightness can be derived from Equation (4.1). An example is shown in Figure 4-5. The integral along the line of sight (Equation (4.15)) must be done numerically, but thanks to the finite number of measurements of the surface brightness for a given cone, it need only be done for a finite number of lines of sight. Additionally, since the expressions for z_{\min} and z_{\max} are analytic in θ_c , ϕ_c , and α , standard minimization routines can be used to find best fit values based on the natural χ^2 from comparing the calculated and measured profiles. The necessary derivatives are also given in Section 5.2.1.

4.6 Phase space distribution

For some extragalactic tidal caustics, line-of-sight velocities can be obtained. Using the coordinate system and projection angles discussed in the previous section, the

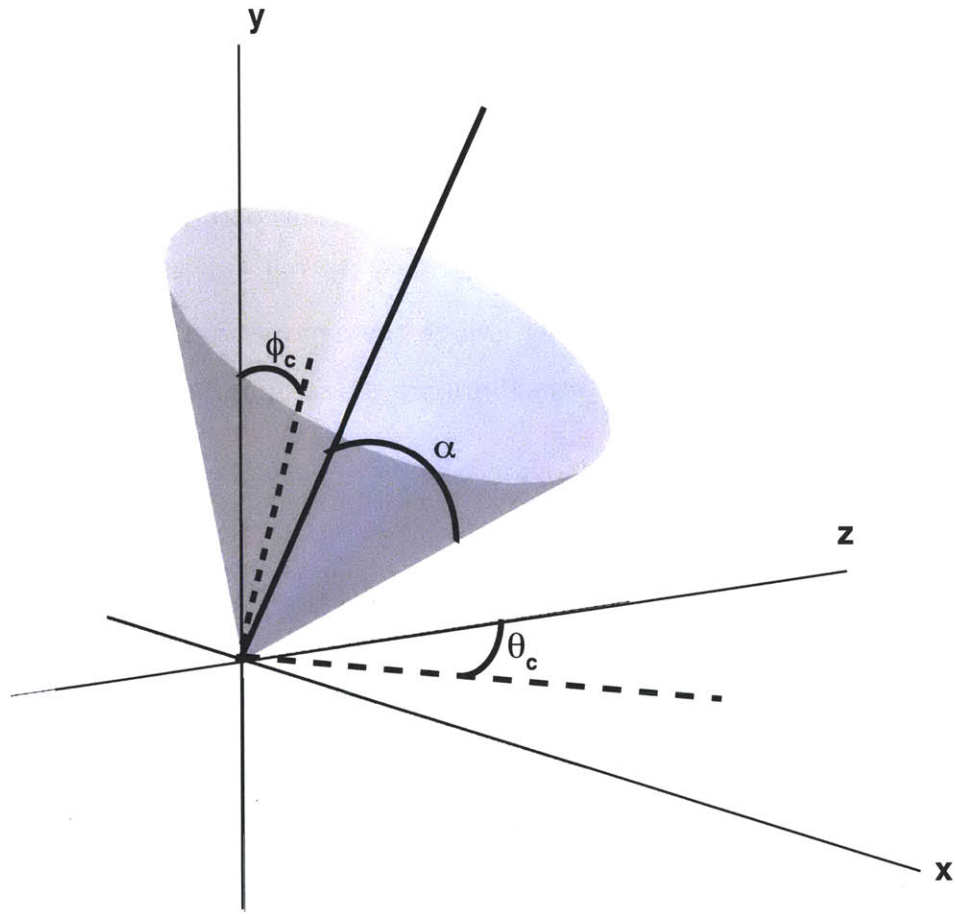


Figure 4-4: Angles describing the orientation of a cone representing the spatial extent of the tidal debris relative to the line of sight (z axis). As in standard spherical polar coordinates, θ_c is the angle between the line of symmetry of the cone and the positive z axis, which points away from the observer, and ϕ_c is the projected angle of the line of symmetry in the x - y plane (the plane of the image). α is the opening angle of the cone. If θ_c is near $\pi/2$, α can be approximately measured in the x - y plane.

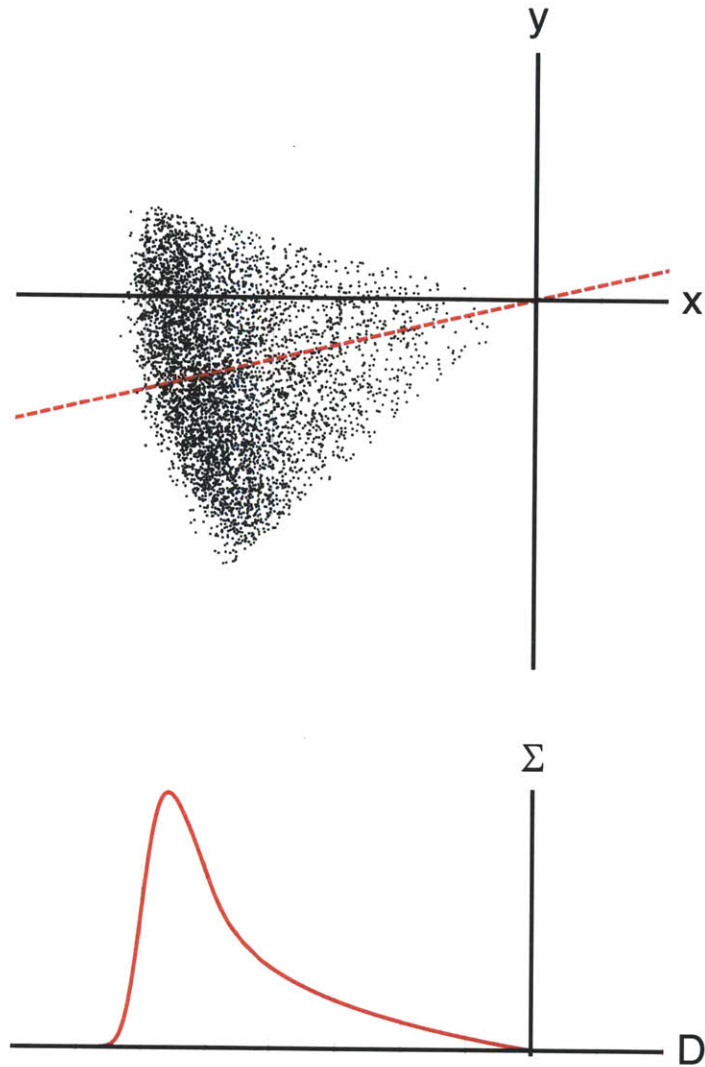


Figure 4-5: Using the simple geometric model described in the section (top), the surface brightness profile in a given direction (red dashed line in top panel) for a given orientation of the cone with respect to the line of sight can be derived analytically (bottom). The model pictured has $\alpha = 0.5$, $\theta_c = 1.3$, $\phi_c = 3.4$, and $\delta r/r_c = 0.3$.

line-of-sight velocity is

$$v_{\text{los}} = \vec{v} \cdot \hat{z} = v_r \cos \theta_c - v_\theta \sin \theta_c \quad (4.16)$$

For the caustic to have a sharp edge in projection, θ_c must be close to $\pi/2$; therefore, measurements of the line-of-sight velocity for such caustics will primarily constrain v_θ and therefore the angular momentum of the tidal debris. If multiple shells are observed, their relative angles on the sky, combined with the shell orientations estimated using the mass profile, can constrain the angular momentum of the satellite galaxy and the degree of flattening of the host potential. This method is discussed in the next section.

If v_r can be measured, either by constraining and removing v_θ from the line-of-sight velocities or by some other means, then the host galaxy's potential can also be constrained by using the shape of the phase space distribution near the caustic edge. Now κ , $v_r(r_c)$, and $\sigma_{v_r}^2$ are also known, leading to a new set of constraints. First, we can get a lower limit on the magnitude of the gravitational force at the caustic surface:

$$|g(r_c)| \geq \frac{1}{2\kappa} \quad (4.17)$$

Using this approximation, we can also verify that the second term in Equation 4.7 is much less than the first, by approximating dg/dr by $1/(2\kappa r_c)$ and $g(r_c)$ with the limit above. Combined with the measured $v_r(r_c)$ we get the condition

$$\frac{2v_r(r_c)^2\kappa}{r_c} \ll 1 \quad (4.18)$$

If this condition is satisfied, we can acceptably use the lower limit on $g(r_c)$ as a measurement of the local gravity of the host galaxy.

Second, having measured κ we can now compute f_0 directly using the value obtained for $f_0/\sqrt{2\pi\kappa}$ from the fit to the density profile, or from conservation of mass using Equation (4.11). Both approaches should yield the same value for f_0 , a useful consistency check.

Finally, we can get a lower limit for the energy spread of the material in the satellite galaxy by using Equation (4.3) and the limit on $g(r_c)$ derived above:

$$\sigma_E^2 \geq \frac{\delta r^2}{4\kappa^2} + \frac{1}{2}\sigma_{v_r}^2 \quad (4.19)$$

4.7 Multiple caustics

If multiple caustics are observed around the same galaxy, the angular momentum of the satellite galaxy can be estimated from the relative angles of the caustics (and/or tidal stream, if one is observed). As discussed in Chapter 1.2.2, the caustics from a single stripping event can be placed in order of orbital phase by comparing r_c ; the caustic with the smallest phase has the largest radius. The possibility of multiple episodes of tidal stripping complicates this picture, but if the tidal tail from each stripping event can be detected, then the shells from each event can still be phase-ordered. Both the angular momentum of the satellite and the additional precession caused by the aspherical potential of the host galaxy contribute to the angular separation of the shells.

Although the material in the satellite galaxy spans a range of angular momenta and energies, a suitable first approximation to the contribution made by the intrinsic angular momentum to the relative angles of the caustics is the angular deflection of the center of mass of the satellite given the radial profile, $V(r)$, of the host galaxy potential and the total energy E and angular momentum ℓ of the satellite. Using the energy equation

$$E = \frac{1}{2}\dot{r}^2 + \frac{\ell^2}{2r^2} + V(r), \quad (4.20)$$

and conservation of angular momentum to change variables to the angle θ rather than the time t , the resulting differential equation is separable with solution

$$\Omega = 2\ell \int_{r_{\min}}^{r_{\max}} \frac{dr}{r^2 \sqrt{2(E - \frac{\ell^2}{2r^2} - V(r))}} \quad (4.21)$$

where r_{\min} is the pericenter distance and r_{\max} is the apocenter distance (for $E < 0$) or

infinite (for $E > 0$). As usual, r_{\min} and r_{\max} are the two positive roots of the equation

$$E = \frac{\ell^2}{2r^2} + V(r). \quad (4.22)$$

In a spherical potential, this approximation works well for the tidal stream and the first few caustics (Figure 4-6), which are at the largest radii and therefore the most likely to be seen. As discussed in the previous section, measurement of the line-of-sight velocity at the caustic surface constrains the angular momentum of the debris. $V(r)$ can be interpolated from the rotation curve of the host galaxy; this avoids introducing a full mass model of the host.

Additional angular separation between the caustics is due to deviations from spherical symmetry in the potential that cause the orbit to precess. To account for flattening of the potential, the orbit of the center of mass of the debris must be integrated numerically. This introduces new parameters: the amount of flattening and the angle of the flattened halo relative to the initial orbital plane, so that the interpolated potential $V(r)$ is replaced by the flattened version $V[r_q(\beta, \gamma)]$, where

$$r_q(\beta, \gamma) = \sqrt{x'(\beta, \gamma)^2 + y'(\beta, \gamma)^2 + z'(\beta, \gamma)^2/q^2} \quad (4.23)$$

and $\vec{r}' \equiv (x', y', z') = \mathcal{R}(\beta, \gamma)\vec{r}$ is the orbital position rotated into the plane of the potential. The variable q ranges from 0 to 1 and describes the flattening. Further tests will be needed to determine how well these parameters can be constrained by measurements of the shells, but β and γ can be estimated from the image and spectra of the host galaxy if they are available, and the center-of-mass integration is sufficient to predict the relative angles of the caustics (Figure 4-7).

4.8 Simulations with a live halo

It is possible that some of the effects we see may be different for simulations that use a live halo; that is, one that is comprised of self-gravitating particles and can respond to the influence of the satellite galaxy. [102] performed numerical experiments with

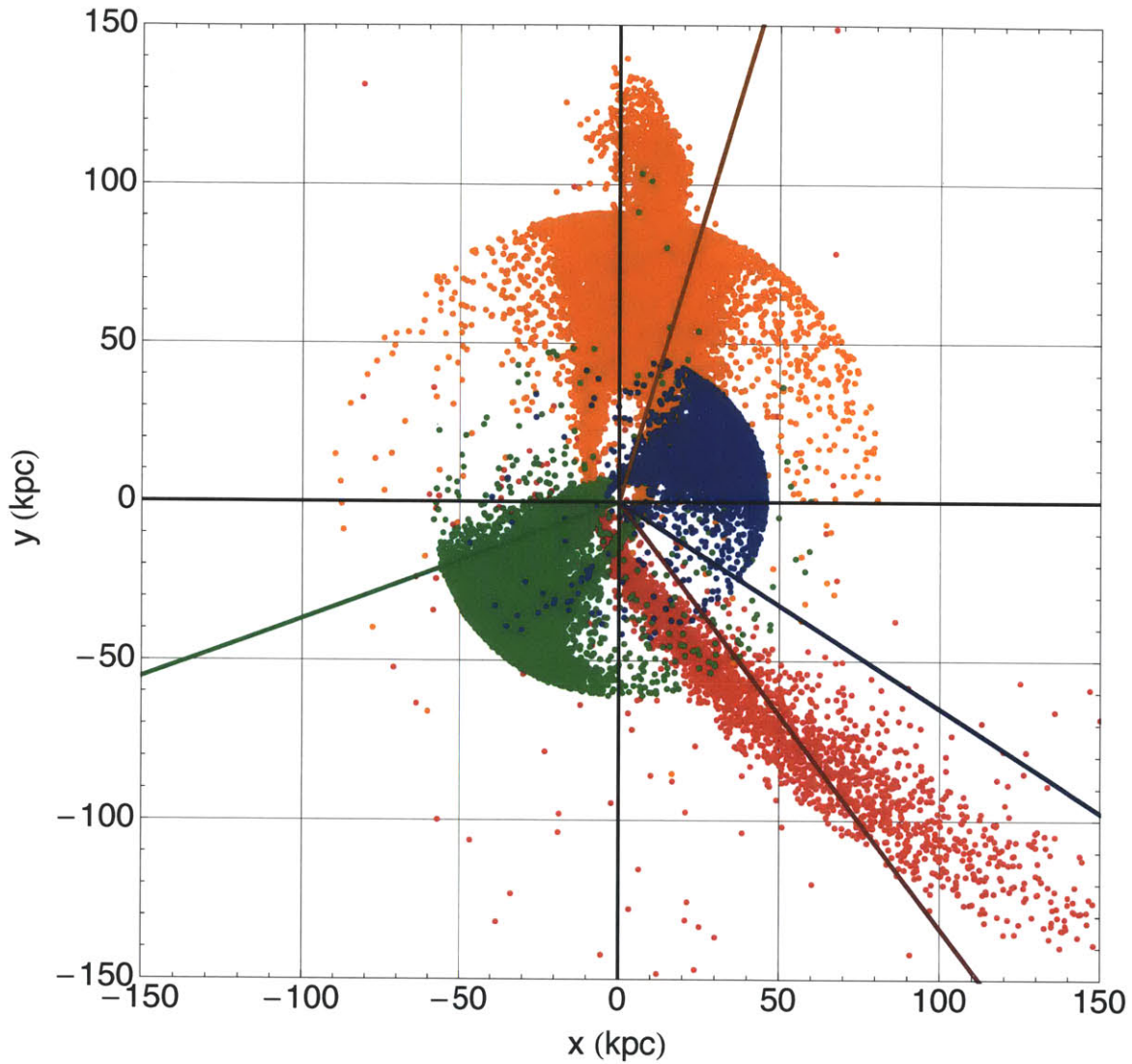


Figure 4-6: In a spherical potential, the scattering approximation can successfully predict the approximate relative angles of the tidal stream (red), first caustic (orange), and second caustic (green), although by the third caustic (blue) the additive errors in angle begin to dominate. Here the mass profile of the host is the isochrone potential. As discussed in the text, the second tidal stream (the orange material extending beyond the shell) is an indicator that two stripping episodes took place; here we consider only material stripped in the first episode.

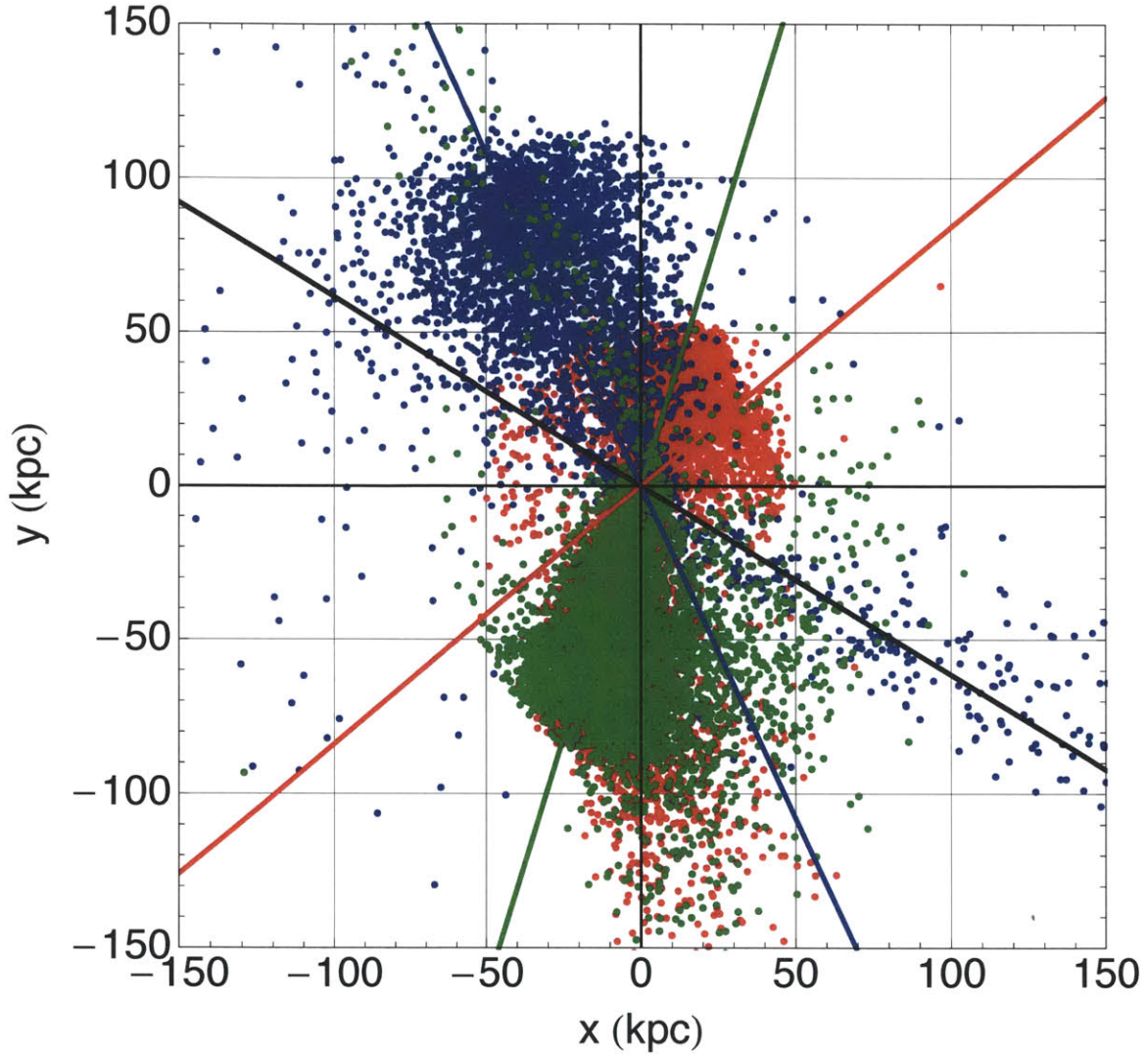


Figure 4-7: An N-body simulation of a minor merger in a flattened cored isothermal potential, $V(\vec{r}) = \frac{1}{2}v_0^2 \ln(R_c^2 + x^2 + y^2 + z^2/q^2) + V_0$, with lines showing the predicted angles of the debris at each time based on the center-of-mass orbit, shows that this technique successfully predicts the relative angles of caustics. The potential in this example had the parameters $v_0 = 450$ km/s, $V_0 = -0.4$ kpc² Myr⁻², $R_c = 8$ kpc, and $q = 0.5$, and is oriented as if the host galaxy were face-on (the x - y plane is the plane of the sky). The same parameters were used for the N-body simulation and the integration of the center of mass motion. Notice that the shells are no longer all equally sharp, thanks to the oscillation of the orbit in the z direction (along the line of sight).

live halos and found that for nearly radial encounters, the material in the satellite galaxy still forms phase-wrapped shells. They also found that depending on the density of the satellite galaxy, tidal heating and dynamical friction cause significant orbital decay that affects the locations and numbers of shells. Significantly, they show that the Chandrasekhar formula for dynamical friction is not a good model for the highly eccentric encounters that create shells: the “wake” of particles that the formula assumes is the primary source of friction is in fact subdominant to interactions with the core of the host galaxy for nearly radial orbits. Instead, restricted three-body methods must be used instead for analytical work.

The simulations of [102], limited by the computational capabilities at the time, used less than 2×10^5 total particles, whereas much larger simulations are possible today. More recent simulations of individual systems with live halos by [31] and [103], at significantly higher resolution, also show the formation of shells, but do not systematically study their properties. Based on this prior work, and on the shells observed in nature, we expect that the shell phenomenon is not destroyed by the inclusion of a live halo. Since the shape of the radial density profile is a direct result of the creation of the shells by phase-wrapping, we expect that the shells will still have the same profile even when a live halo is used.

We investigated the effect of a live halo using the self-consistent halo model of [104] known as the lowered Evans model. The Evans model [105] refers to the analytic phase space distribution function $f_{\text{Ev}}(E, L_z)$, a function of the energy E and the axial angular momentum L_z , that corresponds to the axisymmetric logarithmic potential

$$\Phi_{\text{axi}}(R, z) = -\frac{1}{2}v_0^2 \log(R_c^2 + R^2 + z^2/q^2) \quad (4.24)$$

where (R, φ, z) are cylindrical coordinates, v_0 is a characteristic velocity, and R_c the core radius. The oblateness parameter q is less than 1 for an oblate potential, 1 for a spherical potential, and greater than 1 for a prolate potential. The model is physical for values of q between $1/\sqrt{2}$ and about 1.08 [105]. As is well known, this potential corresponds to a mass distribution that diverges at large r so that the total

mass is infinite. As a result a self-consistent N-body representation of a halo with this potential, which is necessarily finite in mass and extent, cannot be constructed. The *lowered* Evans model addresses this problem by cutting off the phase space distribution f_{Ev} at $E = 0$:

$$f_{\text{LEv}} = \begin{cases} f_{\text{Ev}}(E, L_z) & E < 0 \\ 0 & E \geq 0 \end{cases}, \quad (4.25)$$

guaranteeing that all the matter in the resulting distribution is bound. In this sense the lowered Evans model is a generalization of the spherical King model. For this modified phase space distribution the potential is close to, but no longer exactly, Φ_{axi} ; the potential and spatial density (obtained by integrating f_{LEv}) must be solved for iteratively using the Poisson equation. The program GalactICs, which constructs a self-consistent N-body realization of the lowered Evans model [104], represents the potential and density with multipole expansions for the iterative solution and outputs the expansion coefficients of the final model along with the N-body representation. This allowed us to directly compare simulations with live and static versions of the same host halo, and isolate the effect of the host halo’s gravitational response.

To perform the N-body integrations, we used GADGET for the live halo and a serial tree code for the static halo. More information on both N-body integrators and on GalactICs is available in Appendix C. We set the integration parameters (softening, maximum timestep, tree opening angle) to be as consistent as possible given the differences in the two codes.

The effect of the gravitational response of the host is seen primarily in the effect of dynamical friction on the bound portion of the satellite (Figure 4-8, left panel), which reduces the successive apocenter distances of the bound mass remaining after each pericenter passage. However, it does not qualitatively change the phase space distribution of the unbound tidal stream, which still forms a tightly correlated caustic (Figure 4-8, right panel).

Both caustics are fit to obtain the physical parameters for comparison. First, κ was obtained by fitting a quadratic polynomial to the plot of $r - v_r$ in the region of

the caustic and using Equation 4.7. Then Equation 4.1 was fit to the density profile to obtain δr , r_c , and $f_0/\sqrt{2\pi\kappa}$. The fit parameters are summarized in Table 4.1. The radial density profile of the shell formed in the encounter with the live halo is both wider and taller than that formed in the static potential, but we obtain consistent results in both cases for r_c , κ , and f_0 . In this fit we did not allow r_c to vary as was done in Chapter 2, but the degree of variation in r_c along the caustic was comparable in the live and static cases and comparable to Caustic 1 in Chapter 2.

Parameter	Live halo	Static halo
Mass in shell ($10^8 M_\odot$)	1.6	1.1
δr , kpc	0.49 ± 0.04	0.33 ± 0.04
f_0 , Myr kpc $^{-2}$	14 ± 1	14 ± 2
κ , Myr 2 kpc $^{-1}$	251 ± 2	249 ± 3
r_c from density fit, kpc	26.46 ± 0.03	26.57 ± 0.04
r_c from phase space fit, kpc	26.4 ± 0.4	26.6 ± 0.5

Table 4.1: Comparison of fit parameters for a caustic in a live versus a static halo.

The larger normalization of the caustic in the live case indicates that more mass was stripped from the satellite at its first pericenter passage in the live halo than in the static case; in this case about 40 percent more. Mass loss can be larger in the live halo for two reasons. First, as the satellite passes pericenter it deepens the potential of the host galaxy and attracts more material from the host into the center. This increases the tidal force, leading to more stripping. Second, drag from dynamical friction can strip additional mass from the satellite’s edges when it is far from pericenter. The two processes may be distinguished by considering substructure in the density profile of the caustic. At pericenter, mass is stripped catastrophically and could come from anywhere in the satellite, which is entirely deformed by tides, so its initial spread in velocities is likely to be large. In contrast, dynamical friction strips material preferentially from the outskirts of the satellite where it will have near-zero velocity relative to the satellite’s center of mass. So the material stripped at pericenter will be hotter than that stripped by dynamical friction, and a caustic containing both kinds of material will exhibit a thinner peak on top of a wider structure. In future, the contributions to mass loss from each of these processes can also be estimated

analytically.

The additional width indicates that the energy spread of the stripped material is larger in the case of the live halo. The larger energy spread can be traced back to the same root causes as the change in the amount of mass stripped. If a larger amount of material is stripped at pericenter, the energy spread of the material will be wider, since the increased tidal force can liberate material that was more tightly bound to the satellite. The larger energy spread will lead to a broader caustic. In addition, the satellite will heat up in response to the drag from dynamical friction, further adding to the energy spread of stripped material. Both these effects are evidence of the host galaxy's response to the merger, but they are not large enough to completely destroy the phase space structure that produces the shells.

The radii of the shells are consistent with one another. This is evidence that the differences in the shells between the live and static cases for this time-snapshot are due to the pericenter interaction rather than dynamical friction, since significant drag from dynamical friction would also have reduced the apocenter distances of the material, leading to a discrepancy in r_c . Additionally, consistent results are obtained for f_0 , indicating that the phase space distribution of the satellite was not significantly altered before the first pericenter. Shells that develop from later pericenter passages, on the other hand, could exhibit both effects as friction has more time to act.

4.9 Future work

In order to apply this technique to real data, the surface-brightness measurements must be deprojected to obtain the radial density profile of the material. It is by no means clear whether this is possible in practice. On one hand, even if a simple geometry like the one suggested in Section 4.5 is sufficient to represent one of the shells, there is a geometric degeneracy between the inclination angle θ_c of the cone and the width δr of the caustic. This degeneracy may be partially broken by the fact that the surface brightness profile changes differently when θ_c is changed than when δr is manipulated, but further work is necessary to determine how correlated

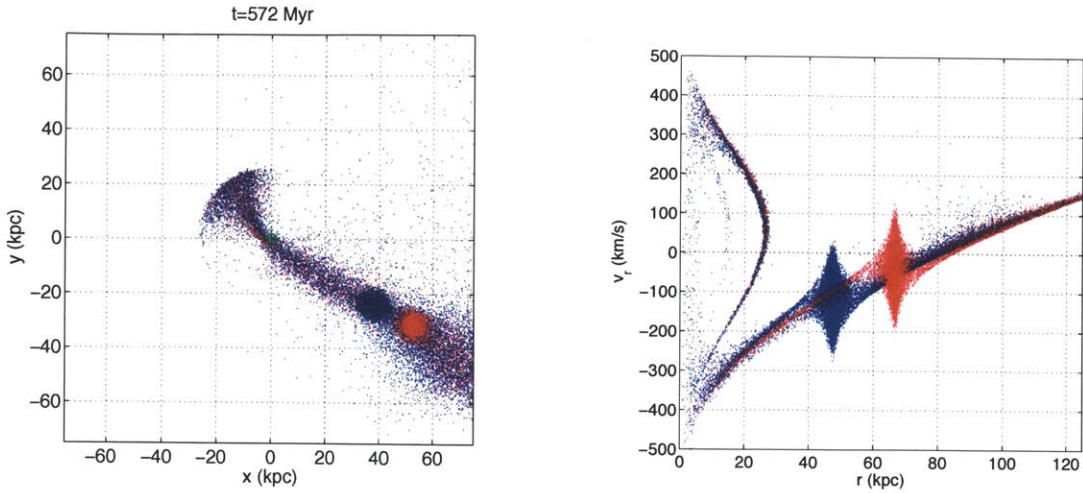


Figure 4-8: Comparing the result of a simulation with a live host halo (blue points) with one using a static halo (red points), we see that dynamical friction affects the orbit of the material still bound in the satellite, but does not disrupt the streaming of the unbound material. The relative sizes of the two bound structures differ because the particles in the live halo were plotted using larger points to highlight them; in fact a comparable amount of mass is stripped from the satellite in both cases.

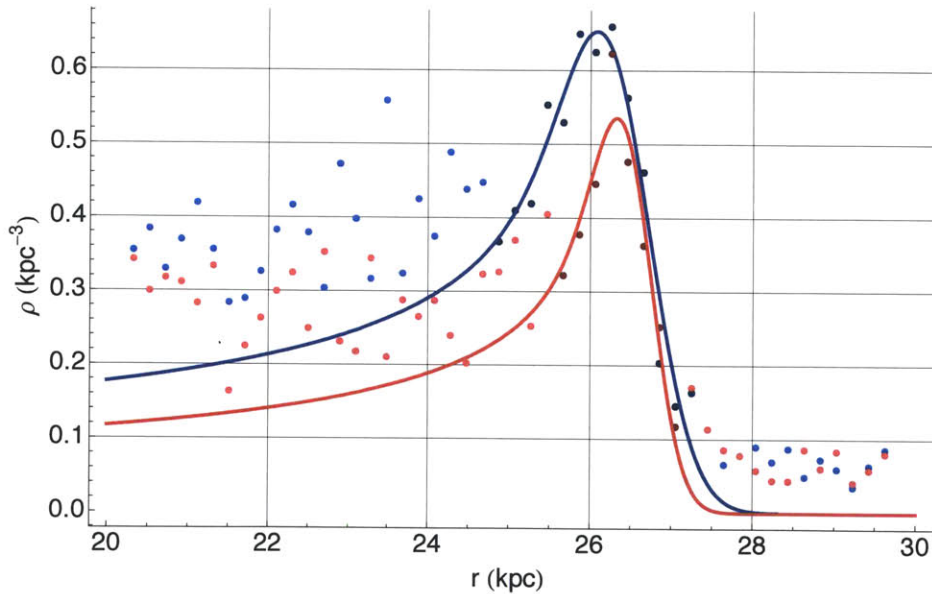


Figure 4-9: The density profile of the caustic from the live halo (blue points) is wider than that in the static halo (red points), but both can be fit with the same density profile (blue and red lines). The points used in the fit are shown in a darker shade.

these parameters are. There is also a possibility that the simple model suggested here, which takes r_c to be a constant, is not sufficient to describe the geometry of the system. We saw in Section 2.3.3 that the caustic radius varies with angle in the N-body model of those shells. Failing to account for this variation may significantly affect the quality of the fit. In Chapter 5, we will implement a fitting procedure for the surface brightness and test it on various sample data to determine the practical exigencies related to deprojection. Only after this is done will it be possible to attempt a fit of real surface brightness data.

4.10 Conclusion

Thanks to the high symmetry of nearly radial minor mergers, we have developed a new framework that can constrain several parameters of the merger based only on images of the shells. Measurements of the line-of-sight velocities improve or break degeneracies in those constraints, and measurements of the relative angles of multiple shells can further constrain the total angular momentum in the system. These results are equally applicable to spherically symmetric potentials (examples in Figures 4-2 and 4-3), axisymmetric potentials (example in Figure 4-7), and realistic mass models of galaxies that include a disk, bulge and halo (example in Figure 4-1). Our conclusions also hold when the gravitational response of the host halo is included. In fact, the relative angles of the shells may be able to restrict the magnitude of departures from spherical symmetry in the total potential. The approach described in this paper also *requires no full N-body simulations* to produce these constraints, but instead directly links particular types of additional data with particular refinements of the constraints using the minimum number of parameters. Therefore, in contrast with other techniques for modeling merging galaxies, this technique is not hampered by incomplete exploration of the parameter space. We expect this approach, although limited, to be of use in studying minor mergers for a wide range of galaxies of different morphological types for which minimal kinematic information is available, extending tests of hierarchical accretion beyond the study of Local Group galaxies.

Chapter 5

Constraints on halo mass profiles from surface brightness measurements of tidal shells

5.1 Introduction

In Chapter 4, we showed that some constraints on the mass profiles of interacting galaxies can be obtained by fitting the surface brightness profiles of tidal shells produced by the merger. The crucial step in obtaining these constraints involves the deprojection of the surface brightness profile; that is, determining the angle at which the shell is oriented along the line of sight. In this chapter, we discuss the technical and practical aspects of deprojection. We first derive the equations necessary to model the surface brightness profile (Section 5.2.1). We then implement the fitting procedure and test its ability to recover the input parameters given a surface brightness map generated from an N-body realization of the model (Section 5.2.2). This test also identifies possible systematic degeneracies between parameters in the model. Finally, we test the same fitting procedure on the N-body models of the two tidal shells in M31 (Section 5.3). These structures were produced via tidal stripping rather than directly sampled from the model, so this test determines how well we expect

our model to fit real data and whether additional parameters should be added. In Section 5.4 we discuss possible sources of real data, and in Section 5.5 we summarize the work described in this chapter.

5.2 Deprojection of two-dimensional measurements

5.2.1 Equations for cone approximation

Tidal shells are formed by material stripped from a small satellite galaxy on a nearly radial orbit around a large host galaxy. The stripping of material happens nearly instantaneously each time the satellite passes pericenter. The newly unbound material, whose angular momentum and energy distributions have been scrambled by the sudden rapid changes in the gravitational potential, then fans out into a shell as it heads toward apocenter. The edge of the shell can be approximated by a section of a sphere, its density by Equation (3.37), and its angular distribution by a cone. In this section the equations necessary to fit this model to surface brightness measurements are presented.

Surface brightness profile

A cone of height h and opening angle α , with its point at the origin and its axis of symmetry along the z axis has the parametric equations (in Cartesian coordinates)

$$x = (h - u) \cos \vartheta \tan \alpha \quad (5.1)$$

$$y = (h - u) \sin \vartheta \tan \alpha \quad (5.2)$$

$$z = h - u \quad (5.3)$$

where the parameter $u \in [0, h]$ describes the distance from the base of the cone and the parameter $\vartheta \in [0, 2\pi]$ describes the azimuthal location on the cone. If the cone is rotated so that its axis of symmetry points along the unit vector

$$\hat{n} = \sin \theta_c \cos \phi_c \hat{x} + \sin \theta_c \sin \phi_c \hat{y} + \cos \theta_c \hat{z} \quad (5.4)$$

then the equations become

$$\begin{aligned}
x &= (h - u) \{ \cos \vartheta [1 - \cos^2 \phi_c (1 - \cos \theta_c)] \tan \alpha \\
&\quad - \sin \vartheta [\cos \phi_c \sin \phi_c \tan \alpha (1 - \cos \theta_c)] \\
&\quad + \sin \theta_c \cos \phi_c \} \tag{5.5}
\end{aligned}$$

$$\begin{aligned}
y &= (h - u) \{ -\cos \vartheta [\cos \phi_c \sin \phi_c \tan \alpha (1 - \cos \theta_c)] \\
&\quad + \sin \vartheta [1 - \sin^2 \phi_c (1 - \cos \theta_c)] \tan \alpha \\
&\quad + \sin \theta_c \cos \phi_c \} \tag{5.6}
\end{aligned}$$

$$z = (h - u) [-\cos (\vartheta - \phi_c) \sin \theta_c \tan \alpha + \cos \theta_c] \tag{5.7}$$

To determine z_{\min} and z_{\max} as a function of x and y , we solve the system of the x and y parametric equations to obtain the parameters u and ϑ . The system can have zero, one, or two solutions depending upon the values of x and y : zero solutions for lines of sight that do not intersect the cone, one solution for lines of sight that intersect the cone once, and two solutions for lines of sight that intersect the cone twice. In cases where there is one solution, the other limit can safely be taken to be $\pm\infty$, where the plus sign is taken if the bulk of the cone is in front of the intersection point (the solution is the lower bound of z), and the minus sign is taken if the cone is behind the intersection (the solution is the upper bound of z). The system can always be solved analytically and we present the solution here. First we will define some auxiliary quantities to make the notation simpler. We extract the θ_c , ϕ_c , and α dependence of Equations (5.5-5.6) into coefficients that need only be calculated once for a given

cone:

$$A_x = \tan \alpha [1 - \cos^2 \phi_c (1 - \cos \theta_c)] \quad (5.8)$$

$$B_x = \tan \alpha \cos \phi_c \sin \phi_c (1 - \cos \theta_c) \quad (5.9)$$

$$C_x = \sin \theta_c \cos \phi_c \quad (5.10)$$

$$A_y = B_x \quad (5.11)$$

$$B_y = \tan \alpha [1 - \sin^2 \phi_c (1 - \cos \theta_c)] \quad (5.12)$$

$$C_y = \sin \theta_c \sin \phi_c, \quad (5.13)$$

so that Equations (5.5-5.6) become

$$x = (h - u) (A_x \cos \vartheta - B_x \sin \vartheta + C_x) \quad (5.14)$$

$$y = (h - u) (-A_y \cos \vartheta + B_y \sin \vartheta + C_y). \quad (5.15)$$

We can solve for ϑ by dividing the two equations, noting that if $x = 0$ or $y = 0$ then we are at the point of the cone ($u = h$) and ϑ is degenerate. We obtain an equation for ϑ in terms of the ratio $\eta \equiv x/y$:

$$\mathcal{S} \sin \vartheta - \mathcal{C} \cos \vartheta + \mathcal{K} = 0, \quad (5.16)$$

where

$$\mathcal{C} \equiv A_x + \eta A_y \quad (5.17)$$

$$\mathcal{S} \equiv B_x + \eta B_y \quad (5.18)$$

$$\mathcal{K} \equiv -C_x + \eta C_y. \quad (5.19)$$

Equation (5.16) can be recast as a quadratic equation for either $\sin \vartheta$ or $\cos \vartheta$. To avoid using inverse trigonometric functions (and the associated difficulties in choosing

the right branch) we simply solve for both and use them in the rest of the solution:

$$(\cos \vartheta)_{\pm} = \frac{\mathcal{K}\mathcal{C} \pm \mathcal{S}\mathcal{D}}{\mathcal{C}^2 + \mathcal{S}^2} \quad (5.20)$$

$$(\sin \vartheta)_{\pm} = \frac{\mathcal{K}\mathcal{S} \pm \mathcal{C}\mathcal{D}}{\mathcal{C}^2 + \mathcal{S}^2} \quad (5.21)$$

where \mathcal{D}^2 is the discriminant

$$\mathcal{D}^2 \equiv \mathcal{C}^2 + \mathcal{S}^2 - \mathcal{K}^2. \quad (5.22)$$

As usual, if $\mathcal{D}^2 < 0$ the point is outside the cone and the equation has no real roots; otherwise it has two real roots. When two solutions exist for ϑ , one or both of them may lead to a value of u outside its allowed range.

Having determined ϑ either Equation (5.5) or Equation (5.6) can be used to determine the quantity $h - u$ that is necessary to find the limits on z :

$$\begin{aligned} (h - u)_{\pm} &= \frac{x}{A_x (\cos \vartheta)_{\pm} - B_x (\sin \vartheta)_{\pm} + C_x} \\ &= \frac{y}{-A_y (\cos \vartheta)_{\pm} + B_y (\sin \vartheta)_{\pm} + C_y} \end{aligned} \quad (5.23)$$

The quantity $h - u$ should be in the range $(0, h]$ —if it is not, that value of ϑ is discarded as a root and one of the limits in z goes to $\pm\infty$.

Finally, the limits z_{\pm} are given by plugging in the valid solutions for u and ϑ :

$$z_{\pm} = (h - u)_{\pm} \left\{ \cos \theta_c - \sin \theta_c \tan \alpha \left[(\cos \vartheta)_{\pm} \cos \phi_c + (\sin \vartheta)_{\pm} \sin \phi_c \right] \right\} \quad (5.24)$$

By inspection, we see that z_{\min} is not always equal to z_- and z_{\max} is not always z_+ since both $(\cos \vartheta)_{\pm}$ and $(\sin \vartheta)_{\pm}$ can take any sign: the roots must be compared and the smaller assigned to z_{\min} . If one root is out of range in u , its out-of-range z value can still be calculated (for this work the height h of the cone is arbitrary because the density function effectively cuts off the cone along a spherical segment) and compared to the value of the in-range root to determine whether the out-of-range limit is z_{\min} or z_{\max} .

Partial derivatives

In constructing a fitting routine, the partial derivatives of Σ with respect to the projection parameters α, θ_c, ϕ_c and to the density profile parameters $r_c, \delta r, \kappa, f_0$ may be needed although not for the fitting routine we use in this work. In the following, we define the shorthand

$$\begin{aligned} [\cos(\vartheta - \phi_c)]_{\max} &\equiv (\cos \vartheta)_{\max} \cos \phi_c + (\sin \vartheta)_{\max} \sin \phi_c \\ [\sin(\vartheta - \phi_c)]_{\max} &\equiv (\sin \vartheta)_{\max} \cos \phi_c - (\cos \vartheta)_{\max} \sin \phi_c \end{aligned}$$

where $(\cos \vartheta)_{\max}$ and $(\sin \vartheta)_{\max}$ are taken to be the roots that lead to the value of z_{\max} , and likewise

$$\begin{aligned} [\cos(\vartheta - \phi_c)]_{\min} &\equiv (\cos \vartheta)_{\min} \cos \phi_c + (\sin \vartheta)_{\min} \sin \phi_c \\ [\sin(\vartheta - \phi_c)]_{\min} &\equiv (\sin \vartheta)_{\min} \cos \phi_c - (\cos \vartheta)_{\min} \sin \phi_c \end{aligned}$$

for the roots leading to the value of z_{\min} . A similar notation is used to denote the appropriate root of $(h - u)$. We also use the shorthand

$$\rho(z_{\max}) \equiv \rho(\sqrt{x^2 + y^2 + z_{\max}^2}) \quad (5.25)$$

in the following, since x and y are understood to be constant. With these definitions, the derivatives with respect to the projection parameters are:

$$\begin{aligned} \frac{\partial \Sigma}{\partial \alpha} &= \frac{\sin \theta_c \sec^2 \alpha}{\Upsilon} \{ \rho(z_{\min}) [\cos(\vartheta - \phi_c)]_{\min} (h - u)_{\min} \\ &\quad - \rho(z_{\max}) [\cos(\vartheta - \phi_c)]_{\max} (h - u)_{\max} \} \quad (5.26) \end{aligned}$$

$$\begin{aligned} \frac{\partial \Sigma}{\partial \theta_c} &= \frac{1}{\Upsilon} \{ \rho(z_{\min}) [(h - u)_{\min} \sin \theta_c + [\cos(\vartheta - \phi_c)]_{\min} \cos \theta_c \tan \alpha] \\ &\quad - \rho(z_{\max}) [(h - u)_{\max} \sin \theta_c + [\cos(\vartheta - \phi_c)]_{\max} \cos \theta_c \tan \alpha] \} \quad (5.27) \end{aligned}$$

$$\begin{aligned} \frac{\partial \Sigma}{\partial \phi_c} &= \frac{\sin \theta_c \tan \alpha}{\Upsilon} \{ \rho(z_{\min}) (h - u)_{\min} [\sin(\vartheta - \phi_c)]_{\min} \\ &\quad - \rho(z_{\max}) (h - u)_{\max} [\sin(\vartheta - \phi_c)]_{\max} \} \quad (5.28) \end{aligned}$$

The derivatives with respect to the profile parameters are all integrals of the derivatives of the density, of the form

$$\frac{\partial \Sigma}{\partial \pi_i} = \frac{1}{\Upsilon} \int_{z_{\min}}^{z_{\max}} \frac{\partial \rho}{\partial \pi_i} dz \quad (5.29)$$

for a given parameter π_i , since the limits of the integral and the variable being integrated over do not depend on any of the π_i and the function ρ , although defined as a piecewise function in Equation 3.37, is continuous over the entire integration range. To compactly write derivatives of ρ with respect to the parameters, we expand the definition of \mathcal{B} to include other Bessel functions:

$$\mathcal{B}_n(u) \equiv \begin{cases} \frac{\pi}{2} [\mathcal{I}_{(2n+1)/4}(u) + \mathcal{I}_{-(2n+1)/4}(u)] & r \leq r_c \\ (-1)^{n+1} \frac{\pi}{2} [\mathcal{I}_{(2n+1)/4}(u) - \mathcal{I}_{-(2n+1)/4}(u)] & r > r_c \end{cases} \quad (5.30)$$

with $u \equiv (r - r_c)^2 / 4\delta r^2$ as before. The definition used in Equation 3.37 is equivalent to \mathcal{B}_0 in this new notation. With this simplification, the derivatives $\partial \rho / \partial \pi_i$ are:

$$\frac{\partial \rho}{\partial r_c} = \frac{f_0}{\sqrt{2\pi\kappa}} e^{-u} \frac{\sqrt{|r - r_c|}}{r - r_c} \left\{ u [2\mathcal{B}_0(u) - \mathcal{B}_2(u) - \mathcal{B}_1(u)] - \frac{1}{2}\mathcal{B}_0(u) \right\} \quad (5.31)$$

$$\frac{\partial \rho}{\partial \delta r} = \frac{f_0}{\sqrt{2\pi\kappa}} e^{-u} \sqrt{|r - r_c|} \frac{u}{\delta r} [2\mathcal{B}_0(u) - \mathcal{B}_2(u) - \mathcal{B}_1(u)] \quad (5.32)$$

$$\frac{\partial \rho}{\partial f_0} = \frac{\rho}{f_0} \quad (5.33)$$

$$\frac{\partial \rho}{\partial \kappa} = -\frac{\rho}{2\kappa} \quad (5.34)$$

5.2.2 Example fit and limitations of the model

Degeneracies in the proposed model may lead to multiple minima in the merit function of a fitting routine. In this case there are three partial to complete degeneracies between parameters that are introduced by the projected geometry. To identify and illustrate them, we perform a fit to sample data generated from the model, comparing the fitted parameters to their true values. We will also discuss the identifiability of the parameters and the reliability of the fit, including the best method for estimating

the confidence intervals on the parameter estimates.

Sample data

To generate the sample data, we use an N-body representation of the mass in the shell. The N-body representation is generated as follows. First, the r -coordinates of the particles are randomly sampled from the scaled mass distribution function (the CDF) of the caustic model (which depends on the parameters r_c , δr , and the normalization $f_0/\sqrt{2\pi\kappa}$). We use $N_p = 20000$ for the N-body realization. Then $\cos\theta$ is chosen for each particle from the uniform distribution on $[\cos\alpha, 1)$ and ϕ is chosen from the uniform distribution on $[0, 2\pi)$. This distributes the particles in a cone with the correct opening angle and radial distribution. The coordinates are converted to Cartesian and rotated to the appropriate orientation angles θ_c and ϕ_c . The final N-body realization used the parameters listed in Table 5.1, and is shown in Figure 5-1.

To imitate measurements of the surface brightness Σ , the FiEstAS density estimator (Section A.2) is used to integrate the number density of particles in the N-body realization along various lines of sight, each defined by a square “pixel” of side length Δ in the plane of the sky. For the tests, we define the surface brightness in pixel i to be

$$\Sigma_i \equiv \frac{N_i}{\Delta^2} \quad (5.35)$$

where N_i refers to the number of bodies in the volume along the lines of sight spanned by pixel i . The surface brightness is considered to be evaluated at the center of a pixel. This calculation has an associated error that results from Poisson fluctuations in N_i :

$$\sigma_i \equiv \sqrt{N_i}/\Delta^2. \quad (5.36)$$

Δ is chosen so that $\sigma_i/\Sigma_i < 0.3$ over most of the cone (i.e., $N_i > 10$) and less than 10 percent for the crucial pixels nearest to the caustic radius ($N_i > 100$). For the fit, we select only the pixels with enough resolution that the error is less than the measured value; that is, those for which the density could actually be estimated. The final set

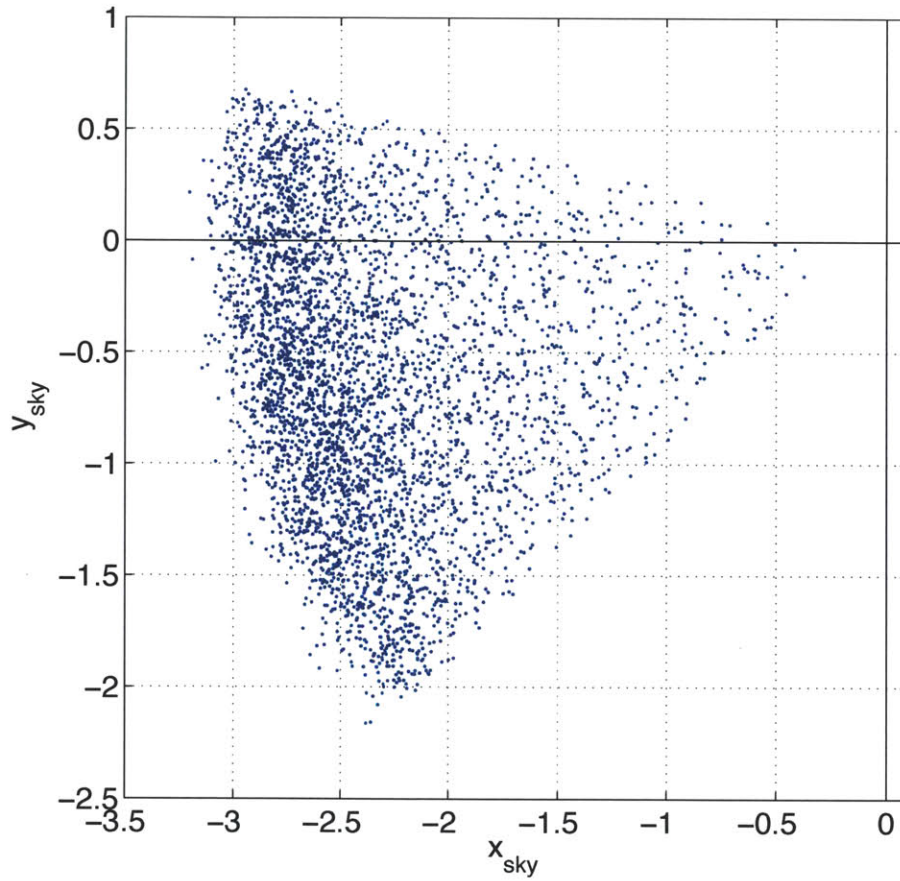


Figure 5-1: N-body realization of the model described in the previous section. One-fifth of the particles are plotted. The parameters used to generate the realization are given in Table 5.1.

of sample data, and the error bars, are shown in Figure 5-2.

Table 5.1: Parameters for sample data. The choice of units is arbitrary for this test data, but dimensions are given in the table.

Parameter	Value	Dimension
r_c	3.0	[L]
δr	0.1	[L]
$f_0/\sqrt{2\pi\kappa}$	$1/\sqrt{2\pi}$	$[\text{L}]^{-4}$
α, ϕ_c, θ_c	0.5, 3.4, 1.5	radians
N_p	2×10^4	particles
Δ	0.1	[L]
N_{pix}	544	data points to fit

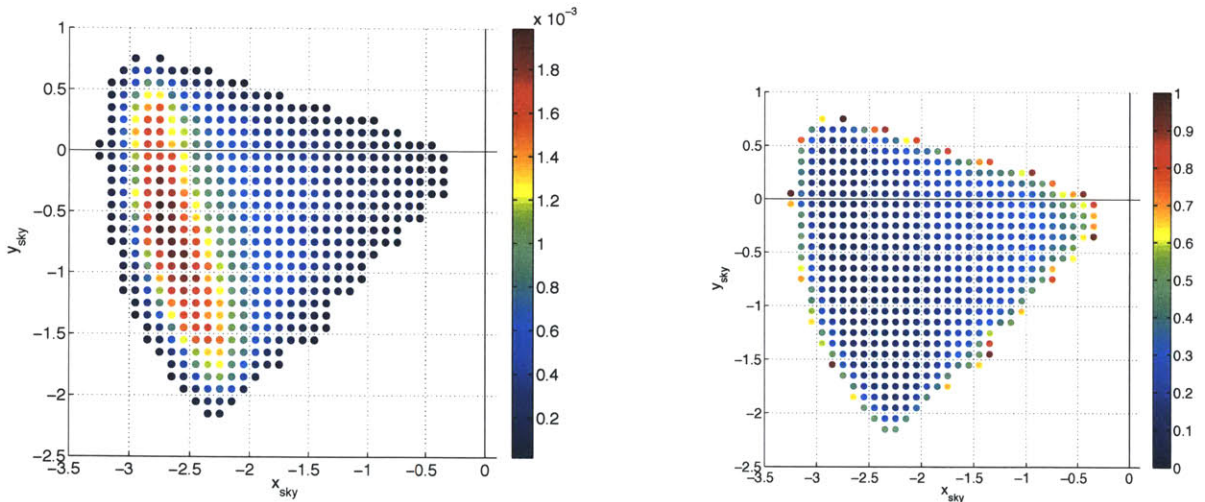


Figure 5-2: Sample data used to test the fitting program. On the left, the color bar shows the magnitudes of the measured “surface brightness” (definition is in the text). On the right, it shows the relative error.

Fitting algorithm

To perform the fit, we used the implementation by [106] of the downhill simplex method [107] to minimize the chi-squared function

$$\chi^2 \equiv \sum_{i=1}^{N_{\text{pix}}} \left(\frac{\hat{\Sigma}_i - \Sigma(\vec{x}_i, \vec{a})}{\sigma_i} \right)^2 \quad (5.37)$$

by manipulating the fit parameters $\vec{a} \equiv \{r_c, \delta r, f_0/\sqrt{2\pi\kappa}, \alpha, \phi_c, \theta_c\}$. This definition of χ^2 compares the measured values $\hat{\Sigma}_i$ of the surface brightness at sky positions \vec{x}_i to the model $\Sigma(\vec{x}_i, \vec{a})$ evaluated at the sky positions of the data points for the current fit parameters. The data points are weighted by the associated Poisson error σ_i . To improve the performance of the fit, we scale all the parameters $\vec{a} \rightarrow \vec{a}'$ by twice the initial guesses so their absolute values are of order 1 to start. If an initial guess was particularly bad and the absolute value of the parameter ranges far from 1, the parameters are rescaled to values near unity before the fit proceeds further.

The downhill simplex method was chosen because it is particularly well suited to understanding the behavior of the fit: it requires a fast method for evaluating χ^2 at many different points in the parameter space, starting values for the parameters, and sizes for the initial perturbation of each parameter. Although it converges much more slowly than directed-descent methods like Levenberg-Marquardt, it doesn't require computation of the first derivatives of χ^2 , which are problematic for our model since they all go to zero outside the area of the projected cone. The slow convergence has the additional advantage that watching the algorithm converge can help build intuition about possible degeneracies in the parameter space, since the convergence tends to slow down and stall around these degeneracies. For more information on the simplex method, see Appendix C.

The fit converges to the input parameters, although sometimes the fitting algorithm terminates prematurely at a false minimum. This possibility is indicated by the large value of χ^2 per degree of freedom, and means that the fit should be restarted with the final parameters to reset the search radius. A representative fit trajectory, including one restart, is shown in Figure 5-3 for initial guesses chosen by referring to Figure 5-2. In particular, the maximum measured surface brightness was the initial guess for the normalization, guesses for r_c , δr , α , and ϕ_c were determined by eye from the plot, and θ_c was started at about $\pi/2$ since the shell is fairly well-defined on the sky. The initial guesses, final recovered parameters, and 68% confidence intervals are compared in Table 5.2. The best-fit model is within one or two error bars of the data for nearly the entire data set (Figure 5-5), which is as expected since the sample data

were drawn from a random representation of the model.

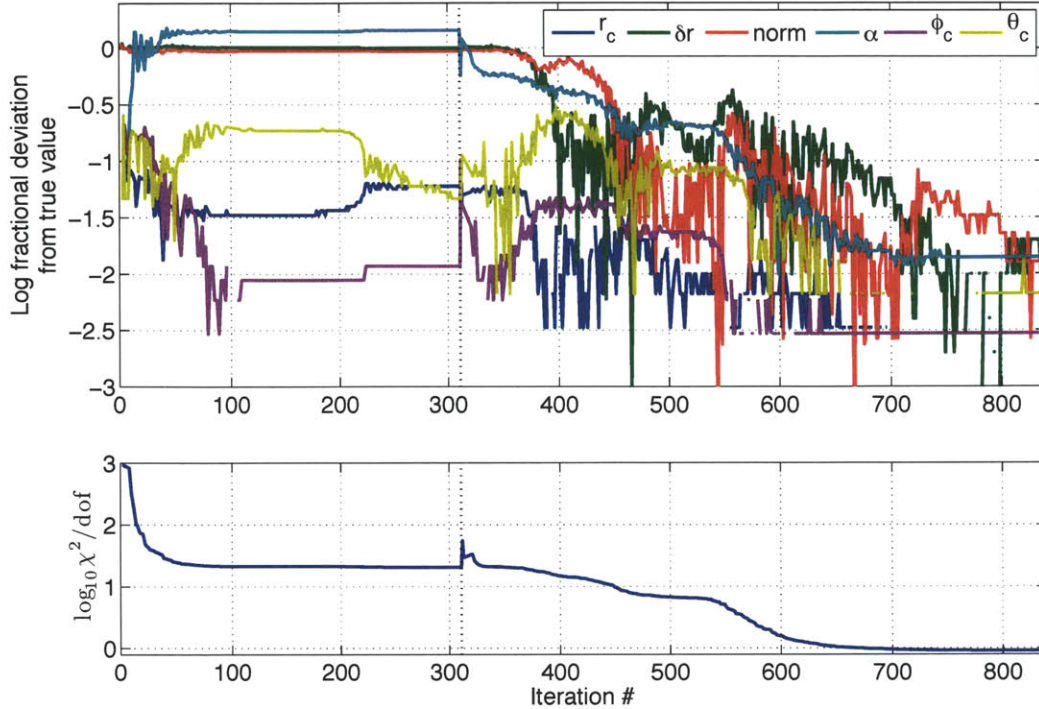


Figure 5-3: Example of convergence of the fitted parameters for the sample data shown in Figure 5-2. The dashed line indicates the restart of the fit after a false minimum. A large value of χ^2 per degree of freedom, as shown in the bottom panel, indicates that a false minimum may have been identified and the fit algorithm should be restarted to check.

Figure 5-3 shows that the last parameters to converge for this particular fit trajectory are the caustic width δr and the normalization, which are the most difficult parameters to fit. The fastest parameter to converge (and one of the easiest to fit) is the angle ϕ_c , which is defined in the plane of the sky and therefore very well constrained by the image in the first place. α and θ_c are adjusted in order to get the final convergence of δr and the normalization. The common behavior of some groups of parameters suggests the existence of correlations between them. To identify the correlations we study the shape of χ^2 around its minimum.

Parameter	Starting Guess	Starting Search Length	Final Value	Standard Error	True Value
r_c	2.7	0.1	3.00	0.03	3.0
δr	0.2	0.01	0.10	0.03	0.1
$f_0/\sqrt{2\pi\kappa}$	1.0×10^{-3}	0.3×10^{-3}	1.0×10^{-3}	0.3×10^{-3}	0.997×10^{-3}
α	0.55	0.314	0.493	0.008	0.5
ϕ_c	3.8	0.314	3.41	0.01	3.4
θ_c	1.57	0.314	1.51	0.06	1.5

Table 5.2: Summary of fit to sample data. Standard error is calculated using Equation (5.41). For error contours for correlated parameters, see Figure 5-4.

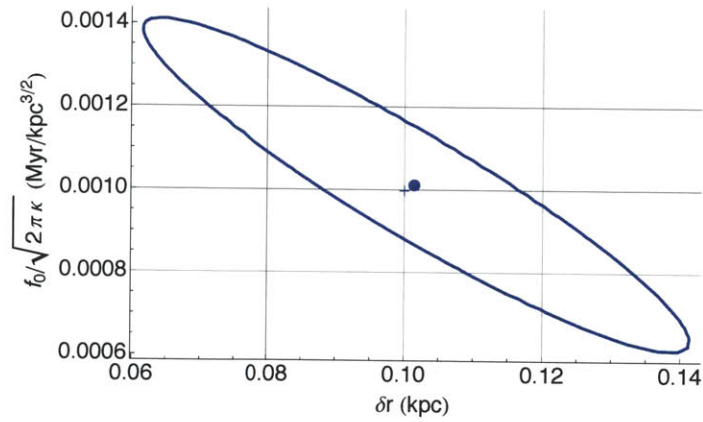


Figure 5-4: Estimates and error contours for correlated parameters δr and $f_0/\sqrt{2\pi\kappa}$. The cross indicates the true values of the parameters; the ellipse shows the 68% confidence interval.

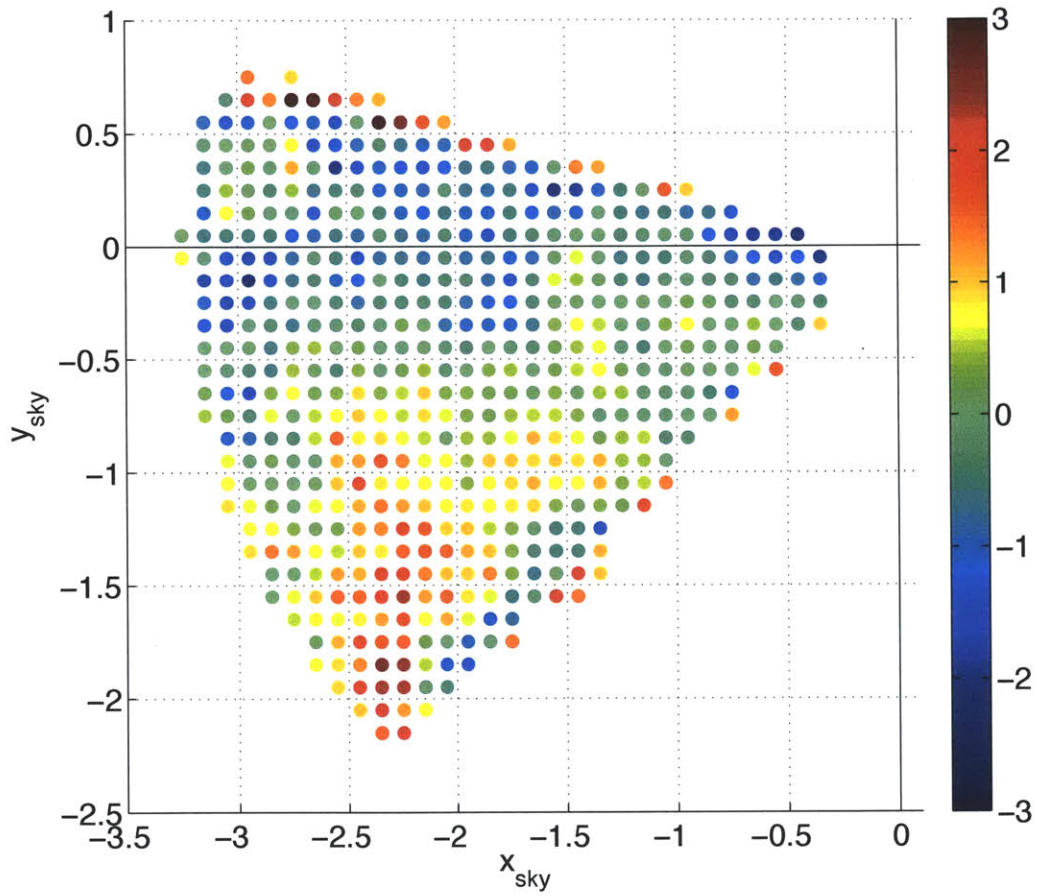


Figure 5-5: Fit residuals scaled by the error on the data points for the best-fit model. The fit is within a few sigma of the data points for nearly the whole fitted region, except at the edges where the surface brightness is least well-determined.

Identifiability and reliability of the parameter estimates

The question of the correlation of parameters in a fitted model is referred to in the statistical literature as the identifiability problem; the related question of how well the fit can recover certain parameters is called reliability. Both questions can be answered using some tools borrowed from the techniques of maximum-likelihood estimation (MLE).

In MLE, the best-fit parameters are identified by maximizing the log-likelihood \mathcal{L} , which is analogous to minimizing χ^2 with the identification $\mathcal{L} \equiv -\ln \chi^2$. Information about the shape of the minimum in MLE comes in the form of the Fisher information matrix, which corresponds to the Hessian of the χ^2 function evaluated at the scaled best-fit point \vec{a}'_0 :

$$\mathcal{I}_{ij} \equiv - \left. \frac{\partial^2 \mathcal{L}(\vec{a}')}{\partial a'_i \partial a'_j} \right|_{\vec{a}'_0} = \frac{1}{\chi^2(\vec{a}'_0)} \left. \frac{\partial^2 \chi^2(\vec{a}')}{\partial a'_i \partial a'_j} \right|_{\vec{a}'_0} \equiv \frac{H_{ij}(\vec{a}'_0)}{\chi^2}, \quad (5.38)$$

since we are presumably at the minimum of the χ^2 function where its first derivative is zero. For the fits in this work, the Hessian was calculated by first-order finite differencing with a relative step length of 1% in each scaled parameter. We calculate

The Hessian approximates the shape of the χ^2 function near \vec{a}'_0 . Taking a Taylor series of χ^2 about \vec{a}'_0 gives

$$\chi^2(\vec{a}') \approx \chi^2(\vec{a}'_0) + \frac{1}{2} \delta a'_i H_{ij}(\vec{a}'_0) \delta a'_j + \mathcal{O}[(\delta a')^3] \quad (5.39)$$

where $\vec{\delta a}' \equiv \vec{a}' - \vec{a}'_0$ is the deviation from the best-fit point. This approximation tells us how the eigensystem of the Hessian can be used to better understand the fit and the recovered parameters. Most simply, since we are at a minimum, the surface must be concave up in all directions in the parameter space; that is, the Hessian matrix must be positive-definite. This test can be used to confirm that the fit algorithm has stopped at a minimum, rather than at a maximum or saddle point.

The eigenvectors of the Hessian describe the correlations between the parameters near the minimum and determine which ones are independently identifiable. For a

system with fully independent parameters, the matrix of eigenvectors will be diagonal. Correlations between parameters appear as off-diagonal terms in the matrix of eigenvectors. The matrix of eigenvectors for the sample data (Figure 5-6, left panel) shows that the fit parameters are nearly independent except for δr and the normalization, which are extremely correlated: the eigenvectors in this subspace are very nearly $(\hat{\delta r} \pm \text{norm})/\sqrt{2}$. This is to be expected, since the ability to determine δr is limited by the spacing of the sample points in the data set Δ (Table 5.1), which is of order δr . As Δ is changed, the values recovered for δr and $f_0/\sqrt{2\pi\kappa}$ can both change, but the total mass around the peak of the caustic (that is, the number of bodies making up the realization) does not. The total mass is proportional to the product of the two quantities, so that the combination $p_+ \propto (f_0/\sqrt{2\pi\kappa} + \delta r)$ and the orthogonal combination $p_- \propto (f_0/\sqrt{2\pi\kappa} - \delta r)$ should form independent parameters in the estimate. The fluctuations in p_+ should be directly related to Poisson fluctuations in the total mass and scale as $1/\sqrt{N}$, independent of Δ . The fluctuations in p_- are sensitive to the ratio between the width and height of the peak instead of its product, and this quantity changes with Δ because Poisson fluctuations in the bins along the peak's leading edge will tend to widen the peak as Δ increases. So fluctuations in p_- should scale as Δ/\sqrt{N} .

An alternative measure of correlations between parameters is the so-called correlation matrix R , which compares the scales over which χ^2 changes (given by the inverse of the Hessian) in the cross-derivative directions to the scale in the parameter directions:

$$R_{ij} \equiv \frac{H_{ij}^{-1}}{\sqrt{H_{ii}^{-1} H_{jj}^{-1}}}. \quad (5.40)$$

where H_{ii} denotes the i th diagonal element of the matrix (not the trace). Like the Hessian, this matrix is symmetric; the off-diagonal terms are the correlation coefficients between parameters. Parameters with correlation coefficients above 0.9 are considered to be significantly correlated (Figure 5-6, right panel).

The relative magnitudes of the eigenvalues of the Hessian give information about how sensitive the fit is to the various parameter combinations described by the eigen-

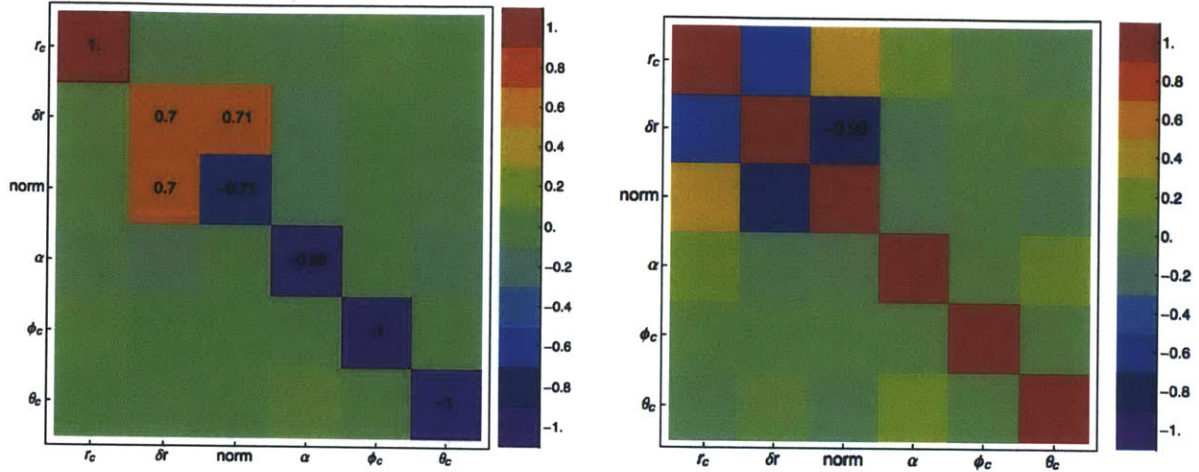


Figure 5-6: Two different tests for correlations between parameters, the matrix of eigenvectors of the Hessian (left) and the correlation matrix R (right), are defined and discussed in the text.

vectors. A large eigenvalue indicates that the corresponding parameter or combination is very well constrained since a small departure in that direction will produce a large change in χ^2 . A small eigenvalue indicates that the fit is not very sensitive to that combination of parameters. For multiparametric nonlinear fits, the range of eigenvalues can be very large: the matrix can be ill-conditioned if this range spans more orders of magnitude than the precision of the numerical calculation. In the fit to the sample data the eigenvalues range over four orders of magnitude (Figure 5-7), easily within the capabilities of double-precision representation. As expected, the two correlated parameters are most difficult to determine: the lowest eigenvalue corresponds to the linear combination p_- ; the next-lowest eigenvalue is for p_+ . Comparing Figures 5-7 and 5-3 shows that the size of the eigenvalues tracks very closely with the order of convergence.

The standard error on the parameter estimates, as given in Table 5.2, is also determined from the Hessian:

$$\sigma_i = \sqrt{2H_{ii}^{-1}}. \quad (5.41)$$

The standard error is equivalent to the 68% confidence interval for the independent parameters in the fit (for a few hundred degrees of freedom, the normal distribution

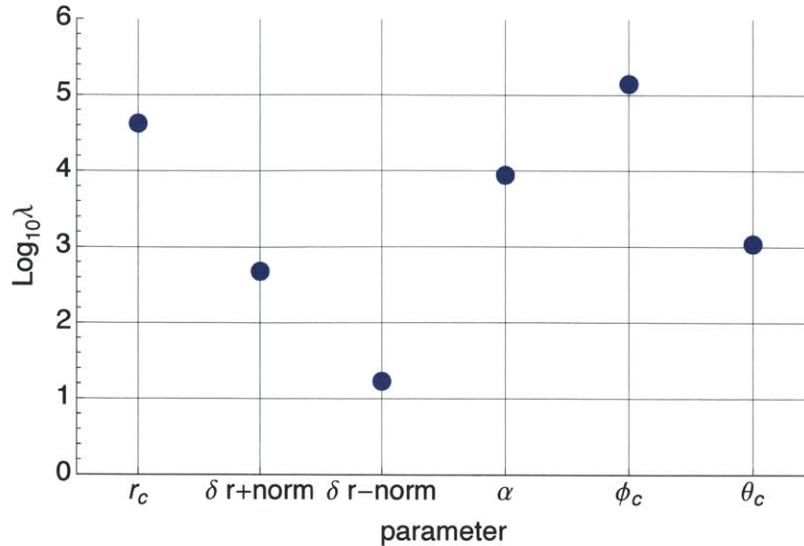


Figure 5-7: Eigenvalues of the Hessian for the sample data.

is a suitable approximation to the chi-squared distribution). In the context of the properties of the Fisher matrix, this is a lower limit on the error (the Crámer-Rao bound). For the correlated parameters, the error contours are shown in Figure 5-4.

Restricted fits

The fit can be made to converge faster if the dimensionality of the parameter space can be reduced. In these tests it was noticed that for θ_c close enough to $\pi/2$ that the structure looks like a shell in projection, some of the parameters can be fairly accurately estimated from the projected image. In particular, guesses for α and ϕ_c from the image are usually correct within about 10 percent. Furthermore, r_c can be estimated fairly accurately as well by comparing the projected surface brightness profile with the shape of the deprojected density profile. Holding these three parameters constant during the fit help it converge faster and more reliably, especially r_c since it is slightly correlated with the two most crucial parameters to fit correctly, δr and θ_c . The peak surface brightness varies roughly as the square root of the caustic width, so the normalization cannot be held constant. However, it helps if a normalized version of the image data is fitted rather than using absolute values, since the fractional variation in the norm can be much larger than that of δr and θ_c , causing it to dominate

the behavior of the χ^2 .

5.3 Fitting an N-body model of two tidal shells

We used the restricted fitting method discussed in the previous section to fit the projected mass density (for a constant mass-to-light ratio, this is proportional to the surface brightness) of the two tidal shells around M31 using the existing N-body model. This test is meant to be an intermediate step between testing the ability of the model to recover input parameters, as demonstrated in the previous section, and fitting real surface brightness measurements. In this demonstration, the shells are produced by gravitational interaction, not sampled directly from the model, so we do not expect the model to be a perfect fit. This test is intended to determine how well it can still recover the right parameters. In this section we present the results of the fit and the analysis of the Hessian eigensystem for each shell. We also compare the results to the fits from Section 2.3.3, which used the full three-dimensional structure of the shells. As in Section 2.3.3, we refer to the older shell as Caustic 1 and the younger one as Caustic 2.

5.3.1 Preparation of sample data

The orientation of the M31 shells with respect to our line of sight is such that only one of them looks like a shell in projection. However, looking down on the system from “above” (in our coordinate system, the line of sight along the y axis) produces a projection with two clear shells. We use this line of sight for the fit in order to have two shells to fit, allowing us to test whether fitting becomes easier or harder as the system evolves. The data is obtained by integrating the estimated density of the N-body representation along this line of sight in the same way as for the sample data in the previous section. The pixel size is chosen so that Poisson errors are about 20-30 percent or less in the region around the edge of each shell.

Once the surface brightness of each shell is obtained, a small amount of additional work is necessary before fitting. We saw in Chapter 2.3.3 that only the region right

around each shell is expected to fit the universal profile, as illustrated in Figure 2-15. So a cut in projected-radius is made to the surface brightness image for each shell to select only the region around the edge. The selection radii are determined by comparing the various surface brightness profiles (versus projected radius) to the universal profile. Finally, we removed points with Poisson errors larger than 100 percent and normalized the remaining data to its maximum value (a step which helps the performance of the fitting program). The data sets submitted to the fitting program are shown in Figure 5-8.

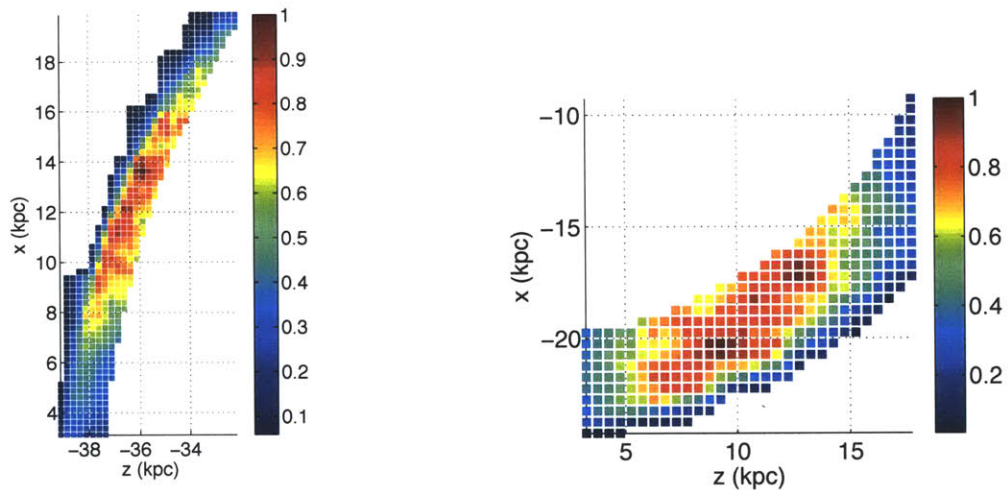


Figure 5-8: Data sets for Caustics 1 (left) and 2 (right) submitted to the fitting program, with the cuts made as described in the text. The color bar now indicates the normalized surface brightness.

5.3.2 Fitting the reoriented shells

We first tried, unsuccessfully, to use the full six-parameter fit to the shells. The fit was not very successful for either shell using this algorithm: the slight systematic unidentifiability of some of the parameters demonstrated in the previous section is compounded by practical unidentifiabilities resulting from imperfect measurements and the assumption that the debris is cone-shaped and spherical. This second assumption was demonstrated to be false in Chapter 2.3.3, where a linear or quadratic

fit in θ and ϕ was used to correct for the asphericity of both shells. As a result, the values of δr and θ_c will be more uncertain than for the sample data drawn from the model; in particular we expect δr to be overestimated. The fitter was not always able to find a local minimum and could not distinguish between the true minimum (the set of parameters measured in Chapter 2.3.3) and other minima; in fact, the true parameters had a larger value of χ^2 than some other combinations, indicating that the range of sensitivities in χ^2 to the various parameters was too large. This was borne out by the range of eigenvalues of the Hessian matrix at the various minima, which often included several that were smaller than the condition number.

The restricted fit using constant estimates of r_c , α , and ϕ_c was more successful. The values held constant were estimated using the surface-brightness data only, not the three-dimensional data. Based on these results we attempted an iterative fit to all six parameters as follows. Once the fit had converged for δr , the normalization, and θ_c , those parameters were then held constant at the fitted values and the other three were fitted. As a third step, we then allowed all six parameters to vary using the restricted fit results as starting points. The progress of the fits is shown in Figure 5-9, and the final results of each fit after all three steps are given in Tables 5.3 and 5-13.

As expected the eigensystem of the Hessian matrix and R for each shell both indicate that there is a significant correlation between δr and the normalization (Figures 5-11 and 5-12, respectively). This correlation is worse than in the fit to sample data generated from the model, as might be expected since a larger value of the pixel size Δ relative to δr was used. The matrices of eigenvectors show that an additional correlation between the angles α and θ_c has started to develop, though its correlation coefficient is not high enough to be of concern (Figure 5-12).

5.3.3 Comparison with full three-dimensional fit

Examining Tables 5.3 and 5.4, we see that the performance of the fit is variable. The width δr is systematically overestimated and the caustic radius r_c is systematically underestimated. These systematic errors are a failure of the model to fully represent

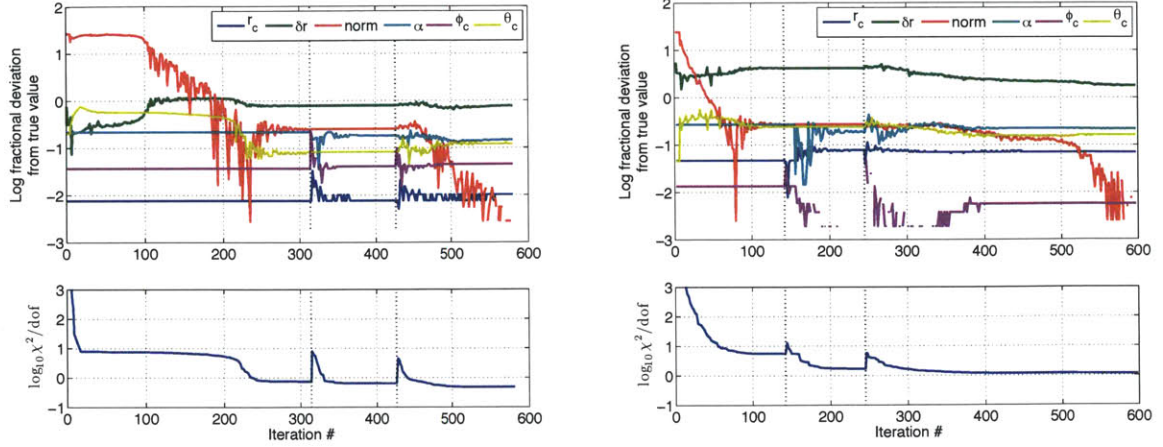


Figure 5-9: Progress of the iterative fit method for Caustics 1 (left) and 2 (right). Successive restarts are denoted with vertical dashed lines. In the first two steps, half the parameters are held constant and the other half allowed to vary; in the final step, all the parameters are allowed to vary. All parameters are compared to the values given in the last column of Tables 5.3 (for Caustic 1) and 5.4 (for Caustic 2) except for the normalization, which is compared to its final value.

Parameter	Starting Guess	Starting Search Length	Final Value	Std. Error	True Value
δ_r	0.3	0.1	0.4	0.2	0.2
$f_0/\sqrt{2\pi\kappa}$	1.0	0.3	0.036	0.015	n/a
θ_c	1.6	0.5	1.49	0.06	1.31
r_c	39.0	1.0	38.9	0.2	39.3
α	0.22	0.05	0.239	0.009	0.280
ϕ_c	2.8	0.5	2.82	0.02	2.73

Table 5.3: Fit information for Caustic 1. The uncertainties for the fitted parameters are calculated from the Hessian as described in the text. Contours for the correlated parameters are shown in the left panel of Figure 5-10. True values of the parameters were calculated using the full N-body representation in three dimensions.

Parameter	Starting Guess	Starting Search Length	Final Value	Std. Error	True Value
δ_r	1.0	0.5	0.65	0.38	0.23
$f_0/\sqrt{2\pi\kappa}$	1.0	0.5	0.040	0.022	n/a
θ_c	1.6	0.5	1.78	0.09	1.53
r_c	24.5	1.0	23.9	0.3	25.7
α	0.3	0.1	0.50	0.03	0.41
ϕ_c	5.2	0.5	5.24	0.04	5.27

Table 5.4: Fit information for Caustic 2. The uncertainties for the fitted parameters are calculated from the Hessian as described in the text. Contours for the correlated parameters are shown in the right panel of Figure 5-10. True values of the parameters were calculated using the full N-body representation in three dimensions.

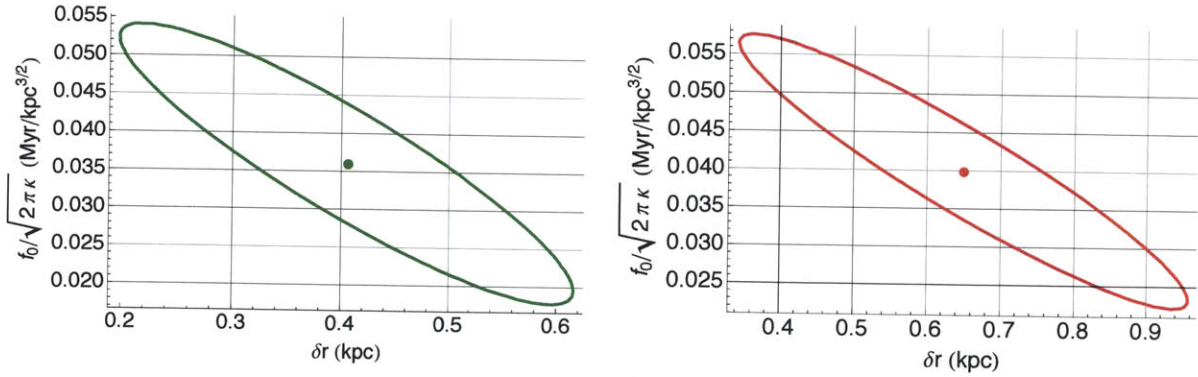


Figure 5-10: 68% confidence ellipses for the correlated parameters δr and $f_0/\sqrt{2\pi\kappa}$ in Caustic 1 (left) and Caustic 2 (right).

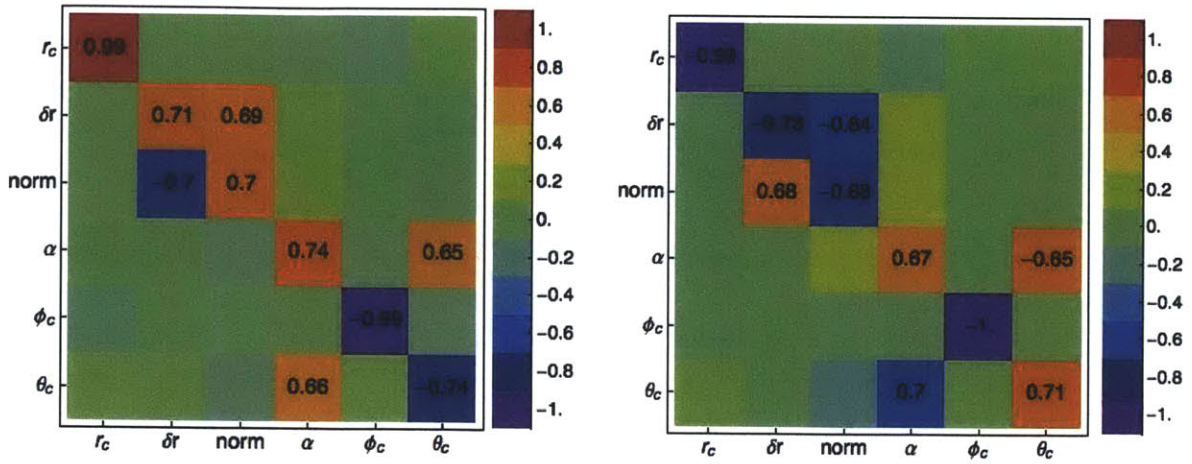


Figure 5-11: Matrices of eigenvectors of the Hessian at the best-fit point for Caustics 1 (left) and 2 (right).

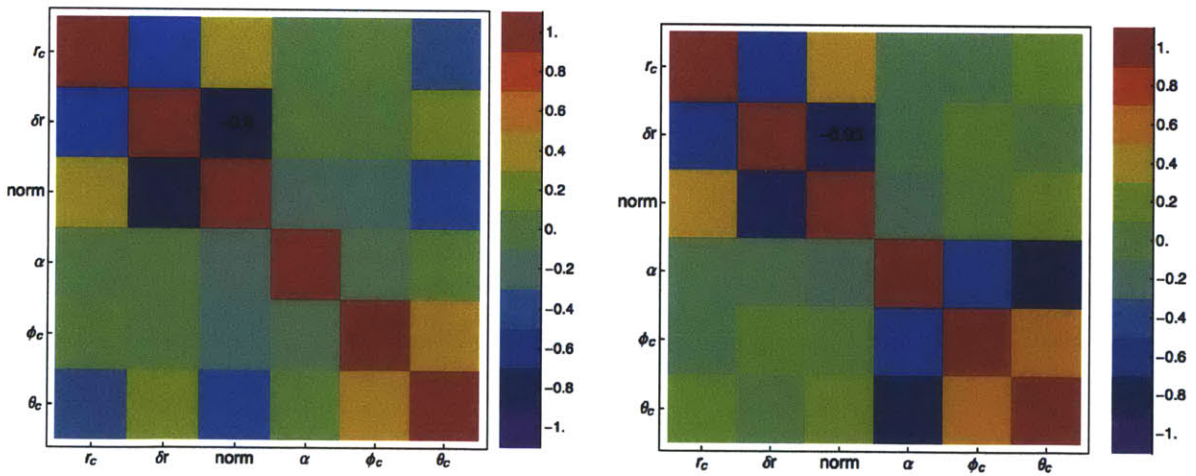


Figure 5-12: Correlation matrices R for Caustics 1 (left) and 2 (right).

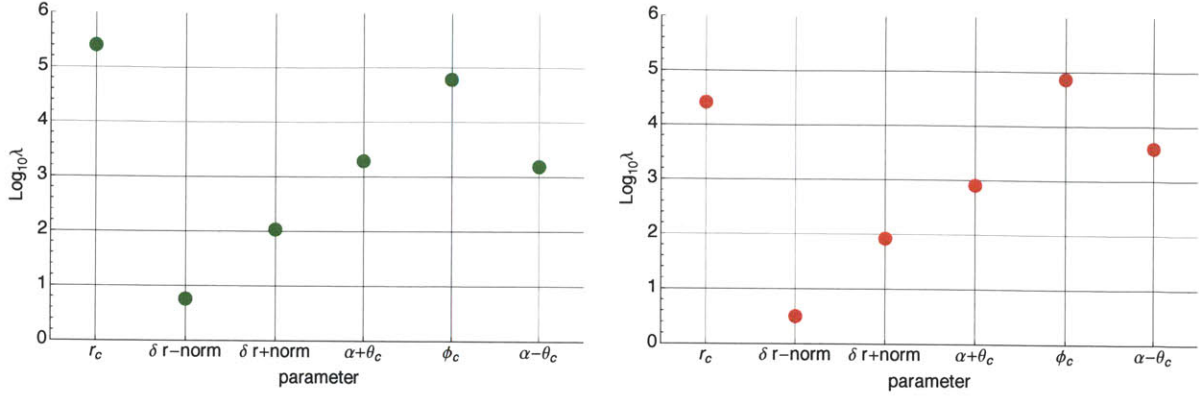


Figure 5-13: Eigenvalues of the Hessian at the best-fit point for Caustics 1 (left) and 2 (right).

the structure of the shells: this problem and its solution are discussed in the next section. The angles θ_c , ϕ_c , and α are not always successfully recovered, which may be partially due to the irregular distribution of the material along the shell (Figure 5-14). Recovery of these parameters is better when the material is more evenly distributed: in caustic 2, ϕ_c and α are recovered accurately but θ_c is not since it is forced to pick one of the two density peaks. The algorithm also has better success when the ranges of θ and ϕ spanned by the caustic are comparable, so that they can both be represented with the single parameter α . Caustic 1 spans nearly double the range in θ as it does in ϕ , leading to difficulty in recovering both parameters properly given only one angle α . The same is true for Caustic 2, but since the fit has selected one of the two density peaks (each of which have roughly half the width of the whole distribution), it can more easily recover both θ (for that peak) and ϕ since these two widths are comparable. This discrepancy in the angular span of the caustic is another symptom of the model's failure to fully describe the shape of the shell.

5.3.4 Systematic error: the tilt

In the N-body model, neither caustic is perfectly spherical: r_c varies linearly or even quadratically with angle, as illustrated in Figure 2-14. In fitting the caustics using three-dimensional data we corrected for the asphericity by fitting a model to r_c

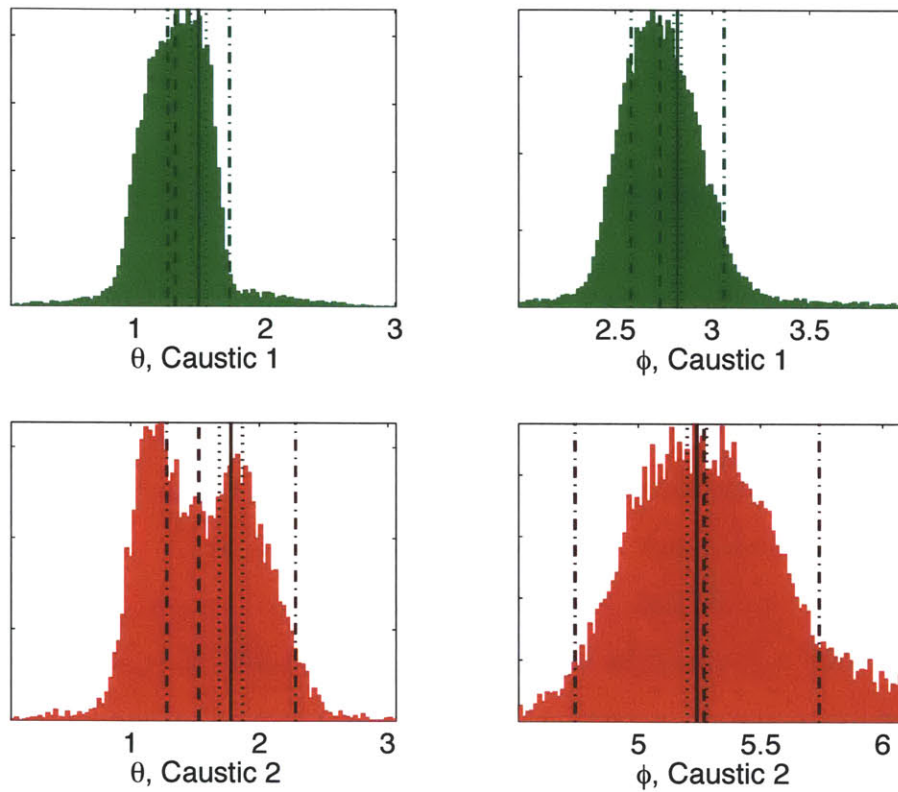


Figure 5-14: Recovery of the orientation angles of the caustics is sensitive to irregularities in the distribution of material, which is also seen in Figure 5-8. The recovered values (solid vertical lines, with uncertainties denoted by dotted lines) are often offset from the midpoint (dashed line) thanks to the irregular distribution of material. The angular extent α (delimited by dot-dashed lines) is assumed in the model to be the same in θ and ϕ , but in reality the two ranges differ. The recovered value must be intermediate between them and generally leans low.

and calculating distances from this model, but in the two-dimensional case we have assumed that r_c is a constant. This leads to a systematic error that overestimates δr . This systematic error can account for the entire discrepancy between the fitted values of δr in Tables 5.3 and 5.4. If the difference in r_c over the span of each caustic is calculated using the fitted equations from Table 2.5, and added to the true width of each caustic from the three-dimensional fit, we see that the combined width matches or exceeds the fitted δr in the two-dimensional case (Table 5.5). In the case of Caustic 2 we saw that the fit picked out one of two density peaks rather than fitting the entire envelope, so the result is closer to half the total width.

To truly estimate the caustic width from a two-dimensional image, therefore, the model must be amended to allow for a tilted caustic surface. The physical reason for this tilt suggests that this could be done by including one additional parameter. The deviations from a spherical caustic result from the variation in energy over the angular span of the shell, as shown in Figure 5-15. Thanks to this correlation, a single variable related to the energy spread of the material should be able to account for the tilt. This variation in energy ΔE corresponds to a spread in radius Δr of approximately

$$\Delta r_c \approx \frac{\Delta E}{g(r_c)}, \quad (5.42)$$

where $g(r_c)$ is the local gravity at the caustic radius, that can account for the entire span of r_c (Table 5.5).

Parameter	Caustic 1	Caustic 2
δr , fitted (from Tables 5.3 and 5.4)	0.4 kpc	0.65 kpc
Δr_c , calculated from Table 2.5 line 2	0.43 kpc	1.1 kpc
ΔE (see Figure 5-15)	$1.0 \times 10^{-3} \text{ (kpc/Myr)}^2$	$2.5 \times 10^{-3} \text{ (kpc/Myr)}^2$
Δr_c , estimated from Equation 5.42	0.8 kpc	1.1 kpc

Table 5.5: Energy variation and corresponding radial variation in caustics 1 and 2, compared with fit results.

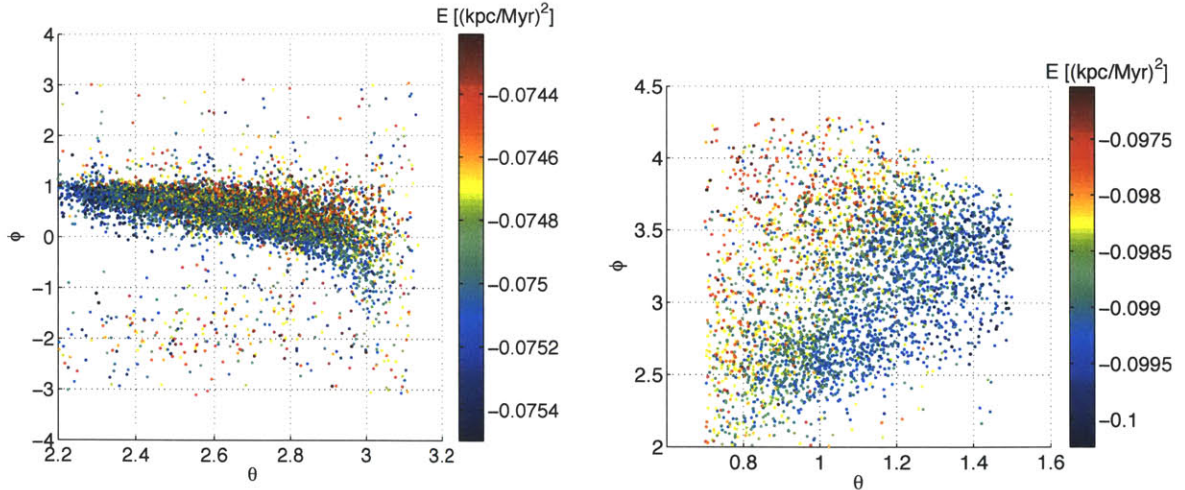


Figure 5-15: Energy (in $(\text{km/s})^2$) varies in Caustics 1 (left) and 2 (right) as some function of the angles θ and ϕ , producing a variation in r_c .

5.3.5 Fit to Caustic 2 in its original orientation

We also attempted to fit Caustic 2 in the orientation corresponding to our actual view of the structure. This fit was more problematic than those attempted in the rotated orientation. The surface brightness of the caustic in the N-body model in this orientation exhibits a bifurcation that is the main cause of the trouble (Figure 5-16, left panel); *this bifurcation is an artifact of the N-body model that is not seen in the star-count map of the shells* [32]. This caustic also displayed the larger asphericity of the two (Figure 2-14). In the end the fit failed to recover the parameters of the caustic, but we include it anyway for two reasons: first, we had hoped to use it as a proxy for real data; second, the fit fails in a characteristic way that demonstrates the model's limits and how the fitting procedure attempts to cope with them.

We prepared the data for fitting in the manner described in Section 5.3.1; additionally, we subdivided the data in the x - y plane with a simple cut at $y = 0$ so that we could also consider each surface-brightness maximum separately. This was motivated by examining the surface-brightness profile, which appears to be a superposition of two peaks (Figure 5-16, right panel).

We attempted a fit of the entire structure first, using the method described in

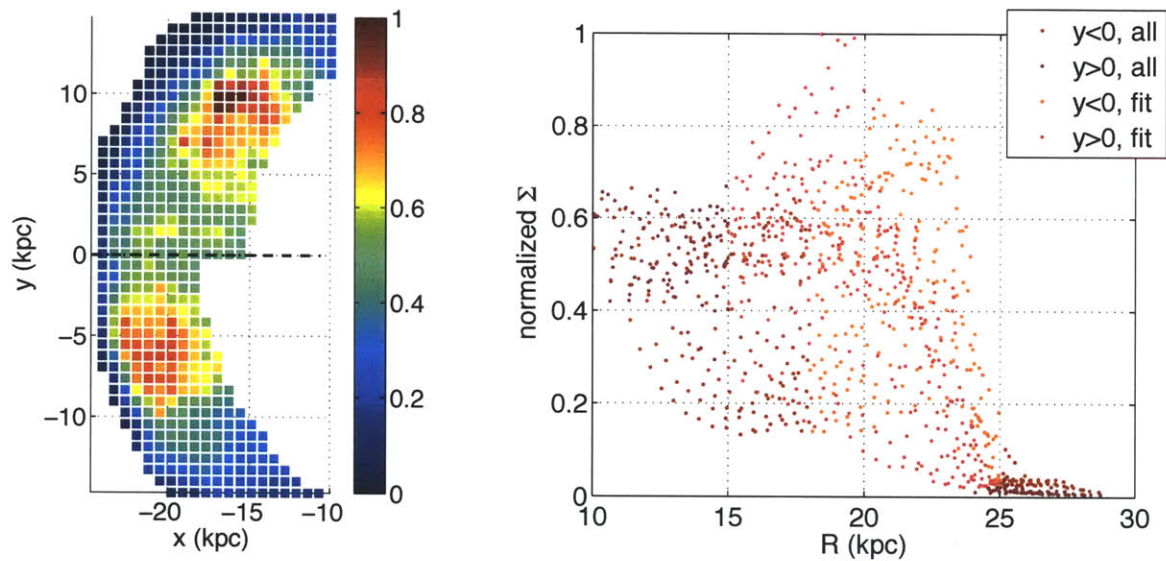


Figure 5-16: The surface brightness map of the N-body model of Caustic 2, viewed in its proper orientation with respect to the observer, has a bifurcation in the shell edge that is not observed in the data used to create the model (left panel). In addition to attempting a fit of the entire shell, we also tried to fit the upper and lower pieces (delineated by the dashed line in the left panel) separately. This was motivated by the double-peaked structure in the surface brightness profile (right panel).

Section 5.3.2. The partially restricted, iterative method converged to a saddle point rather than a minimum when given reasonable guesses for r_c , α , and ϕ_c . We also attempted the fit allowing all the parameters to vary. In this case the fit identified a solution that is more or less uniform across the entire data set.

We also tried fitting each subset of data separately even though this would obviously give conflicting results. This was slightly more successful but the fit was not able to reproduce the shape of the bifurcation in either case, settling instead on a solution that reproduced the shape of the background (Figure 5-17). As expected, most of the parameters are not reliably recovered by the fit. In both cases the eigenvalue for the normalization is much larger than any of the others, causing the Hessian matrix to be so ill-conditioned that none of the other eigenvalues are larger than the condition number. It seems that, unable to suitably reproduce the exact shape of either half, the fit relies on adjusting the normalization to reduce the χ^2 value. This is characteristic of the model's failure to account for the asphericity or the bifurcation in this example.

The bifurcation that is primarily responsible for the failure to recover the parameters in this example would probably also prevent good recovery even if the asphericity was accounted for. Good candidate images for this type of analysis should therefore display a minimum of additional structure at the shell surface.

5.4 Future work

In order to use the method demonstrated in this chapter to constrain real mergers, the model should be changed to include a parameter that accounts for shell asphericity. Then the fitting procedure should be re-tested to make sure that no new structural degeneracies have arisen in the model. We expect that adding an asphericity parameter will actually improve the ability of the model to recover the various parameters from real caustics, since the asphericity exacerbates the correlations between parameters. After this step the model will be ready to use to fit real surface-brightness data.

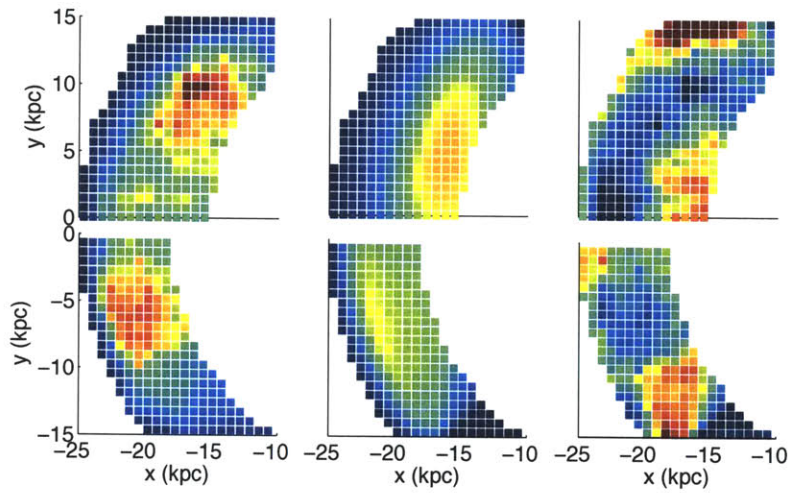


Figure 5-17: The fit to each half of the caustic attempts to fit the background density rather than account for the structure in the shell. The top row shows the upper half of the caustic; the bottom row shows the lower half (as in Figure 5-16). The leftmost column shows the surface brightness data being fitted on a normalized scale from 0 (blue) to 1 (red). The center column shows the “best-fit” model on the same scale. The rightmost column shows $(\Sigma(\vec{x}_i) - \hat{\Sigma}_i) / \sigma_i$ at each point, on a scale from -3 (blue) to 3 (red), with zero colored green.

Once the model is ready there is plenty of real data to be fit. Semi-analytic modeling suggests that observations must reach a depth of at least 26 magnitudes per arcsecond to see the brightest tidal streams [31]. Space-based observations can easily reach this depth and have been used to study systems of tidal caustics around several elliptical galaxies [e.g. 42]. The catalog of [41] includes many bright examples of elliptical galaxies that show these features, some of which have been re-observed from space or the ground. More recently, ground-based observations have reached this brightness threshold and observed systems with tidal caustics, including a large number of elliptical galaxies in the co-added Stripe 82 in the Sloan Digital Sky Survey [40] and 7 spiral galaxies observed by [29] with a consortium of wide-field amateur telescopes. Any of these data sets would be good places to look for candidates; in particular, the spiral “Umbrella Galaxy” NGC 4651 and the elliptical galaxy NGC 3923 have already been extensively observed. The approach described in this work can be applied to galaxies with a wide range of morphologies as long as the surface brightness profiles of the shells can be obtained. In future work we will choose and analyze some systems of shells observed around galaxies as a pilot study for the method.

5.5 Conclusion

In this chapter we have discussed how the deprojection of the shells can be modeled and implemented, and demonstrated a few examples of the quality of the fits achieved by a simple implementation. The most difficult parameters to fit are the width δr and the inclination angle θ_c , which are somewhat correlated systematically. For real shells the correlation is exacerbated by the slight asphericity of the shells, leading to difficulty in determining the ill-defined caustic radius and an over-estimation of δr . This shortcoming of the model should be corrected in order to recover the correct δr ; fortunately the asphericity can probably be accounted for by adding only one new parameter to the model since it arises from smooth variations in the energy and angular momentum of material along the caustic.

Chapter 6

Conclusions

In this thesis we have developed and demonstrated the use of a new formalism for describing tidal structures created by merging galaxies with low orbital angular momentum. Because of their special properties, we chose to call these structures “tidal caustics.” Here we summarize our results.

In Chapter 2, we described how the high densities achieved in tidal caustics motivated us to test whether they were a likely source of gamma-ray radiation from self-interacting dark matter particles. We tested estimators for the integrated density-squared in regions where the density contrast is high by comparing analytic calculations to numerical estimates of this quantity for a one-dimensional caustic. We found that unless all the features in the density distribution are known to be fully resolved, the best estimator uses a simple nearest-neighbors method to find the squared density and a constant Riemann volume to perform the integral. We used this method to estimate the signal from the tidal caustics, finding that the total gamma-ray flux from the shells is three orders of magnitude lower than emission from the dark halo, and too low to be detected by Fermi.

In Chapter 3 we presented the full derivation of the density profile for a one-dimensional caustic in a system with an initially Gaussian velocity dispersion and force-free motion starting from correlated initial positions and velocities. We found that the form of the density profile is insensitive to the details of the motion and is an equally good description of caustics created by the radial motion of material in a

gravitational potential. The reason for this remarkable universality is the underlying similarity in phase space: the profile is valid as long as the initial distribution is confined to a tiny fraction of the available phase space and the subsequent evolution conserves phase space volume. For the case of the tidal disruption of a satellite galaxy on a radial orbit, we related the shape parameters of the density profile to the initial energy spread of the satellite galaxy and the local gravity and gravity gradient of the host galaxy. We also derived an expression for the integrated squared density of the caustic.

In Chapter 4, motivated by the distinctive shapes of tidal caustics in images, we determined how to constrain parameters of the mergers that create them based only on the information available from the surface brightness profile of one caustic. We found that two joint constraints on the parameters identified in Chapter 3 could be obtained if the surface brightness profile could be deprojected. If multiple shells or line-of-sight velocity information is available, some information about the orbital angular momentum of the merger and the flattening of the host potential can be obtained, though this is more model-dependent than the density profile information.

In Chapter 5 we discussed how to deproject the surface brightness profile using a simple cone-shaped model of the caustic. We found that the width δr and the inclination angle θ_c are systematically correlated in this model; the slight asphericity of the caustics exacerbates the problem, leading to difficulty in determining the ill-defined caustic radius and an over-estimation of δr . However, despite these limitations we found that the deprojection method could still recover parameters from surface brightness profiles of caustics from a realistic N-body model. Furthermore, we demonstrated that the asphericity of the caustics can be accounted for by introducing a single extra parameter to the model; this is outlined for future work.

Radial mergers, from their discovery as shell galaxies nearly 30 years ago, have promised to better our understanding of galaxy formation and its connection to cosmology. The work in this thesis has demonstrated the advantages inherent in exploiting the symmetry of these mergers to make the most of limited information. The results open up 30 years of spectacular images to dynamical analysis, connecting shell

galaxies to the great cosmic web and gravity's slow shaping of all the structures in it.

Appendix A

Rate Estimators

Here we describe the five rate estimators tested in this work. They are based on three types of density estimators (spherical and Cartesian nearest-neighbor algorithms and a spherical smoothed kernel method) and two methods for determining Riemann volumes (constant size and adaptive tessellation).

A.1 Nearest Neighbor

An N-body representation of a continuous number density distribution $n(\vec{x})$ is a Poisson point process with a spatially varying mean. As such, all estimators (e.g., \hat{n}) of the density and its higher moments (n^2 , n^3 , etc.) obey the statistics of point processes. In the case of a uniform distribution, these are simply the well known Poisson statistics, and the simplest estimator calculates the density in terms of the distance to the N_s^{th} particle, called the nearest neighbor distance. For a given value of the smoothing number N_s there is a particular nearest neighbor distance r_{N_s} for each particle, and the density near the particle is estimated using

$$\hat{n}_b = \frac{3N_s}{4\pi r_{N_s}^3}, \quad (\text{A.1})$$

However, if we compute $E(\hat{n}_b)$ by integrating over the probability distribution describing the distances between particles, we find that

$$E(\hat{n}_b) = \frac{N_s}{N_s - 1}n, \quad (\text{A.2})$$

indicating that the estimator is biased since $E(\hat{n}_b) \neq n$. This is a result of the random fluctuations in the distance r_{N_s} from particle to particle, which obey Poisson statistics. Even in cases where the density is not uniform a similar effect is present.

The Poisson bias of the estimator (A.1) can be easily eliminated in the case of the uniform distribution by noting that $E(\hat{n}_b)$ differs from n by a constant factor only. Dividing by this factor produces the estimator

$$\hat{n}_u = \frac{3(N_s - 1)}{4\pi r_{N_s}^3}, \quad (\text{A.3})$$

which has $E(\hat{n}) = n$.

We wish to construct a minimally biased estimator for the rate, $\Gamma = \int n^2 dV$. It is well known that in a Poisson distribution $E(\hat{n}_u^2)$, the expectation value of the square of the unbiased density estimator in equation (A.3), is not equal to n^2 ; still, an unbiased estimator for n^2 in the case of the uniform distribution does exist:

$$\hat{n}_{u,i}^2 = (N_s - 1)(N_s - 2) \left(\frac{3}{4\pi r_{N_s}^3} \right)^2 \quad (\text{A.4})$$

with N_s and r_{N_s} defined as before. Then $E(\hat{n}_{u,i}^2) = n^2$.

Using $\hat{n}_{u,i}^2$, we can construct an unbiased estimator for Γ by using a Riemann sum over N_V identical volumes dV to approximate the volume integral, so that

$$\hat{\Gamma}_u = dV \sum_{i=1}^{N_V} \hat{n}_{u,i}^2 \quad (\text{A.5})$$

where $\hat{n}_{u,i}^2$ is given by evaluating Equation (A.4) at the center of subvolume i , and the total volume $V = N_V dV$. Using Equation (A.5), $E(\hat{\Gamma}_u) = \Gamma_{\text{true}}$. This estimator provides a useful check that the code is functioning properly.

We also tested the same density estimation method with an adaptive Riemann volume, in which each particle occupies its own box ($N_V = N_p$). Each particle's Riemann volume contains all the space closer to that particle than any other. The size of such a Riemann volume is also affected by Poisson statistics, so this rate estimator will not be unbiased even if Equation (A.4) is used to calculate the density-squared. For simplicity, we represent the N_s dependence of the additional bias from the adaptive box size as a prefactor, to be determined numerically, and use Equation (A.4) to estimate the density-squared:

$$\hat{\Gamma}_n = \frac{1}{1 + b_n(N_s)} \sum_{i=1}^{N_p} \hat{n}_{u,i}^2 dV_i \quad (\text{A.6})$$

A.2 FiEstAS

A variation on the nearest-neighbor estimator is implemented by Ascasibar and Binney (2005) in their algorithm FiEstAS. We refer to this estimator as f . It too uses the N_s th nearest neighbor, but instead of a spherical volume considers the volume of the Cartesian box enclosing N_s particles when calculating the density, so that for a particle i :

$$\hat{n}_{f,i} = \frac{N_s}{V_{N_s,i}} \quad (\text{A.7})$$

Conveniently, the construction of the tree used in calculating dV_i also chooses the Riemann volume adaptively in the same manner as for the estimator n . Now there are two contributions to the bias: the Poisson bias from using $(\hat{n}_{f,i})^2$ to estimate the density-squared and the Poisson bias from determining the adaptive Riemann volumes. For simplicity, we represent the N_s -dependence of both contributions with a single prefactor, so the rate estimator is

$$\hat{\Gamma}_f \equiv \frac{1}{1 + b_f(N_s)} \sum_{i=1}^{N_p} (\hat{n}_{f,i})^2 dV_i \quad (\text{A.8})$$

A.3 Kernel-based

We also tested two kernel-based rate estimators. Kernel-based density estimators use a weighted sum to smooth over the N_s nearest particles, so that the estimated density at location \vec{x} is

$$\hat{n}_s(\vec{x}) = \sum_{j=1}^{N_s} W(\vec{x}_j - \vec{x}, \vec{h}) \quad (\text{A.9})$$

The smoothing vector \vec{h} is a generalized nearest-neighbor distance. The vector has length $|\vec{x}_{N_s} - \vec{x}|$. For a one-dimensional spherical kernel $\vec{h} \equiv r_N \hat{r}$ and the kernel function W is nonzero when $|\vec{x}_j - \vec{x}| < r_N$. For a three-dimensional kernel, $\vec{h} \equiv \vec{x}_{N_s} - \vec{x}$ and W is nonzero when $|\vec{x}_j - \vec{x}| < |\vec{h}|$.

The type of kernel chosen can have a significant effect on the bias and variance. [108] have tested the bias and variance of density estimators with a variety of kernels on uniform density distributions, and we use their notation here. The one-dimensional kernel can be written in the form

$$W(\vec{r}, \vec{h}) = \frac{fW(u)}{V_h} \quad (\text{A.10})$$

where \vec{r} is the distance from the target location, u is the scaled distance

$$u = r/h, \quad (\text{A.11})$$

f is the kernel normalization

$$\frac{1}{f} = \int_0^1 W(u) 4\pi u^2 du, \quad (\text{A.12})$$

and V_h is the volume enclosed by the smoothing length \vec{h} .

[108] found that the Epanechnikov kernel

$$W(u) = \begin{cases} 1 - u^2 & 0 \leq u \leq 1 \\ 0 & \text{otherwise} \end{cases} \quad (\text{A.13})$$

has the smallest bias and variance in estimating the density in the case of a uniform distribution. We use this density estimator and an adaptive Riemann volume for the rate estimator s , and again collect the N_s -dependence of the bias in a prefactor:

$$\hat{\Gamma}_s \equiv \frac{1}{1 + b_s(N_s)} \sum_{i=1}^{N_p} [\hat{n}_s(r_i)]^2 dV_i \quad (\text{A.14})$$

Again, $1/(1 + b_s)$ is the bias when using a constant Riemann volume.

[58] used an adaptation of the kernel-based method to estimate the rate from simulations of the Milky Way's dark halo and halo substructure. Starting with Equation (A.14), they make the substitution $\hat{n}_s(r_i)dV_i = 1$ (there is one particle per Riemann volume). We examine this variation of the kernel-based method, referred to as estimator d . We include a bias-correcting prefactor that is equal to 1 in [58]:

$$\hat{\Gamma}_d \equiv \frac{1}{1 + b_d(N_s)} \sum_{i=1}^{N_p} \hat{n}_s(r_i) \quad (\text{A.15})$$

From a Poisson-statistics standpoint, the substitution implicitly assumes that $E(\hat{n})^2 = E(\hat{n}^2)$, yet the estimator itself is linear instead of quadratic in the density. For this reason it is expected to behave differently than the rate estimator (A.14).

Appendix B

Calculation of the Minimum Nearest-neighbor Distance

In this appendix we derive an expression for the expectation value of the minimum nearest-neighbor distance, $E(\hat{r}_{N_s, \min})$, in the case of a uniform density distribution of particles. The scaling of this value with the smoothing number N_s and the number N_p of particles in the simulation subsequently determines the maximum density that can be both represented in an N-body realization with N_p particles and calculated using the nearest-neighbor estimator with N_s nearest neighbors. $E(\hat{r}_{N_s, \min})$ is the first order statistic of the estimator \hat{r}_{N_s} , which is related to the nearest-neighbors density estimator \hat{n} by

$$\hat{r}_{N_s} = \left(\frac{4\pi\hat{n}}{3N_s} \right)^{-1/3}. \quad (\text{B.1})$$

We start with the PDF of the nearest-neighbor distance,

$$p_{\hat{r}_{N_s}}(\rho)d\rho = \frac{\exp(-4\pi n\rho^3/3)}{(N_s - 1)!} \left(\frac{4\pi n\rho^3}{3} \right)^{N_s-1} d\left(\frac{4\pi n\rho^3}{3} \right), \quad (\text{B.2})$$

which can be derived from directly integrating over the joint PDF for the nearest N_s particles. The PDF for each particle is Poisson. To calculate the first order statistic

we also need the CDF of the nearest neighbor distance,

$$P_{\hat{r}_{N_s}}(\mu) = \frac{1}{(N_s - 1)!} [\Gamma(N_s) - \Gamma(N_s, 4\pi n\mu^3/3)] = 1 - \frac{\Gamma(N_s, 4\pi n\mu^3/3)}{\Gamma(N_s)}, \quad (\text{B.3})$$

where $\Gamma(N)$ and $\Gamma(N, x)$ are the complete and incomplete gamma functions, respectively.

The PDF of $\hat{r}_{N_s, \min}$ is that of the first order statistic of the PDF of the nearest-neighbor distance:

$$p_{\hat{r}_{N_s, \min}}(\nu) d\nu = N_p \{1 - P_{\hat{r}_{N_s}}(\nu)\}^{N_p - 1} p_{\hat{r}_{N_s}}(\nu) d\nu. \quad (\text{B.4})$$

Substituting the expressions for the PDF and CDF of the nearest-neighbor distance,

$$p_{\hat{r}_{N_s, \min}}(\mu) d\mu = \frac{N_p}{(N_s - 1)!} \exp(-4\pi n\mu^3/3) \left(\frac{4\pi n\mu^3}{3}\right)^{N_s - 1} \left(\frac{\Gamma(N_s, 4\pi n\mu^3/3)}{\Gamma(N_s)}\right)^{N_p - 1} d\left(\frac{4\pi n\mu^3}{3}\right). \quad (\text{B.5})$$

Changing variables to $y = 4\pi n\mu^3/3$ gives us a simpler expression:

$$p_{\hat{r}_{N_s, \min}}(y) dy = \frac{N_p}{(N_s - 1)!} e^{-y} y^{N_s - 1} \left[\frac{\Gamma(N_s, y)}{\Gamma(N_s)}\right]^{N_p - 1} dy. \quad (\text{B.6})$$

y represents the average number of particles in a sphere of radius μ .

The expectation value of $\hat{r}_{N_s, \min}$ is then determined by weighted integration over its PDF:

$$E(\hat{r}_{N_s, \min}) = \left(\frac{3}{4\pi n}\right)^{1/3} \frac{N_p}{[(N_s - 1)!]^{N_p}} \int_0^\infty e^{-y} y^{N_s - 2/3} [\Gamma(N_s, y)]^{N_p - 1} dy, \quad (\text{B.7})$$

using the definition of μ in terms of y and the fact that $p_{\hat{r}_{N_s, \min}}(y) dy = p_{\hat{r}_{N_s, \min}}(\mu) d\mu$.

We now use two approximations for the gamma function. The first is the asymptotic expansion for the incomplete gamma function at large y :

$$\Gamma(N_s, y) \approx e^{-y} [y^{N_s - 1} + \mathcal{O}(y^{N_s - 2})]; \quad (\text{B.8})$$

the second is Stirling's approximation for the complete gamma function at large

argument N :

$$\Gamma(N + 1) = N! \approx N^N e^{-N}. \quad (\text{B.9})$$

Using the asymptotic expansion, we may rewrite the integral:

$$E(\hat{r}_{N_s, \min}) = \left(\frac{3}{4\pi n} \right)^{1/3} \frac{N_p}{[(N_s - 1)!]^{N_p}} \int_0^\infty e^{-N_p y} y^{N_p(N_s - 1) + 1/3} dy \quad (\text{B.10})$$

Making the change of variables $t = N_p y$, we find

$$E(\hat{r}_{N_s, \min}) = \left(\frac{3}{4\pi n} \right)^{1/3} \frac{N_p}{[(N_s - 1)!]^{N_p} N_p^{N_p(N_s - 1) + 4/3}} \int_0^\infty e^{-t} t^{N_p(N_s - 1) + 1/3} dt \quad (\text{B.11})$$

which, using the definition of the complete gamma function, is

$$E(\hat{r}_{N_s, \min}) = \left(\frac{3}{4\pi n} \right)^{1/3} \frac{\Gamma[N_p(N_s - 1) + 4/3]}{[(N_s - 1)!]^{N_p} N_p^{N_p(N_s - 1) + 4/3}} = \left(\frac{3}{4\pi n} \right)^{1/3} \frac{[N_p(N_s - 1) + 1/3]!}{[(N_s - 1)!]^{N_p} N_p^{N_p(N_s - 1) + 1/3}} \quad (\text{B.12})$$

Now we use Stirling's approximation to simplify the factorials in the ratio. We define $N_\alpha \equiv N_p(N_s - 1)$ to keep things shorter:

$$(N_\alpha + 1/3)! \approx (N_\alpha + 1/3)^{N_\alpha + 1/3} e^{-(N_\alpha + 1/3)} \quad (\text{B.13})$$

$$[(N_s - 1)!]^{N_p} \approx [(N_s - 1)^{N_s - 1} e^{-(N_s - 1)}]^{N_p} = (N_s - 1)^{N_\alpha} e^{-N_\alpha} \quad (\text{B.14})$$

Substituting these two approximations back into the expression for $E(\hat{r}_{N_s, \min})$, we find

$$E(\hat{r}_{N_s, \min}) = \left(\frac{3}{4\pi n} \right)^{1/3} \frac{(N_\alpha + 1/3)^{N_\alpha + 1/3} e^{-(N_\alpha + 1/3)}}{N_p^{N_\alpha + 1/3} (N_s - 1)^{N_\alpha} e^{-N_\alpha}} = \left(\frac{3(N_s - 1)}{4\pi n e} \right)^{1/3} \left(\frac{N_\alpha + 1/3}{N_\alpha} \right)^{N_\alpha + 1/3} \quad (\text{B.15})$$

In the limit $N_\alpha \gg 1/3$, the second term approaches $e^{1/3}$, so to leading order, $E(\hat{r}_{N_s, \min}) \propto N_s^{1/3}$:

$$E(\hat{r}_{N_s, \min}) \approx \left(\frac{3(N_s - 1)}{4\pi n} \right)^{1/3} \quad (\text{B.16})$$

The *minimum* expected distance to the N_s^{th} particle is roughly equal to the *average* expected distance to the $(N_s - 1)^{\text{th}}$ particle, in the limit where both N_s and N_p are

much greater than 1.

Although the criterion $N_p \gg 1$ is generally satisfied, we are interested in values of N_s that do *not* satisfy the criterion $N_s \gg 1$: in fact, we wish to use the smallest value of N_s possible while retaining a good RMS error. We must therefore estimate the scaling with N_s in our region of interest by tabulating values of the integral in Equation (B.7) at various values of N_s and N_p . The scaling in the region of interest can then be approximated by fitting the values in the region of interest to a power law whose index is a free parameter. We find that for $2 \leq \log_{10} N_p \leq 5$ and $10 \leq N_s \leq 50$, the N_s -dependence roughly obeys a power law $\hat{r}_{N_s, \min} \propto N_s^\gamma$ with index $\gamma = 0.51 \pm 0.06$, where γ depends slightly on N_p . We confirm that the scaling of $E(\hat{r}_{N_s, \min})$ with N_p is suitably consistent with the prediction (Figure B-1).

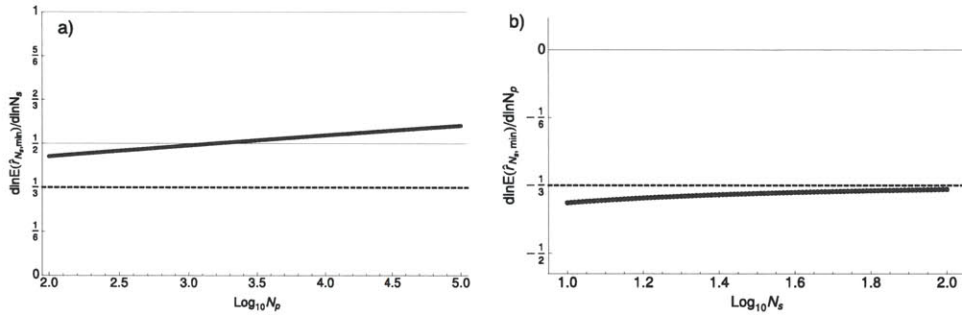


Figure B-1: The N_s -dependence of $E(\hat{r}_{N_s, \min})$ (a, thick line) is not close to the asymptotic prediction (dashed line) in the range of interest, and varies somewhat with N_p . We take an average slope of $1/2$ (solid thin line) for the power law index in N_s . However, the N_p -dependence of $E(\hat{r}_{N_s, \min})$ (b) is close to and approaches the asymptotic prediction (dashed line), so we use the asymptotic slope of $-1/3$ for the N_p scaling relation.

Appendix C

Technical information

This appendix contains technical details about the various algorithms, computer programs, and numerical methods used in this thesis.

C.1 N-body integration

In this work, the dynamics of systems of dark matter particles and/or stars is simulated using N-body integration. Numerous different methods for carrying out such integrations exist, but the ones used in this thesis all have a few things in common.

C.1.1 Smoothing

In all cases in this thesis the gravitational force between two bodies in the simulation is calculated using a smoothed version of Newton's law. At distances smaller than the smoothing length, the force is modified to avoid the singularity at zero distance as well as unphysical interactions between pairs of bodies (since in all cases each body in the integration represents many stars or dark matter particles). A simple version of this modification replaces Newton's force law with this form, known as the Plummer model:

$$\vec{F}_G = \frac{Gm_1m_2}{(d_{12}^2 + \epsilon_s^2)^{3/2}} \vec{d}_{12} \quad (\text{C.1})$$

where G is the gravitational constant, m_1 and m_2 are the masses of the two bodies, \vec{d}_{12} is the distance between them, and ϵ_s is the softening length. In the limit $\epsilon_s \rightarrow 0$, and for $d_{12} \gg \epsilon_s$, this expression is equivalent to the unmodified force law.

C.1.2 Force computation

Computing the forces between all N_p bodies in a simulation pairwise is prohibitively time-consuming in large simulations since the effort of this procedure scales as N_p^2 . In practice, most simulations of more than a few thousand bodies compute pairwise forces only between bodies that are very close together, and approximates the force contributions from more distant bodies with a multipole expansion. All the simulations in this thesis use a method that organizes the force computation with a data structure known as a Barnes-Hut tree [109]. Each node of the tree corresponds to a region in space, and the nodes are linked together in levels by repeatedly dividing the region in half in each dimension. The top level consists of one node representing the whole simulation volume, while at the lowest level are N_p “leaves” that each represent a region containing one body. To decide how accurately to compute the force on one of the bodies, the algorithm considers the solid angle subtended by the region associated with each node in the tree from the point of view of the body. If the solid angle subtended by a given node is large, the algorithm descends to the next level of subdivisions and checks again. For the closest bodies it subdivides all the way to the leaves and computes forces pairwise; otherwise, it subdivides until the solid angle subtended by a node is smaller than some threshold β , usually called the “opening angle”, and then computes the force on the body in question from a multipole expansion of the gravitational force from all the mass in that node. Using this algorithm, the forces on all the particles can be calculated more efficiently, so that the time to perform the calculation scales as $N_p \ln N_p$. To put this in perspective, a simulation with 1 million particles will run 10^5 times faster using a tree code than it will using direct summation! Compared to the amount of error introduced by using a softening length, the error from this approximation is usually fairly small for standard values of β around 0.4 (about 20 degrees).

C.1.3 Time integration

The N-body system is advanced forward in time by integrating the equations of motion in discrete steps of some length Δt . The simulations in this thesis use an integration method that staggers the advancement of the positions and momenta of the bodies by $\Delta t/2$. In this method, incrementing the momenta is called a “kick” and incrementing the positions is referred to as a “drift”. The timestepping scheme starts by kicking for $\Delta t/2$ to offset the steps, and finishes by kicking for half a step to catch up. Starting with initial conditions (x_0, p_0) , the integration for each body, simplified to one dimension, proceeds as follows:

$$\text{Kick half a timestep: } p(\Delta t/2) = p_0 + F(x_0)\Delta t/2 \quad (\text{C.2})$$

$$\text{Drift one timestep: } x(\Delta t) = x_0 + \frac{p(\Delta t/2)\Delta t}{m} \quad (\text{C.3})$$

$$\text{Kick one timestep: } p(3\Delta t/2) = p(\Delta t/2) + F[x(\Delta t)]\Delta t \quad (\text{C.4})$$

$$\text{Repeat the last two steps: } \vdots \quad \vdots \quad (\text{C.5})$$

$$\text{Drift one timestep: } x(t_f) = x(t_f - \Delta t) + p(t_f - \Delta t/2)\Delta t \quad (\text{C.6})$$

$$\text{Kick half a timestep: } p(t_f) = p(t_f - \Delta t/2) + F[x(t_f)]\Delta t/2. \quad (\text{C.7})$$

where m is the mass of the body and F is the force. If Δt is the same for all bodies over the entire integration time, this algorithm (known as “kick-drift-kick”) has a truncation error in the energy of order Δt^2 even though the integration is only to first order in Δt (Gadget 2 paper). Furthermore, the discrete equations of motion used for this time-integration scheme can be derived from a Hamiltonian (different from the continuous Hamiltonian), so the error in the integration is a Hamiltonian as well. This means that the errors on conserved quantities in the real physical system (like the energy, angular momentum, and Poincaré invariant) are bounded. These features have led to widespread use of symplectic integrators like this one in N-body simulations.

In practice, the wide range of densities (and therefore timescales) encountered in a typical N-body simulation make it impractical either to use the same value

of Δt for all particles in the system or to use the same value for the duration of the simulation. The symplectic algorithm described above can be modified to use different timestep lengths for each particle as long as they are all related by factors of two: it is no longer exactly symplectic, but still performs better than standard methods for integrating differential equations [110]. Changing the length of the base timestep as the simulation proceeds also destroys the exact symplecticity of the algorithm but the real deviations are again fairly slight. The usual criterion for the timestep length is that the distance over which the particle is accelerated during one timestep must be smaller than one softening length:

$$\frac{1}{2}|a|\Delta t^2 < \eta^2\epsilon_s, \text{ or } \Delta t < \eta\sqrt{\frac{2\epsilon_s}{|a|}} \quad (\text{C.8})$$

C.2 Computer programs used in this thesis

C.2.1 Serial tree code

For small-scale numerical simulations, we adapted an N-body code written by Will Farr. The code used a Barnes-Hut tree for force calculations and a single adaptive timestep length for all particles in the system following the criterion in Equation (C.8), advancing the particles with the kick-drift-kick symplectic scheme. The computationally-intensive parts of the code were written in the C programming language and called from a wrapper program written in Scheme. The code was modified to include an analytic part in the force calculations that could represent the static potential of the host galaxy. This program was written to be run on only one CPU at a time.

C.2.2 Axisymmetric potentials

To represent a realistic galactic potential for the host galaxy, we used the ‘‘GalPot’’ library of functions accompanying [111]. These functions construct multipole expansions of axisymmetric potentials that are used to compute the force at any point.

There are three types of potentials represented that are intended to correspond to the density distribution in different parts of a typical spiral galaxy. All of them depend mainly on the modified radius variable μ , defined in terms of the equatorial radius R and the height z out of the equatorial plane, and the flattening parameter q , so that for a given component i ,

$$\mu_i \equiv \sqrt{R^2 + z^2/q_i}. \quad (\text{C.9})$$

The flattening parameter is 1 for a spherical distribution, between 0 and 1 for an oblate distribution, and larger than 1 for a prolate distribution.

The bulge or nucleus component, which is thought to dominate at small galactocentric radii, is represented by an exponentially truncated power law in density with central density $\rho_{b,0}$, truncation radius r_b , scale radius a_b , and power-law index α_b :

$$\rho_{\text{bulge}}(R, z) = \rho_{b,0} \left(\frac{\mu_b}{a_b} \right)^{-\alpha_b} e^{-\mu_b^2/r_b^2} \quad (\text{C.10})$$

The disk component is intended to represent the density distribution in the disk of a spiral galaxy (though without the spiral arms). Old stars, young stars, and gas all seem to form disks with different scale heights z_d , lengths R_d , and surface densities Σ_d , but all display an exponential drop-off in both R and z . A multi-component disk can be constructed by superposition of the form:

$$\rho_{\text{disk}}(R, z) = \frac{\Sigma_d}{2z_d} \exp\left(\frac{-R}{R_d} - \frac{|z|}{z_d}\right) \quad (\text{C.11})$$

The halo component, intended to represent the distribution of dark matter in the galaxy, is represented by a double power-law in density, with indices α_h and β_h , central density $\rho_{h,0}$, and scale radius a_h :

$$\rho_{\text{halo}}(R, z) = \rho_{h,0} \left(\frac{m}{a_h} \right)^{-\alpha_h} \left(1 + \frac{m}{a_h} \right)^{\alpha_h - \beta_h} \quad (\text{C.12})$$

The NFW profile [95, 94], which adequately describes the density profiles of dark

matter halos in cosmological simulations, corresponds to $\alpha_h = 1$, $\beta_h = 3$.

The code to calculate the forces from these density distributions is written in C++. We modified the serial tree code described in the previous section to include forces calculated with this library to the interbody gravitational forces calculated with the tree.

C.2.3 Self-consistent N-body realizations

A few of the simulations done for this thesis used an N-body realization of the host halo as well as the satellite galaxy, in order to estimate the magnitude and type of error introduced by ignoring the gravitational response of the host galaxy. Using a “live” host halo is extremely computationally costly because of the inherent instability of discrete representations of smooth density distributions. To avoid unphysical effects, we use the rule of thumb that the number of bodies used to represent the halo must be at least large enough that the relaxation time of the halo is several times longer than the run time of the simulation. The relaxation time is derived by calculating the rate of close encounters between bodies within the halo and estimating how many close encounters would be needed to change a typical body’s velocity significantly [112], so it essentially measures the time over which the discreteness of the halo would start to have large-scale effects on its phase space distribution. We focus on the core of the halo because in our case, we are simulating mergers in which the satellite galaxy has multiple close encounters with the host’s core, so it is crucial that these close encounters are not affected by large-scale discreteness effects. Thus in the cases studied in this thesis we use

$$t_{\text{relax}} \approx \frac{0.1 N_p r_c}{\ln N_p \sigma_c} \quad (\text{C.13})$$

for the relaxation time, where r_c refers to the halo’s core radius and σ to the velocity dispersion in the core, so the ratio approximates the time it takes for a body in the core of the halo to make one orbit. For a simulation that runs for one billion years, a halo with the characteristic size and mass of the Milky Way must use 5 to 10 million particles (Figure C-1).

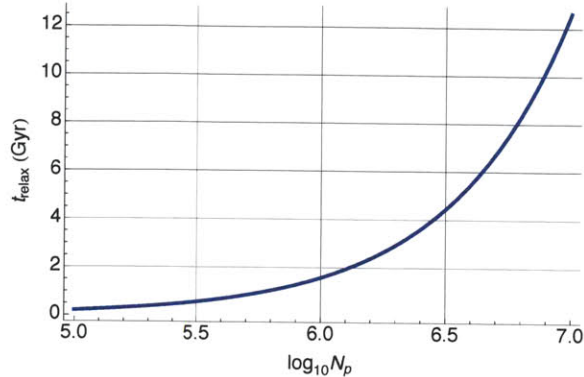


Figure C-1: Relaxation time in the core of a Milky-Way-like halo as a function of resolution.

An additional difficulty in using a live halo is in generating a self-consistent random realization. Most plausible models of the density distribution of dark matter halos are infinite in extent; truncating them at some finite radius to construct a random realization does not work because the phase space distribution function (DF) for the infinite version is then inconsistent with the density distribution. Instead, a DF for the truncated version must be constructed from scratch and used to generate the random realization. For this work we used the DF for the lowered Evans model [104], which is a truncated version of the cored axisymmetric logarithmic potential

$$\Psi(\mu) = \sigma^2 \ln \left(\frac{r_c^2 + \mu^2}{r_h^2} \right) \quad (\text{C.14})$$

with characteristic velocity σ , core radius r_c , and scale radius r_h [113]. GalactICs, a program that constructs random realizations of this model, as well as a multipole expansion of the associated self-consistent potential, force, and density functions, was publicly released by [114]. We used this program, written in a combination of Fortran and C, to generate random realizations of a host halo for use in our simulations.

C.2.4 GADGET

Because such high numbers of particles are required for a successful simulation with a live halo, we chose to use the publicly released parallel N-body integration code

Gadget-2 [115]. This program has multiple modes for different types of integrations; we used it in mode 1 (vacuum boundary conditions, flat Newtonian space), which uses a Barnes-Hut tree as described above but also parallelizes the force calculations by distributing them over multiple processors. Gadget-2 uses a kernel to smooth the gravitational force instead of Equation (C.1); the kernel is normalized so that the smoothing length ϵ_s is consistent with its meaning in Plummer-smoothed simulations but the force is exactly Newtonian for $d_{12} > 2.8\epsilon_2$. The criterion for subdividing a region of the tree is also slightly different: instead of a constant opening angle the code estimates the relative force error produced by treating a given node as one point mass and adjusts the opening angle to keep this relative error constant. This criterion made the force calculation more efficient [115]. The integration is performed with the kick-drift-kick scheme, with adjustments for individual and adaptive timesteps. To check for consistency we compared small- N_p simulations using Gadget-2 and the serial tree code and obtained the same result with both codes.

C.2.5 Beowulf Cluster

Parallel simulations were run on the Beowulf cluster co-sponsored by the MIT Kavli Institute for Space Research and the Astrophysics Division of the Department of Physics. Computer programs running on the cluster are scheduled using the Platform Lava job management software (<http://www.hpccommunity.org/section/lava-5>) installed as part of the Kusu "Platform OCS" 5 cluster management software package (<http://www.hpccommunity.org/section/kusu-45>). The cluster consists of 24 Linux machines (worker nodes) with 8 CPUs each for a total of 192 cores, connected by gigabit Ethernet cables, and a server machine (queen node) running Red Hat Enterprise Linux Server release 5.3. 15 of the machines have 2.0 GHz CPUs and the remaining 9 have 2.5 GHz CPUs. Gadget-2 performs best on a homogeneous cluster, so for this work we used the 9 2.5-GHz machines only (72 cores). Parallel Gadget-2 uses the Message Passing Interface (a copy of the standard is available at <http://www.mcs.anl.gov/research/projects/mpi/mpi-standard/mpi-report-2.0/mpi2-report.htm>) to send commands between cores; the work in this thesis was run using the

Lam-MPI implementation of the interface (version 7.1.2; <http://www.lam-mpi.org>). The author thanks Paul Hsi for administration of the cluster and for troubleshooting assistance (this was the first time that the current incarnation of the cluster was used for a massively parallel application).

C.3 Downhill Simplex Optimization Method

Downhill simplex optimization, also known as the Nelder-Mead method [107], is a method of optimizing a merit function without computing its derivatives. The method is particularly useful when the merit function is relatively quick to evaluate and the derivatives are not well-behaved. It does not converge efficiently but is very robust. “Simplex” refers to the geometrical object comprising a vertex in N -dimensional space (in this case, a particular set of N parameters) and the edges leading to N neighboring vertices. In the case of this method, the edges are oriented along the N directions in parameter space. The simplex is initialized with a guess (the starting vertex) and initial edge lengths in each parameter direction (search radii for each parameter). In the algorithm, the simplex is represented as a set of $N + 1$ points in parameter space. The figure of merit (the thing being minimized) is calculated at each of these points and they are sorted from lowest to highest; then the edge leading to the highest point in the simplex is reflected through the face containing the lowest point (that is, the simplex is aimed “downhill”) and the procedure is repeated. In practice, this step is taken as a trial; if it does not reduce the figure of merit the simplex is also contracted or expanded. Calculating the figure of merit at a set of neighboring points is equivalent to a very rough estimate of the local partial derivatives of the merit function by finite differencing; through this procedure, edges of the simplex along steeply sloped directions are contracted while those along shallow slopes are expanded to improve the convergence.

One shortcoming of most multidimensional minimizers is their tendency to stop before finding a true local minimum, especially in cases where there is a long shallow “valley” in the merit function (this corresponds to a degeneracy between two or more

parameters). In the case of the downhill can be addressed by restarting the optimizer from the convergence point with re-initialized edge lengths.

We used the implementation of the downhill simplex method from [106], called “amoeba”. Originally written in Fortran, we used the version translated into C with two modifications: an increase in the floating-point precision of the optimizer (from float to double!) and the addition of a monitoring function that printed the current best-fit parameters and figure of merit at each step taken by the optimizer.

Bibliography

- [1] F. Zwicky. Die Rotverschiebung von extragalaktischen Nebeln. *Helvetica Physica Acta*, 6:110–127, 1933.
- [2] H. J. Rood, V. C. A. Rothman, and B. E. Turnrose. Empirical Properties of the Mass Discrepancy in Groups and Clusters of Galaxies. *ApJ*, 162:411–+, November 1970.
- [3] J. P. Ostriker, P. J. E. Peebles, and A. Yahil. The size and mass of galaxies, and the mass of the universe. *ApJ Letters*, 193:L1–L4, 1974.
- [4] D. Larson, J. Dunkley, G. Hinshaw, E. Komatsu, et al. Seven-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Power Spectra and WMAP-Derived Parameters. *arXiv:1001.4635*, January 2010.
- [5] M. Tegmark, D. J. Eisenstein, M. A. Strauss, D. H. Weinberg, et al. Cosmological constraints from the SDSS luminous red galaxies. *Phys. Rev. D*, 74(12):123507, 2006.
- [6] A. G. Riess, A. V. Filippenko, P. Challis, A. Clocchiatti, et al. Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant. *AJ*, 116:1009–1038, 1998.
- [7] B. P. Schmidt, N. B. Suntzeff, M. M. Phillips, R. A. Schommer, et al. The High-Z Supernova Search: Measuring Cosmic Deceleration and Global Curvature of the Universe Using Type IA Supernovae. *ApJ*, 507:46–63, 1998.
- [8] G. Bertone, D. Hooper, and J. Silk. Particle dark matter: evidence, candidates and constraints. *Physics Reports*, 405:279–390, 2004.
- [9] M. Battaglia, A. De Roeck, J. Ellis, F. Gianotti, et al. Updated post-wmap benchmarks for supersymmetry. *European Physical Journal C*, 33:273, 2004.
- [10] P. Gondolo, J. Edsjö, P. Ullio, L. Bergström, M. Schelke, and E. A. Baltz. DarkSUSY: computing supersymmetric dark matter properties numerically. *J. Cosm. Astropart. Phys.*, 7:8, 2004.
- [11] Z. Ahmed et al. Search for weakly interacting massive particles with the first five-tower data from the cryogenic dark matter search at the soudan underground laboratory. *Phys. Rev. Letters*, 102:011301, 2009.

- [12] J. Angle et al. Constraints on inelastic dark matter from xenon10. *Phys. Rev. D*, 80:115005, 2009.
- [13] A. A. Abdo et al. Observations of Milky Way dwarf spheroidal galaxies with the Fermi-Large Area Telescope detector and constraints on dark matter models. *ApJ*, 712:147–158, 2010.
- [14] J. Chang, J. H. Adams, H. S. Ahn, G. L. Bashindzhagyan, et al. An excess of cosmic ray electrons at energies of 300-800GeV. *Nature*, 456:362–365, 2008.
- [15] A. A. Abdo et al. Measurement of the cosmic ray positron+electron spectrum from 20GeV to 1TeV with the Fermi Large Area Telescope. *Phys. Rev. Letters*, 102:181101, 2009.
- [16] J. Lavalle et al. Indirect search for dark matter with the CELESTE experiment. astro-ph/0601298, January 2006.
- [17] C.-P. Ma and E. Bertschinger. Cosmological Perturbation Theory in the Synchronous and Conformal Newtonian Gauges. *ApJ*, 455:7, 1995.
- [18] R. Cen, N. Y. Gnedin, and J. P. Ostriker. A Hydrodynamic Treatment of the Cold Dark Matter Cosmological Scenario with a Cosmological Constant. *ApJ*, 417:387, 1993.
- [19] D. Spolyar, K. Freese, and P. Gondolo. Dark Matter and the First Stars: A New Phase of Stellar Evolution. *Physical Review Letters*, 100(5):051101, 2008.
- [20] M. Boylan-Kolchin, V. Springel, S. D. M. White, A. Jenkins, and G. Lemson. Resolving cosmic structure formation with the Millennium-II Simulation. *MNRAS*, 398:1150–1164, 2009.
- [21] M. Davis, G. Efstathiou, C. S. Frenk, and S. D. M. White. The evolution of large-scale structure in a universe dominated by cold dark matter. *ApJ*, 292:371–394, 1985.
- [22] B. Moore, S. Ghigna, F. Governato, G. Lake, et al. Dark Matter Substructure within Galactic Halos. *ApJ Letters*, 524:L19–L22, 1999.
- [23] J. F. Navarro, A. Ludlow, V. Springel, J. Wang, M. Vogelsberger, S. D. M. White, A. Jenkins, C. S. Frenk, and A. Helmi. The diversity and similarity of simulated cold dark matter haloes. *MNRAS*, 402:21–34, 2010.
- [24] M. Maciejewski, M. Vogelsberger, S. D. M. White, and V. Springel. Bound and unbound substructures in Galaxy-scale Dark Matter haloes. *arXiv:1010.2491*, October 2010.
- [25] M. Geha, M. R. Blanton, M. Masjedi, and A. A. West. The Baryon Content of Extremely Low Mass Dwarf Galaxies. *ApJ*, 653:240–254, 2006.

- [26] N. F. Martin, R. A. Ibata, S. C. Chapman, M. Irwin, and G. F. Lewis. A Keck/DEIMOS spectroscopic survey of faint Galactic satellites: searching for the least massive dwarf galaxies. *MNRAS*, 380:281–300, 2007.
- [27] J. D. Simon and M. Geha. The Kinematics of the Ultra-faint Milky Way Satellites: Solving the Missing Satellite Problem. *ApJ*, 670:313–331, 2007.
- [28] V. Belokurov, D. B. Zucker, N. W. Evans, G. Gilmore, et al. The Field of Streams: Sagittarius and Its Siblings. *ApJ Letters*, 642:L137–L140, 2006.
- [29] D. Martínez-Delgado, R. J. Gabany, K. Crawford, S. Zibetti, et al. Stellar Tidal Streams in Spiral Galaxies of the Local Volume: A Pilot Survey with Modest Aperture Telescopes. *AJ*, 140:962–967, 2010.
- [30] A. W. McConnachie, M. J. Irwin, R. A. Ibata, et al. The remnants of galaxy formation from a panoramic survey of the region around M31. *Nature*, 461:66–69, 2009.
- [31] J. S. Bullock and K. V. Johnston. Tracing Galaxy Formation with Stellar Halos. I. Methods. *ApJ*, 635:931–949, 2005.
- [32] M. A. Fardal, P. Guhathakurta, A. Babul, and Alan W. McConnachie. Investigating the Andromeda stream – III. a young shell system in M31. *MNRAS*, 380:15–32, 2007.
- [33] M. A. Fardal, A. Babul, J. J. Geehan, and P. Guhathakurta. Investigating the Andromeda stream – II. orbital fits and properties of the progenitor. *MNRAS*, 366:1012, 2006.
- [34] D. R. Law and S. R. Majewski. The Sagittarius Dwarf Galaxy: A Model for Evolution in a Triaxial Milky Way Halo. *ApJ*, 714:229–254, 2010.
- [35] E. Bertschinger. Self-similar secondary infall and accretion in an Einstein-de Sitter universe. *ApJS*, 58:39, 1985.
- [36] J. A. Fillmore and P. Goldreich. Self-similar gravitational collapse in an expanding universe. *ApJ*, 281:1–8, 1984.
- [37] R. Mohayaee and S. F. Shandarin. Gravitational cooling and density profile near caustics in collisionless dark matter haloes. *MNRAS*, 366:1217–1229, 2006.
- [38] S. F. Shandarin and Ya. B. Zeldovich. The large-scale structure of the universe: turbulence, intermittency, structures in a self-gravitating medium. *Reviews of Modern Physics*, 61(2):185–222, 1989.
- [39] S. Tremaine. The geometry of phase-mixing. *MNRAS*, 307:877–883, 1999.
- [40] S. Kaviraj. Peculiar early-type galaxies in the Sloan Digital Sky Survey Stripe82. *MNRAS*, 406:382–394, 2010.

- [41] D. F. Malin and D. Carter. A catalog of elliptical galaxies with shells. *ApJ*, 274:534–540, 1983.
- [42] G. Sikkema, D. Carter, R. F. Peletier, M. Balcells, et al. HST/ACS observations of shell galaxies: inner shells, shell colours and dust. *A&A*, 467:1011–1024, 2007.
- [43] R. E. Sanderson and E. Bertschinger. Seen and unseen tidal caustics in the Andromeda galaxy. *ApJ*, 725:1625, 2010.
- [44] S. J. Asztalos et al. Improved rf cavity search for halo axions. *Phys. Rev. D*, 69:011101, 2004.
- [45] D. Kinion, I. G. Irastorza, and K. van Bibber. Searches for Astrophysical and Dark Matter Axions. *Nuclear Physics B Proceedings Supplements*, 143:417–422, 2005.
- [46] T. Gerialis. Cast results and axion review. for the CAST collaboration. arXiv:0905.4273, 2009.
- [47] Z. Ahmed et al. Search for axions with the cdms experiment. *Phys. Rev. Letters*, 103:141802, 2009.
- [48] J. L. Feng. Dark Matter Candidates from Particle Physics and Methods of Detection. *Ann. Rev. Astron. Astrophys.*, 48:495–545, 2010.
- [49] E. A. Baltz, M. Battaglia, M. E. Peskin, and T. Wizansky. Determination of dark matter properties at high-energy colliders. *Phys. Rev. D*, 74(10):103521, 2006.
- [50] D. Hooper and E. A. Baltz. Strategies for Determining the Nature of Dark Matter. *Annual Review of Nuclear and Particle Science*, 58:293–314, 2008.
- [51] O. Adriani et al. An anomalous positron abundance in cosmic rays with energies 1.5–100 GeV. *Nature*, 458:607, 2009.
- [52] O. Adriani et al. New measurement of the antiproton-to-proton flux ratio up to 100 GeV in the cosmic radiation. *Phys. Rev. Letters*, 102:051101, 2009.
- [53] P. Scott, J. Conrad, J. Edsjö, L. Bergström, et al. Direct constraints on minimal supersymmetry from Fermi-LAT observations of the dwarf galaxy Segue 1. *J. Cosm. Astropart. Phys.*, 1:31, 2010.
- [54] L. Bergström, P. Ullö, and J. H. Buckley. Observability of gamma rays from dark matter neutralino annihilations in the Milky Way halo. *Astroparticle Physics*, 9:137–162, 1998.
- [55] A. A. Abdo et al. Fermi/Large Area Telescope Bright Gamma-Ray Source List. *ApJS*, 183:46–66, 2009.

- [56] L. E. Strigari, S. M. Koushiappas, J. S. Bullock, and M. Kaplinghat. Precise constraints on the dark matter content of milky way dwarf galaxies for gamma-ray experiments. *Phys. Rev. D*, 75(8):083526, 2007.
- [57] S. Peirani, R. Mohayaee, and J. A. de Freitas Pacheco. Indirect search for dark matter: prospects for glast. *Phys. Rev. D*, 70:043503, 2004.
- [58] J. Diemand, M. Kuhlen, and P. Madau. Dark matter substructure and gamma-ray annihilation in the Milky Way halo. *ApJ*, 657:262, 2007.
- [59] M. Kuhlen, J. Diemand, and P. Madau. Glast and dark matter substructure in the Milky Way. In S. Ritz, P. Michelson, and C. Meegan, editors, *The first GLAST symposium*, pages 135–138. American Institute of Physics, 2007.
- [60] R. Mohayaee, S. Shandarin, and J. Silk. Dark matter caustics and the enhancement of self-annihilation flux. *J. Cosm. Astropart. Phys.*, 5:15, 2007.
- [61] T. Bringmann, L. Bergström, and J. Edsjö. New gamma-ray contributions to supersymmetric dark matter annihilation. *Journal of High Energy Physics*, 1:49, 2008.
- [62] L. Pieri, G. Bertone, and E. Branchini. Dark matter annihilation in substructures revised. *MNRAS*, 384:1627, 2008.
- [63] M. Vogelsberger, S. D. M. White, A. Helmi, and V. Springel. The fine-grained phase-space structure of cold dark matter halos. *MNRAS*, 385:236, 2008.
- [64] L. Hernquist and P. J. Quinn. Formation of shell galaxies. i. spherical potentials. *ApJ*, 331:682–698, 1988.
- [65] L. Hernquist and P. J. Quinn. Formation of shell galaxies. ii. nonspherical potentials. *ApJ*, 342:1–16, 1989.
- [66] C. J. Hogan. Particle annihilation in cold dark matter micropancakes. *Phys. Rev. D*, 64(6):063515, 2001.
- [67] P. Sikivie. Evidence for ring caustics in the Milky Way. *Physics Letters B*, 567:1–8, 2003.
- [68] L. Pieri and E. Branchini. γ -ray flux from dark matter annihilation in galactic caustics. *J. Cosm. Astropart. Phys.*, 5:7, 2005.
- [69] A. Natarajan and P. Sikivie. Inner caustics of cold dark matter halos. *Phys. Rev. D*, 73(2):023510, 2006.
- [70] A. Natarajan. Weakly interacting massive particle annihilation in caustics. *Phys. Rev. D*, 75(12):123514, 2007.
- [71] A. Natarajan and P. Sikivie. Further look at particle annihilation in dark matter caustics. *Phys. Rev. D*, 77(4):043531, 2008.

- [72] L. D. Duffy and P. Sikivie. Caustic ring model of the MilkyWay halo. *Phys. Rev. D*, 78(6):063508, 2008.
- [73] N. Afshordi, R. Mohayaee, and E. Bertschinger. Hierarchical phase space structure of dark matter haloes: Tidal debris, caustics, and dark matter annihilation. *Phys. Rev. D*, 79(8):083526, 2009.
- [74] R. Rando and for the Fermi LAT Collaboration. Post-launch performance of the Fermi Large Area Telescope. *arXiv:0907.0626*, July 2009.
- [75] A. J. Izenman. Nonparametric density estimation. In G. Casella, S. Fienberg, and I. Olkin, editors, *Modern Multivariate Statistical Techniques*, Springer Texts in Statistics, chapter 4. Springer New York, 2008.
- [76] P. Hall and J. S. Marron. Estimation of integrated squared density derivatives. *Statistics & Probability Letters*, 6(2):109 – 115, 1987.
- [77] P. J. Bickel and Y. Ritov. Estimating integrated squared density derivatives: Sharp best order of convergence estimates. *Sankhyā: The Indian Journal of Statistics, Series A*, 50(3):381–393, 1988.
- [78] L. Birgé and P. Massart. Estimation of integral functionals of a density. *The Annals of Statistics*, 23(1):11–29, 1995.
- [79] B. Laurent. Efficient estimation of integral functionals of a density. *The Annals of Statistics*, 24(2):659–681, 1996.
- [80] H. V. Martinez and M. M. Olivares. Estimation of quadratic functionals of a density. *Statistics & Probability Letters*, 42(4):327–332, 1999.
- [81] E. Giné and D. M. Mason. Uniform in bandwidth estimation of integral functionals of the density function. *Scandinavian Journal of Statistics*, 35(4):739 – 761, 2008.
- [82] E. Giné and R. Nickl. A simple adaptive estimator of the integrated square of a density. *Bernoulli*, 14(1):47–61, 2008.
- [83] E. Tchetgen, L. Li, J. Robins, and A. van der Vaart. Minimax estimation of the integral of a power of a density. *Statistics & Probability Letters*, 78(18):3307–3311, 2008.
- [84] E. Fix and J. L. Hodges. Discriminatory analysis. nonparametric discrimination: Consistency properties. *International Statistical Review / Revue Internationale de Statistique*, 57(3):238–247, 1989.
- [85] D. O. Loftsgaarden and C. P. Quesenberry. A nonparametric estimate of a multivariate density function. *The Annals of Mathematical Statistics*, 36(3):1049–1051, 1965.

- [86] L. P. Devroye and T. J. Wagner. The strong uniform consistency of nearest neighbor density estimates. *The Annals of Statistics*, 5(3):536–540, 1977.
- [87] D. S. Moore and J. W. Yackel. Consistency properties of nearest neighbor density function estimators. *Annals of Statistics*, 5:1, 1977.
- [88] Y. P. Mack and M. Rosenblatt. Multivariate k-nearest neighbor density estimates. *Journal of Multivariate Analysis*, 9(1):1 – 15, 1979.
- [89] A. J. Izenman. Recent developments in nonparametric density estimation. *Journal of the American Statistical Association*, 86(413):205–224, 1991.
- [90] T. J. Wu, C. F. Chen, and H. Y. Chen. A variable bandwidth selector in multivariate kernel density estimation. *Statistics & Probability Letters*, 77:462–467, 2007.
- [91] B. W. Lindgren. *Statistical Theory*. MacMillan, New York, 3rd edition, 1976.
- [92] Y. Ascasibar and J. Binney. Numerical estimation of densities. *MNRAS*, 356:872–882, 2005.
- [93] J. J. Geehan, M. A. Fardal, A. Babul, and P. Guhathakurta. Investigating the Andromeda stream – I. simple analytic bulge-disc-halo model for M31. *MNRAS*, 366:996, 2006.
- [94] Julio F. Navarro, Carlos S. Frenk, and Simon D. M. White. The structure of cold dark matter halos. *ApJ*, 462:563, 1996.
- [95] J. F. Navarro, C. S. Frenk, and S. D. M. White. A Universal Density Profile from Hierarchical Clustering. *ApJ*, 490:493, 1997.
- [96] N. Fornengo, L. Pieri, and S. Scopel. Neutrino annihilation into gamma rays in the milky way and in external galaxies. *Phys. Rev. D*, 70(10):103529, 2004.
- [97] M. P. Wand. Data-based choice of histogram bin width. *The American Statistician*, 51:59–64, 1996.
- [98] M. L. M. Collins, S. C. Chapman, M. J. Irwin, N. F. Martin, et al. A keck/deimos spectroscopic survey of the faint m31 satellites andix, andxi, andxii and andxiii. *MNRAS*, 407:2411–2433, 2010.
- [99] L. Hernquist and P. J. Quinn. Shell galaxies and alternatives to the dark matter hypothesis. *ApJ*, 312:17–21, 1987.
- [100] L. Hernquist and P. J. Quinn. Shells and dark matter in elliptical galaxies. *ApJ*, 312:1–16, 1987.
- [101] R. Ibata, M. Irwin, G. Lewis, et al. A giant stream of metal-rich stars in the halo of the galaxy M31. *Nature*, 412:49–52, 2001.

- [102] P. Seguin and C. Dupraz. Dynamical friction in head-on galaxy collisions. II. N-body simulations of radial and non-radial encounters. *A&A*, 310:757–770, 1996.
- [103] M. Mori and R. M. Rich. The Once and Future Andromeda Stream. *ApJ Letters*, 674:L77–L80, 2008.
- [104] K. Kuijken and J. Dubinski. Lowered Evans Models - Analytic Distribution Functions of Oblate Halo Potentials. *MNRAS*, 269:13, 1994.
- [105] N. W. Evans. Simple galaxy models with massive haloes. *MNRAS*, 260:191–201, 1993.
- [106] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical Recipes in C*, chapter 10, pages 408–412. Cambridge University Press, 2 edition, 1988.
- [107] J. A. Nelder and R. Mead. A Simplex Method for Function Minimization. *The Computer Journal*, 7(4):308–313, 1965.
- [108] S. Sharma and M. Steinmetz. Multidimensional density estimation and phase-space structure of dark matter haloes. *MNRAS*, 373:1293–1307, 2006.
- [109] J. Barnes and P. Hut. A hierarchical $O(N \log N)$ force-calculation algorithm. *Nature*, 324:446–449, 1986.
- [110] T. Quinn, N. Katz, J. Stadel, and G. Lake. Time stepping N-body simulations. *astro-ph/9710043*, October 1997.
- [111] W. Dehnen and J. Binney. Mass models of the Milky Way. *MNRAS*, 294:429, 1998.
- [112] J. Binney and S. Tremaine. *Galactic Dynamics*. Princeton Series in Astrophysics. Princeton University Press, 2 edition, 2008.
- [113] J. Binney. Resonant excitation of motion perpendicular to galactic planes. *MNRAS*, 196:455–467, 1981.
- [114] K. Kuijken and J. Dubinski. Nearly Self-Consistent Disc / Bulge / Halo Models for Galaxies. *MNRAS*, 277:1341, 1995.
- [115] V. Springel. The cosmological simulation code GADGET-2. *MNRAS*, 364:1105–1134, 2005.