SEEN AND UNSEEN TIDAL CAUSTICS IN THE ANDROMEDA GALAXY

R. E. Sanderson and E. Bertschinger
MIT Department of Physics and Kavli Institute for Space Research, Cambridge, MA 02139, USA; robyn@mit.edu
Received 2010 June 21; accepted 2010 October 19; published 2010 November 30

ABSTRACT

Indirect detection of high-energy particles from dark matter interactions is a promising avenue for learning more about dark matter, but is hampered by the frequent coincidence of high-energy astrophysical sources of such particles with putative high-density regions of dark matter. We calculate the boost factor and gamma-ray flux from dark matter associated with two shell-like caustics of luminous tidal debris recently discovered around the Andromeda galaxy, under the assumption that dark matter is its own supersymmetric antiparticle. These shell features could be a good candidate for indirect detection of dark matter via gamma rays because they are located far from the primary confusion sources at the galaxy’s center, and because the shapes of the shells indicate that most of the mass has piled up near the apocenter. Using a numerical estimator specifically calibrated to estimate densities in N-body representations with sharp features and a previously determined N-body model of the shells, we find that the largest boost factors do occur in the shells but are only a few percent. We also find that the gamma-ray flux is an order of magnitude too low to be detected with Fermi for likely dark matter parameters, and about two orders of magnitude less than the signal that would have come from the dwarf galaxy that produces the shells in the N-body model. We further show that the radial density profiles and relative radial spacing of the shells, in either dark or luminous matter, is relatively insensitive to the details of the potential of the host galaxy but depends in a predictable way on the velocity dispersion of the progenitor galaxy.

Key words: dark matter – galaxies: individual (M31) – galaxies: kinematics and dynamics – gamma rays: galaxies – methods: analytical – methods: numerical

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 (Bertone et al. 2005), there are to date no conclusive detections. Countless experiments are attempting to do so, whether by detecting dark matter particles directly (Asztalos et al. 2004; Kinion et al. 2005; Geralis et al. 2009; Angle et al. 2009; Ahmed et al. 2009a, 2009b; Feng 2010), creating them in the laboratory (Baltz et al. 2006; Hooper & Baltz 2008), or observing the standard-model-by-products of interactions between them (Chang et al. 2008; Adriani et al. 2009a, 2009b; Abdo et al. 2009a, 2010; Scott et al. 2010). 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 supersymmetric partner 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 (Bertone et al. 2005).

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 (Bergström et al. 1998). Although the signal from pairwise dark matter annihilation may be highest in those regions, it also suffers from confusion with high-energy astrophysical sources such as pulsars and X-ray binaries, which tend to be concentrated wherever the density of luminous matter is high (Abdo et al. 2009b). 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 (Strigari et al. 2007). 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 (Peirani et al. 2004; Diemand et al. 2007; Kuhlen et al. 2007; Mohayaee et al. 2007; Bringmann et al. 2008; Pieri et al. 2008; Vogelsberger et al. 2008). 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. Hernquist & Quinn (1988, 1989) 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. Fillmore & Goldreich (1984) and Bertschinger (1985) demonstrated that radial infall of gravitating, cold, collisionless matter forms a series of infinite-density peaks at successive radii, known as caustics. Mohayaee & Shandarin (2006) 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., Hogan 2001; Sikivie 2003; Pieri & Branchini 2005; Natarajan & Sikivie 2006; Mohayaee et al. 2007; Natarajan 2007; Natarajan & Sikivie 2008; Duffy & Sikivie 2008; Afshordi et al. 2009) or to otherwise determine ways in which dark matter in caustics could be detected (Duffy & Sikivie 2008; Natarajan & Sikivie 2006; Sikivie 2003). 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., Malin & Carter 1983) 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 (Fardal et al. 2007), 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 (Rando et al. 2009). Furthermore, an $N$-body model of the shells already exists (Fardal et al. 2006) 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 (Kuhlen et al. 2007). 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). 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 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. THE OPTIMAL ESTIMATOR FOR HIGH DENSITY CONTRAST

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

$$\Gamma = \int \rho^2 dV, \quad (1)$$

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 $\rho$. Generally, probability density functions are defined as those that are everywhere positive and normalized to one (as in Izenman 2008, chap. 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 (Hall & Marron 1987; Bickel & Ritov 1988; Birgé & Massart 1995; Laurent 1996; Martinez & Olivares 1999; Giné & Mason 2008; Giné & Nickl 2008; Tchetgen et al. 2008, and many others).

Density estimators studied in the literature are divided into two classes: parametric (in which a particular functional form for $\rho$ is assumed) and nonparametric (in which assumptions about the form of $\rho$ 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 $\rho$ and/or calculate other quantities from it (Izenman 2008). Among the wide variety of nonparametric estimators available, nearest-neighbor estimators (Fix & Hodges 1951) 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 $N_{th}$ nearest particle to the location where the density is
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, \( \Gamma_{\text{true}} \). For the tests in this work, we used density distributions for which \( \Gamma_{\text{true}} \) may be calculated analytically. Bickel & Ritov (1988) 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 \( N_p \to \infty \), provided that \( N_f/N_p \to 0 \) in the same limit. They can also be considered as part of the larger class of adaptive kernel estimators (Moore & Yackel 1977) and are even more closely related when \( V_d \) is replaced by a weighted sum over the \( N_f \) particles (Mack & Rosenblatt 1979). However, because the function they return may not be normalizable, nearest-neighbor methods are more suited to individual density estimates at a point than to recovery of the entire function (Izenman 1991). All the estimators we test in this work are based on either the simple nearest-neighbor 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-neighbor estimator is its root-mean-squared (rms) error,

\[
\text{r.m.s.e.} = \frac{1}{\Gamma_{\text{true}}} \sqrt{E[(\hat{\Gamma} - \Gamma_{\text{true}})^2]},
\]

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, \( \Gamma_{\text{true}} \). For the tests in this work, we used density distributions for which \( \Gamma_{\text{true}} \) may be calculated analytically. Bickel & Ritov (1988) 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 \( N_p^{-1/2} \); Giné & Nickl (2008) 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 simplier rule than the one suggested by Giné & Nickl 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 Giné & Nickl to several orthogonal dimensions is not trivial (Wu et al. 2007).

As is common in the literature, we consider the rms error in two parts: the bias and standard deviation (Lindgren 1976), where

\[
(r.m.s.e.)^2 = b^2 + (\text{std}(\hat{\Gamma}))^2.
\]

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 \( \Gamma_{\text{true}} \), so

\[
b = E\left(\frac{\hat{\Gamma} - \Gamma_{\text{true}}}{\Gamma_{\text{true}}}\right) = \frac{E(\hat{\Gamma})}{\Gamma_{\text{true}}} - 1
\]

The number of particles in each random realization (shown as points in an example in Figure 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 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 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 using improper rates can be eliminated or reduced by using a constant rate estimate of the density, which we discuss in detail in Appendix A.
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 2.1). We examined the rms error it has been measured 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).

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 2.3). 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 $\Gamma_{\text{true}}$ (Section 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 2.5).

### 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 (A5)). It uses a constant Riemann volume $dV$ for the integral and the distance to the $N_{\text{ns}}^\text{th}$ nearest-neighboring particle, denoted $r_{N_{\text{ns}}}^\text{th}$, 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 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), using smaller boxes in higher-density regions, but this method introduces bias because the box dimensions, like $r_{N_{\text{ns}}}^\text{th}$, are then subject to Poisson statistics. We chose $\hat{\Gamma}_n$ (Appendix A, Equation (A6)) 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 $E(n^2) \neq E(\hat{n}^2)$, 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 (A8), and $\hat{\Gamma}_s$, Equation (A14)). 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 (A15)). All the rate estimators are summarized in Table 1.

We calculated the expectation value and standard deviation of each rate estimator using random realizations of a uniform distribution, as described above. Equations (3) and (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 3). Furthermore, using an adaptive Riemann volume does change the bias (green diamonds in Figure 3). The more complicated estimators require

![Figure 2. Example of adaptive volume tessellation referred to in the text, produced by the space-filling tree in the FiEstAS estimator (Ascasibar & Binney 2005). The shading (colors from red to blue in online version) indicates 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.](Image 321x317 to 567x563)

### Table 1

Various Estimators of the Gamma-ray Emissivity $\Gamma$

| Estimator | Definition | Long Name | See Equation(s) |
|-----------|------------|-----------|----------------|
| $\hat{\Gamma}_u$ | $dV \sum_{i=1}^{N_{\text{p}}} n_{i,i}^2$ | Unbiased nearest-neighbor estimator with uniform box size | (A4), (A5) |
| $\hat{\Gamma}_n$ | $\propto N_{\text{p}} \hat{\Gamma}_n^2 dV_i$ | Unbiased nearest-neighbor estimator with adaptive box size | (A4), (A6) |
| $\hat{\Gamma}_f$ | $\propto \sum_{i=1}^{N_{\text{p}}} (\hat{f}_i(r_i))^2 dV_i$ | FiEstAS estimator (Ascasibar & Binney 2005) with adaptive box size | (A7), (A8) |
| $\hat{\Gamma}_s$ | $\propto \sum_{i=1}^{N_{\text{p}}} (\hat{n}_i(r_i))^2 dV_i$ | Epanechnikov kernel density estimator with adaptive box size | (A9), (A13), (A14) |
| $\hat{\Gamma}_d$ | $\propto \sum_{i=1}^{N_{\text{p}}} \hat{n}_i(r_i)$ | Method from Diemand et al. (2007) | (A15) |

**Note.** See Appendix A for longer descriptions of the various estimators.
Figure 3. All the rate estimators except for $\hat{\Gamma}_u$ (red circles) require slightly different but similarly shaped bias corrections. Equation (6) is an adequate ansatz for the $N_s$-dependence of the correction: an example corresponding to Equation (6) is shown as the (black) dashed line and the (green) filled points show the effect of using Equation (6) to correct $\hat{\Gamma}_f$. Fit parameters are listed in Table 2; Table 1 explains the label abbreviations.

Figure 4. Once estimators are corrected for Poisson bias using the fitting formula in Equation (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 1 explains the label abbreviations.

(A color version of this figure is available in the online journal.)

Table 2

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

Notes. Error ranges indicate the 95% confidence level. The form of the bias correction is given in Equation (7). See Appendix A and Table 1 for descriptions of the various estimators.

Even more substantial correction (cyan triangles, blue pentagrams, and purple hexagrams in Figure 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,

$$\frac{E[\hat{\Gamma}^2]}{E(n^2)} = \frac{N_s^2}{(N_s - 1)(N_s - 2)},$$

generalized to the form

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

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 3).

Fitting $E(\hat{\Gamma}_u)/\Gamma_{\text{true}}$ to a function of the form in Equation (7) determined $b_u$, 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. We used the fits to correct for the Poisson bias in all the remaining work discussed in this paper.

2.2. Performance of Estimators on the Uniform Distribution

All the estimators can be corrected to measure $\Gamma_{\text{true}}$ for a uniform density with an accuracy of better than 1%, although the $f$ and $s$ estimators have higher-order uncorrected behavior (Figure 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 5 to Figure 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 2.3.

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

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

and the standard deviations are related by the same factor:

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

From this last expression, we see that if the uncorrected estimator underestimates $\Gamma_{\text{true}}$, correcting it will increase the standard deviation. However, all the uncorrected estimators overestimate $\Gamma_{\text{true}}$ (Figure 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 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-neighbor estimator with a given \( N_s \), as can be clearly seen in Figure 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 \( \hat{\Gamma}_n \) 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-neighbor 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 (orthodidal) used in \( \hat{\Gamma}_s \); we discuss this possibility in Section 2.4. The rms error of \( \hat{\Gamma}_u \) 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 \( \hat{\Gamma}_n \) and \( \hat{\Gamma}_f \), on samples with high density contrast. These tests are described in the next few sections.

### 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

\[
E(\hat{r}_{N_s,\min}) \propto \left( \frac{N_s}{N_p} \right)^{1/3}, \quad N_s, N_p \gg 1. \tag{10}
\]

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 \( N_p \gg 10^2 \), but the scaling with \( N_s \) only approaches the asymptotic limit for values much too large to be practical (\( N_s \gg 10^4 \)). For reasonable values of \( N_s \), the power-law index must be determined numerically as discussed in Appendix B:

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

The corresponding maximum density then scales as

\[
n_{max} \propto \frac{N_s}{r_{N_s,\min}^3} \propto \frac{N_p}{N_s^{3\gamma-1}}. \tag{12}
\]

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 (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 as

\[
\hat{n}_{ul} = \begin{cases} \hat{n} & E(\hat{n}) < n_{max} \\ n_{max} & E(\hat{n}) > n_{max}. \end{cases}
\]

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 \( \hat{\Gamma}_{ul} \) is

\[
\hat{\Gamma}_{ul} = \hat{\Gamma} \bigg|_{n < n_{max}} + \frac{n_{max}^2 V_{n > n_{max}}}{b + 1}, \tag{14}
\]

where \( b \) is a factor with the form of Equation (7) and the appropriate fitted constants from Table 2. There is therefore an undersampling bias, \( b_{ul} \), in the rate estimator that depends on both \( N_p \) and \( N_s \). The \( N_p \)- and \( N_s \)-dependences enter two ways:
in the criterion for separating the Riemann sum and directly in the rate calculation for one of the terms:

\[ h_{ns} = \frac{\hat{r}_{ns}}{\Gamma_{true}} - 1 \]

\[ = E(\hat{r}) \frac{n_{\text{max}}^2 V_{n=n_{\text{max}}} V_{n=n_{\text{max}}}}{\Gamma_{true}(b+1)} - 1. \] (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 \( h_{ns} \leq 0 \). In the limit where the realization is fully resolved, the bias should be zero since \( V_{n=n_{\text{max}}} = 0 \), independent of \( N_s \) and \( N_p \).

In general, the \( N_p \) - and \( N_s \) - dependences in Equation (15) are complicated since some unknown fraction of the total volume is underresolved. To determine the undersampling bias, we tested \( \Gamma_{ns} \), \( \hat{r}_{ns} \), and \( \hat{r}_{s} \), including their respective corrections for Poisson bias, on \( N \)-body realizations of a one-dimensional caustic for which \( \Gamma_{true} \) can be calculated analytically. The caustic has an adjustable sharpness represented in terms of a velocity dispersion \( \sigma \): the smaller \( \sigma \) 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:

\[ \rho(x, t) = \frac{\rho_0}{\sqrt{2\pi \sigma^2 t^2}} \frac{|x-x_c|}{t} e^{-(x-x_c)^2/4\sigma^2 t^2} \times B \left[ \frac{(x-x_c)^2}{4\sigma^2 t^2} \right] \] (16)

with

\[ B(u) = \begin{cases} \frac{\pi}{\sqrt{2}} [I_{-1/4}(u) + I_{1/4}(u)] & x < x_c, \\ \frac{\pi}{\sqrt{2}} [I_{-1/4}(u) - I_{1/4}(u)] & x > x_c, \end{cases} \] (17)

where \( x_c = 1/4\alpha t \) is the position of the caustic and \( I \) is a modified Bessel function of the first kind. Table 3 briefly explains the parameters \( \rho_0, \alpha, \) and \( \sigma \) 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 Appendix C. Equation (16) is given in terms of the mass density \( \rho \), 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 \( \rho_0 \) 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 \( \Gamma_{true} \).

Given the values in Table 3, \( \Gamma_{true} \) can be calculated by performing a single numerical integral (Appendix C). 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 (16) for different \( \sigma \).

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 6). The general form of \( x(q_x) \) is given in Equation (C6). The form for a particular sample is determined by setting the parameters \( \alpha \) and \( t \), which also set the location of the peak density of the caustic, \( x_c \), and by setting \( \sigma \), the width of the normal distribution from which the random components of the particles’ initial velocities are drawn. The values of \( \alpha, t, \) and \( \sigma \) we used are summarized in Table 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 \( \Gamma_{true} \) to account for the incomplete sampling in these cases, as noted in Appendix C.

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 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 $\sigma$. Table 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%.

### 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 in our model caustic by the parameter $\sigma$.

As is expected, using higher resolution (a larger $N_p$) leads to better rate estimates of sharper features (Figure 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 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 underresolved, and only weakly when they are fully resolved (Figure 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 (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 underresolved regime and only very slightly affects the fully resolved case (Figure 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

![Figure 6](image-url)

**Figure 6.** 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 shaded (yellow in online version) 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 3.

(A color version of this figure is available in the online journal.)

![Figure 7](image-url)

**Figure 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 white (dark or bright blue in the online version) are considered fully resolved because the bias is less than the typical standard deviation of about 2%. In the online version, 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 $|b_{us}| > 0.2$.

(A color version of this figure is available in the online journal.)
Figure 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 underresolved features (b). If the realization is fully resolved, the increase in rms error is detectable but extremely small (b, inset).

(A color version of this figure is available in the online journal.)

Figure 9. 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).

(A color version of this figure is available in the online journal.)

as described in Section 2.2 of Ascasibar & Binney (2005) 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 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 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 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.

2.5. 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 \( \sigma \). 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 \( N \), dominates the rms error. We have already established that a smaller \( N_s \) improves performance in the underresolved regime without substantially increasing the rms error for resolved distributions. Figure 10 shows that all three estimators achieve the convergence rate of \( N_s^{-1/2} \) predicted by Bickel & Ritov (1988) and Giné & Nickl (2008) 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 11). The nearest-neighbor estimator with constant Riemann volume (red circles in Figure 11) converges slightly faster than the other two but once the distribution is resolved they are nearly indistinguishable from one another (Figure 11, inset). The FiEstAS method, thanks to the space-filling tree organizing the particles, is faster than the nearest-neighbor estimators, so if the distribution is known to be completely resolved (the regime shown in the inset of Figure 11), then this method can be used without loss of performance to take advantage of its greater speed. However, in situations where
Figure 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. (A color version of this figure is available in the online journal.)

parts of the distribution may be underresolved 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.

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 3.1) and present derivations of formulae for the boost factor and gamma-ray flux (Section 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 to correct our numerical estimates of the boost factor and flux (Section 3.3), and present the results (Sections 3.4 and 3.6).

3.1. N-body Model

For this work we use the $N$-body model of the tidal shell system in Fardal et al. (2007). The model uses a Plummer sphere as the progenitor of the tidal debris, orbiting in a static, three-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 Geehan et al. 2006 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 (Fardal et al. 2006). Finally, the mass and size of the progenitor were constrained with an $N$-body model of the stream using the previously determined orbit (Fardal et al. 2007). 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. Fardal et al. (2007) 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.

Figure 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). (A color version of this figure is available in the online journal.)

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.

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.

3.2.1. 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:

$$\Gamma = \int n_s^2 dV = \int (n_h + n_s)^2 dV = \int (n_h^2 + 2n_h n_s + n_s^2) dV,$$

$$\equiv \Gamma_{hh} + \Gamma_{sh} + \Gamma_{ss}. \quad (18)$$

Equation (18) shows that the boost factor $\beta$ depends on the halo dark matter density:

$$\beta \equiv \frac{\Gamma_{tot} - \Gamma_{hh}}{\Gamma_{hh}} = \frac{1}{\Gamma_{hh}} \int n_s (2n_h + n_s) dV. \quad (19)$$

If $n_h \gtrsim n_s$, the additional emissivity from the tidal debris is dominated by the first term in Equation (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 (Geehan et al. 2006)
with the addition of a small core with size $r_{\text{core}}$ 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 (Navarro et al. 1996, 1997):

$$n_h(r) = \frac{n_{h,0}}{[(r + r_{\text{core}})/r_h][1 + (r + r_{\text{core}})/r_h]^2},$$  \hspace{2cm} (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 Ghecan et al. 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 $n_r$.

### 3.2.2. Gamma-ray Flux

We use the notation of Fornengo et al. (2004, hereafter FPS) to present the calculation of the gamma-ray flux. The differential flux of photons in an infinitesimal band of photon energy $E_\gamma$, $d\Phi_\gamma/dE_\gamma$, can be factored into a contribution from “particle physics” that specifies the spectrum of the radiation and an achromatic contribution $\Phi^{\text{cosmo}}$ from “cosmology”—the shape, size, density, and distance of the dark matter—that sets the normalization, as in Equation (1) of FPS:

$$d\Phi_\gamma/dE_\gamma = d\Phi^{\text{SUSY}}/dE_\gamma \Phi^{\text{cosmo}}.$$  \hspace{2cm} (21)

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

$$dN_\gamma/dt = \Phi^{\text{cosmo}} \int_{E_{\text{th}}}^{m_\chi} d\Phi^{\text{SUSY}}(E_\gamma) A_{\text{eff}}(E_\gamma) dE_\gamma,$$  \hspace{2cm} (22)

where $A_{\text{eff}}(E_\gamma)$ is the effective area of the detector, $E_{\text{th}}$ is the lowest detectable energy, and $m_\chi$ is the dark matter mass. For the Fermi LAT, $E_{\text{th}} = 30 \text{ MeV}$, and above $1 \text{ GeV}$ the effective area for diffuse events is roughly independent of energy to within about 10% of the mean value over the energy range of integration (Rando et al. 2009). Furthermore, the supersymmetric calculations of the particle physics contribution in Gondolo et al. (2004), 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 $A_{\text{eff}}$ is independent of energy. With this simplification we can work with the total flux for now and later multiply it by $A_{\text{eff}}$ to get the detection rate in the LAT.

Also following FPS, the particle-physics contribution is

$$\frac{d\Phi^{\text{SUSY}}}{dE_\gamma} \approx \langle \sigma v \rangle \frac{n_\gamma}{2m_\gamma^2} dE_\gamma,$$  \hspace{2cm} (23)

where $\langle \sigma v \rangle$ is the velocity-averaged cross section and $N_\gamma$ is the yield, called $Q_\gamma$ in Peirani et al. (2004) and FPS. We have moved the constant factor $1/4\pi r$ to be part of $\Phi^{\text{cosmo}}$ because it is most easily understood as part of the attenuation of the gamma-ray flux over distance.

To evaluate $N_\gamma$, we use a subset of the benchmark models of Battaglia et al. (2004) to span the space of supersymmetric WIMP candidates. Gondolo et al. (2004) have calculated the gamma-ray yields above $1 \text{ GeV}$ for 10 of the 12 models in Battaglia et al. (2004): 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, $N_{\gamma,\text{cont}}(\sigma v)$, listed in Table 1 of Gondolo et al. (2004). 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 (22) and (23), we find that

$$\Phi^{\text{SUSY}} = \int_{E_\gamma}^{m_\chi} \frac{d\Phi^{\text{SUSY}}}{dE_\gamma} dE_\gamma = \frac{N_{\gamma,\text{cont}}(\langle \sigma v \rangle)}{2m_\gamma^2}.$$  \hspace{2cm} (24)

For the models in Gondolo et al. (2004), $N_{\gamma,\text{cont}}(\langle \sigma v \rangle)$ is in the range $10^{-28}$ to $10^{-24} \text{ cm}^3 \text{s}^{-1}$, and $m_\chi$ is generally given in GeV. So a useful scaling of this formula in typical units is

$$\Phi^{\text{SUSY}} = 1.54 \times 10^{-8} \text{ cm}^3 \text{kpc}^{-1} \text{s}^{-1} \text{GeV}^{-2} \times \left( \frac{N_{\gamma,\text{cont}}(\langle \sigma v \rangle)}{10^{-29} \text{cm}^3 \text{s}^{-1}} \right) \left( \frac{m_\chi}{1 \text{GeV}} \right)^{-2}.$$  \hspace{2cm} (25)

In addition to the benchmarks, we also use the most optimistic value of $\Phi^{\text{SUSY}}$ from FPS. According to their Figure 8, $\Phi^{\text{SUSY}} \lesssim 10^{-8}$ for all the models explored, and the maximum occurs for $m_\chi \approx 40 \text{ GeV}$ or $N_{\gamma,\text{cont}}(\langle \sigma v \rangle) \approx 10^{-25} \text{ cm}^3 \text{s}^{-1}$. We use these values to represent the most optimistic estimate of the flux.

The astrophysical factor $\Phi^{\text{cosmo}}$ depends on the square of the dark matter mass density $\rho_\chi$:

$$\Phi^{\text{cosmo}} = \frac{1}{4\pi d^2} \int dV \rho_\chi^2(x, y, z),$$  \hspace{2cm} (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 $\int dV \rho^2 dV$. We can calculate this quantity from the simulation results, which consist of the locations of the $N_\text{p}$ simulation particles, each with mass $m_p$. A numerical density estimator gives the number density $n_\rho$ of the simulation particles as a function of position, which is related to the mass density $\rho_\chi$ by mass conservation:

$$\rho_\chi = m_\chi n_\chi = m_p n_\rho$$  \hspace{2cm} (27)

so that

$$\int \rho_\chi^2 dV = m_p^2 \int n_\rho^2 dV$$  \hspace{2cm} (28)

and

$$\Phi^{\text{cosmo}} = \frac{m_p^2}{4\pi d^2} \int n_\rho^2 dV \equiv \frac{m_p^2}{4\pi d^2} E(\hat{\Gamma}).$$  \hspace{2cm} (29)

It is useful to rewrite expression (29) with the units and values used in the $N$-body representation. For the simulation of Fardal et al. (2007), $m_p = 1.68 \times 10^4 \text{M}_\odot$ and $d = 785 \text{ kpc}$. $E(\hat{\Gamma})$ is calculated in units of $\text{kpc}^{-3}$. So a useful version of Equation (29) for this work is

$$\Phi^{\text{cosmo}} = (1.87 \times 10^{-14} \text{ GeV}^2 \text{ kpc}^{-6}) \times \left( \frac{10^4 \text{M}_\odot}{d} \right)^2 \left( \frac{E(\hat{\Gamma})}{785 \text{ kpc}} \text{ kpc}^{-3} \right)^{-2}.$$  \hspace{2cm} (30)

The total flux of gamma rays for a given model of dark matter and density distribution, scaled to typical values in the problem,
composed of particles with a different... relative to M31’s center, let 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 \( N_{\text{peri}} \) (shading/colors are the same as in Figure 12). The youngest caustic contains the particles that have undergone the largest number of orbits. (A color version of this figure is available in the online journal.)

Figure 13. Looking at the projected phase space of \( r \) and \( v_r \), measured relative to M31’s center, we can slightly increase the perceived width of the caustic and cause \( \sigma \) to be overestimated. In order to use our analytical caustic to estimate \( \sigma \) 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 \( \phi \) by binning the particles in \( r \) and \( \phi \) and finding the \( r \)-bin with the highest number of counts for each slice in \( \phi \). The bins were chosen as small as possible for resolution while still being able to identify the peak \( r \) bin for each \( \phi \). Once the peak \( r \) was obtained as a function of \( \phi \), 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 \( \phi \). 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 \( \phi \) direction to find and subtract \( dr_c(\phi) \). Figure 14 illustrates this process.

Correcting for the asphericity in this manner determines the radius \( r_c(\theta, \phi) \) of the caustic surface in two steps, so that

\[
r_c(\theta, \phi) \equiv r_c(\phi) + dr_c(\phi).
\]

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 (Wand 1996) to choose a starting bin width and fit Equation (C37) 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 (C37) to each density profile. The model has a total of four parameters: the caustic width \( \delta r \), \( x_c \), the initial phase-space density \( f_0 \), and the phase-space curvature \( \kappa \), that are discussed in detail in Appendix C. 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 \( \kappa \) to vary in the fit, but checked the fitted values of \( x_c \), \( \delta r \), 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 \( \delta r \), using Equation (C36) and about 10%–20% of the particles in each shell just around the caustic peak. For reference, the covariance term in Equation (C35) was two orders of magnitude smaller than the other terms.

For each caustic, we determined \( \kappa \) 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 \( \kappa \) using the local gravity at each caustic, using Equation (C31). Both are shown in the table and agree with each other. Keeping the value of \( \kappa \) 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 5.

As expected, the model fits the data well all the way through the peak in both cases (Figure 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 Appendix C 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 \( \delta r \) 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 Appendix C.

To finish calibrating the rate calculation, we compared the characteristic widths \( \delta r \) 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 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 2.5, we find that for \( \log_{10} N_p = 4.25 \), the rms error at \( \log_{10} \sigma = -0.75 \) is 4.0% using \( N_s = 10 \) with the FiEstAS estimator and 3.4% with the uniform estimator. For \( \log_{10} N_p = 3.75 \), the rms error at \( \log_{10} \sigma = -0.75 \) is 5.8%.
Figure 15. Radial density profiles of caustics 1 (left upper panel; green in online version) and 2 (right upper panel; red in online version) can be fit surprisingly well with the functional form in Equation (16) after they have been corrected for asphericity with the process illustrated in Figure 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 5.

(A color version of this figure is available in the online journal.)

Table 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) | $38.4 \pm 0.21\phi + 0.14\phi$ | $26.0 - 1.89 + 0.34\phi^2 + 0.3\phi$ | As defined in Equation (32) |
| $x_c$ (kpc) | $0.228 \pm 0.005$ | $0.113 \pm 0.009$ | From fit to density profile after asphericity corrections |
| $\delta r$, estimated (kpc) | 0.23 | 0.39 | Using Equation (C36) |
| $\delta r$, fitted (kpc) | $0.201 \pm 0.005$ | $0.232 \pm 0.009$ | From fit to density profile |
| $\kappa$, estimated$^a$ | $4.6 \times 10^{-4}$ | $2.4 \times 10^{-4}$ | Estimated using Equation (C31) |
| $\kappa$, fitted$^a$ | $(4.2 \pm 0.2) \times 10^{-4}$ | $(2.31 \pm 0.04) \times 10^{-4}$ | From fit to phase-space profile |
| $f_0$, estimated$^b$ | 0.37 | 0.096 | Estimated using mass conservation |
| $f_0$, fitted$^b$ | $(0.33 \pm 0.01)$ | $(0.086 \pm 0.003)$ | From fit to density profile |

Notes. Error ranges on fitted parameters indicate the 95% confidence interval.

$^a$ Units are kpc (km s$^{-1}$)$^{-2}$.

$^b$ Units are kpc$^{-3}$ (km s$^{-1}$)$^{-1}$ sr$^{-1}$.

using $N_s = 10$ and the uniform estimator. All these errors are dominated by the standard deviation, not the undersampling bias.

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, $\Gamma_{ss}$. 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. $\Gamma_{ss}$, 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 (A7)) to estimate $n_s$, and evaluating Equation (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. $\Gamma_{ss}$ was calculated analytically by integrating Equation (20) over the simulation volume. We find that the boost factor $\beta = 2.4 \times 10^{-3}$ for $\Gamma$ integrated over the entire simulation volume, independent of the particle physics model.

By far the largest contribution to $\Gamma_{hh}$ 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°) 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%.

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 16; the maximum is 2.5%, independent of the particle physics model.

### 3.5. Astrophysical Factor

The “astrophysical factor” is the quantity $\Phi^{\text{cosmo}}$ defined by Equations (26), (29), and (30). For a given model of dark matter, comparing values of $\Phi^{\text{cosmo}}$ gives the relative strength of different mass distributions as sources of high-energy particles via self-annihilation. As described in Section 3.4, we used $\Gamma_{\gamma}$ to calculate $\Phi^{\text{cosmo}}$ 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 6). For comparison, we also calculated $\Phi^{\text{cosmo}}$ for the dwarf galaxy used in the N-body model. The dwarf is represented as a Plummer sphere with mass $M_p = 2.2 \times 10^6 M_\odot$ and scale radius $b = 1.03$ kpc. Integrating the density-squared of the dwarf over volume shows that

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

(A color version of this figure is available in the online journal.)

### Table 6

| Object                  | $\Phi^{\text{cosmo}}$ (GeV$^2$ kpc cm$^{-6}$) |
|-------------------------|---------------------------------------------|
| Shell shell (using $\Gamma_s$) | $9.9 \times 10^{-9}$                      |
| Shell halo (using $\Gamma_{sh}$) | $8.5 \times 10^{-7}$                      |
| Plummer dwarf (using $\Gamma_{\text{dwarf}}$) | $4.1 \times 10^{-5}$                      |
| “Ursa Minor” (based on Strigari et al. 2007) | $6.9 \times 10^{-7}$                      |

### Table 7

| Quantity, $n$ | $A'$ | $B'$ | $C'$ | $D'$ | $E'$ | $F'$ | $G'$ | $H'$ | $I'$ | $J'$ | $K'$ | $L'$ | Upper Limit |
|--------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|------------|
| $n_{\gamma}$, GeV$^2$ | 242.8 | 94.9 | 158.1 | 212.4 | 148.0 | 388.4 | 138.1 | 309.1 | 554.2 | 181.0 | 40  |
| $N_\gamma(\sigma v)$ | 120 | 782 | 195 | 63.6 | 86.5 | 6303 | 930 | 7.08 $\times 10^4$ | 1.87 $\times 10^4$ | 1.30 $\times 10^4$ |
| $\Phi^{\text{MSSYb}}$ | 3.14 | 134 | 12.0 | 2.18 | 72.7 | 0.885 | 510 | 15.0 | 356 | 882 | 1.26 $\times 10^4$ |
| $\Phi^{\text{halo}, \text{ss}}$ | 3.38 | 144 | 13.0 | 2.34 | 78.3 | 0.953 | 550 | 16.2 | 383 | 950 | 1.3530 |
| $\Phi^{\text{halo}, \text{sh}}$ | 0.00780 | 0.341 | 0.0307 | 0.00554 | 0.185 | 0.00225 | 1.30 | 0.0382 | 0.906 | 2.25 | 32.0 |
| $\Phi^{\text{halo}, \text{hh}}$ | $<10^{-3}$ | 0.00424 | $<10^{-3}$ | $<10^{-3}$ | 0.003 | $<10^{-3}$ | 0.0101 | $<10^{-3}$ | 0.0112 | 0.0279 | 0.398 |
| $\Phi^{\text{add}, \text{all}}$ | 0.00810 | 0.345 | 0.0310 | 0.00561 | 0.187 | 0.00228 | 1.32 | 0.0387 | 0.917 | 2.28 | 32.4 |
| $\Phi^{\text{halo}, \text{all}}$ | 3.39 | 145 | 13.0 | 2.35 | 78.5 | 0.955 | 551 | 16.2 | 384 | 953 | 13560 |
| $\Phi^{\text{halo}, \text{bh}}$ | 1.92 | 81.8 | 7.35 | 1.33 | 44.4 | 0.5402 | 311 | 9.17 | 217 | 539 | 7672 |
| $\Phi^{\text{halo}, \text{sh}}$ | 0.00509 | 0.217 | 0.0195 | 0.00353 | 0.118 | 0.00144 | 0.827 | 0.0244 | 0.577 | 1.43 | 20.4 |
| $\Phi^{\text{halo}, \text{hh}}$ | $<10^{-3}$ | 0.00362 | $<10^{-3}$ | $<10^{-3}$ | 0.00996 | $<10^{-3}$ | 0.0138 | $<10^{-3}$ | 0.00961 | 0.0238 | 0.339 |
| $\Phi^{\text{add}, \text{all}}$ | 0.00518 | 0.221 | 0.0198 | 0.00359 | 0.120 | 0.00146 | 0.841 | 0.0248 | 0.586 | 1.46 | 20.7 |
| $\Phi^{\text{add}, \text{all}}$ | 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$^{-29}$ cm$^3$ s$^{-1}$. $\Phi^{\text{MSSYb}}$ has units 10$^{-11}$ cm$^3$ kpc$^{-1}$ s$^{-1}$ GeV$^{-2}$. All $\Phi^{\text{g}}$ have units 10$^{-11}$ $\gamma$ cm$^{-2}$ s$^{-1}$. For reference, the Fermi point source sensitivity for photons with $E > 100$ MeV is on the order of 10$^{-9}$ $\gamma$ cm$^{-2}$ 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 $n_{\text{shell}}$ in each integration volume element, then evaluating $n_{\text{halo}}$ 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.
- $^g$ $\Phi^{\text{g}, \text{add}} \equiv \Phi^{\text{g}, \text{hh}} + \Phi^{\text{g}, \text{sh}}$.
- $^h$ $\Phi^{\text{g}, \text{total}} \equiv \Phi^{\text{g}, \text{bh}} + \Phi^{\text{g}, \text{sh}} + \Phi^{\text{g}, \text{ss}}$.
- $^i$ nc: a central region is excluded from the calculation, as described in the text.
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 (Collins et al. 2010). 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 10 times that estimated for Ursa Minor and its analogs in M31).

3.6. Gamma-ray Signal

Beyond calculating $\Gamma$ and $\beta$, we used Equation (31) to estimate the flux of gamma rays in the Fermi LAT for the various benchmarks in Gondolo et al. (2004). The results are shown in Table 7 along with the parameters used to calculate $\Phi^{SUSY}$ 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 7. With this technique we made two-dimensional maps of the expected gamma-ray emission using the most optimistic value of $\Phi^{SUSY}$ (labeled as "Upper Limit" in Table 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 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 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 (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. 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 7). The total signal is comparable to that predicted for an ultra-faint satellite of M31.

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 density (the "rate") from an $N$-body realization is to use the simple nearest-neighbor 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 2.1. 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 $N_\ell > 10$. The correct calibration of the estimator for Poisson bias can change the estimated result by up to 10% for reasonable values of $N_\ell$. The correct calibration for undersampling bias can change the result by a factor of two or more if the simulation is underresolved; 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% 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 7). The total signal is comparable to that predicted for an ultra-faint satellite of M31.

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 orientations are a good limit on the projected angular momentum 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 orientations are a good limit on the projected angular momentum 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 orientations are a good limit on the projected angular momentum 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 orientations are a good limit on the projected angular momentum space according to its total energy.

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 C, the resulting form is identical to that obtained by Mohayaee & Shandarin (2006) 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 (Martínez-Delgado et al. 2010), this technique may provide a way to constrain the properties of luminous matter in dwarf galaxies by examining the tidal debris they produce, as will be discussed in an upcoming paper (R. E. Sanderson 2011, in preparation).

Although the M31 tidal debris is probably not a candidate for indirect detection, we hope that our discussion of how to estimate such signals from N-body realizations will improve those estimates in future work. We also hope it may inspire attempts to develop optimized estimators for this quantity, similar to the way that optimized density estimators have been developed, since no optimization other than simple bias correction was applied to the estimators used in this work, and connect the astrophysics community with the body of statistical literature on such estimators. N-body modeling will undoubtedly prove an indispensable component of the prediction and interpretation of direct and indirect detections of dark matter.

The authors acknowledge support from NASA grant NNG06GG99G. The numerical experiments in the paper were performed using the MIT Kavli Institute computing cluster, which is supported in part by the Kavli Foundation. The authors thank Paul Hsi for maintaining, troubleshooting, and upgrading the cluster. R.E.S. thanks Will Farr for the use of his N-body integrator and for many helpful conversations.

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 \( n(x) \) 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 \( N_{th} \) 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

\[
\hat{n}_{b} = \frac{3N_s}{4\pi r^2_{N_s}} \tag{A1}
\]

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

\[
E(\hat{n}_b) = \frac{N_s}{N_s - 1} n \tag{A2}
\]

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 (A1) 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

\[
\hat{n}_u = \frac{3(N_s - 1)}{4\pi r^3_{N_s}} \tag{A3}
\]

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

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 (A3), is not equal to \( n^2 \); still, an unbiased estimator for \( n^2 \) in the case of the uniform distribution does exist:

\[
\hat{n}^2_u = (N_s - 1)(N_s - 2) \left( \frac{3}{4\pi r^3_{N_s}} \right)^2 \tag{A4}
\]

with \( N_s \) and \( r_{N_s} \) defined as before. Then \( E(\hat{n}_u^2) = n^2 \).

Using \( n^2 \), we can construct an unbiased estimator for \( \Gamma \) by using a Riemann sum over \( N_s \) identical volumes \( dV \) to approximate the volume integral, so that

\[
\hat{\Gamma}_u = dV \sum_{i=1}^{N_v} \hat{n}^2_{u,i} \tag{A5}
\]
where $r_{n,i}^2$ is given by evaluating Equation (A4) at the center of subvolume $i$, and the total volume $V = N_v dV$. Using Equation (A5), $E(\hat{\Gamma}_n) = \Gamma_{\text{true}}$. 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 (A4) is used to calculate the density squared. For simplicity, we represent the $N_s$ dependence of the additional bias from determining the density-squared and the Poisson bias from using ($\hat{n}_s(r_i)$)$^2$ to estimate the density in the case of a uniform distribution. We use this density estimator and an adaptive Riemann volume for the rate estimator $n$, and again collect the $N_s$-dependence of the bias in a prefactor:

$$\hat{\Gamma}_n = \frac{1}{1 + b_n(N_s)} \sum_{i=1}^{N_v} (\hat{n}_s(r_i))^2 dV_i. \quad \text{(A7)}$$

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 $(\hat{n}_s(r_i))^2$ 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

$$\hat{\Gamma}_f \equiv \frac{1}{1 + b_f(N_s)} \sum_{i=1}^{N_v} (\hat{n}_f(r_i))^2 dV_i. \quad \text{(A8)}$$

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 $\hat{x}$ is

$$\hat{n}_s(\hat{x}) = \sum_{j=1}^{N_s} W(\hat{x}_j - \hat{x}, \hat{h}). \quad \text{(A9)}$$

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

The type of kernel chosen can have a significant effect on the bias and variance. Sharma & Steinmetz (2006) 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

$$W(\hat{r}, \hat{h}) = \frac{f W(u)}{V_h}, \quad \text{(A10)}$$

where $\hat{r}$ is the distance from the target location, $u$ is the scaled distance $u = r/h$, $f$ is the kernel normalization

$$\frac{1}{f} = \int_0^1 W(u)4\pi u^2 du, \quad \text{(A12)}$$

and $V_h$ is the volume enclosed by the smoothing length $\hat{h}$. Sharma & Steinmetz found that the Epanechnikov kernel 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 $n$, and again collect the $N_s$-dependence of the bias in a prefactor:

$$\hat{\Gamma}_s \equiv \frac{1}{1 + b_s(N_s)} \sum_{i=1}^{N_v} (\hat{n}_s(r_i))^2 dV_i. \quad \text{(A14)}$$

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

Diemand et al. (2007) 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 (A14), 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 Diemand et al.:

$$\hat{\Gamma}_d \equiv \frac{1}{1 + b_d(N_s)} \sum_{i=1}^{N_v} \hat{n}_s(r_i). \quad \text{(A15)}$$

From a Poisson-statistics standpoint, the substitution implicitly assumes that $E(\hat{n})^2 = E(n^2)$, 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 (A14).

APPENDIX B

CALCULATION OF THE MINIMUM NEAREST-NEXTBORDER DISTANCE

In this appendix, we derive an expression for the expectation value of the minimum nearest-neighbor distance, $E(\hat{r}_{N_n \text{, min}})$, 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. $E(\hat{r}_{N_s \text{, min}})$ is the first order statistic of the estimator $\hat{r}_{N_s}$, which is related to the nearest-neighbor density estimator $\hat{n}$ by

$$\hat{r}_{N_s} = \left(\frac{4\pi \hat{n}}{3N_s}\right)^{-1/3}. \quad \text{(B1)}$$

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

$$p_{r_{N_s}}(\rho)d\rho = \frac{\exp(-4\pi n \rho^3/3)}{(N_s - 1)!} \left(\frac{4\pi n \rho^3}{3}\right)^{N_s - 1} d\rho \left(\frac{4\pi n \rho^3}{3}\right)^{N_s - 1}, \quad \text{(B2)}$$
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,

$$P_{r_{N_s}}(\mu) = \frac{1}{(N_s - 1)!} [(N_s - (N_s, 4\pi n \mu^3/3)]$$

$$= 1 - \frac{(N_s, 4\pi n \mu^3/3)}{(N_s)}, \quad (B3)$$

where ($N$) and ($N, x$) 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:

$$P_{\hat{r}_{N_s, \min}}(v)dv = N_p [1 - P_{r_s}(v)]^{N_s - 1} P_{r_s}(v)dv. \quad (B4)$$

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

$$P_{\hat{r}_{N_s, \min}}(\mu)d\mu = \frac{N_p}{(N_s - 1)!} \exp(-4\pi n \mu^3/3) \left(\frac{4\pi n \mu^3}{3}\right)^{N_s - 1}$$

$$\times \left(\frac{(N_s, 4\pi n \mu^3/3)}{(N_s)}\right)^{N_s - 1} d\left(\frac{4\pi n \mu^3}{3}\right). \quad (B5)$$

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

$$P_{\hat{r}_{N_s, \min}}(y)dy = \frac{N_p}{(N_s - 1)!} e^{-y} y^{N_s - 1} \left(\frac{(N_s, y)}{(N_s)}\right)^{N_s - 1} dy. \quad (B6)$$

y represents the average number of particles in a sphere of radius $\mu$.

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

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

$$\times \int_0^\infty e^{-y} y^{N_s - 2/3} \left(\frac{(N_s, y)}{(N_s)}\right)^{N_s - 1} dy, \quad (B7)$$

using the definition of $\mu$ 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 $y$:

$$(N_s, y) \approx e^{-y} [y^{N_s - 1} + O(y^{N_s - 2})]; \quad (B8)$$

the second is Stirling’s approximation for the complete gamma function at large argument $N$:

$$(N + 1) = N! \approx N^N e^{-N}. \quad (B9)$$

Using the asymptotic expansion, we may rewrite the integral:

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

$$\times \int_0^\infty e^{-N_s y} y^{N_s(N_s - 1) + 1/3} dy. \quad (B10)$$

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

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

$$\times \int_0^\infty e^{-t} N_p y^{N_s(N_s - 1) + 1/3} dt \quad (B11)$$

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

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

$$= \left(\frac{3}{4\pi n}\right)^{1/3} \frac{[N_p(N_s - 1) + 1/3]}{[(N_s - 1)!]^{N_s}} N_p^{N_s(N_s - 1) + 1/3}. \quad (B12)$$

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

$$[(N_s - 1)!]^{N_s} \approx [(N_a + 1/3)!]^{N_s} e^{-(N_s + 1/3)} \quad (B13)$$

$$[(N_s - 1)!]^{N_s} \approx [(N_a - 1)!]^{N_s - 1} e^{-(N_s - 1)} N_s = (N_a - 1) e^{-N_a}. \quad (B14)$$

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

$$E(\hat{r}_{N_s, \min}) = \left(\frac{3}{4\pi n}\right)^{1/3} \frac{(N_a + 1/3)!^{N_s} e^{-(N_s + 1/3)}}{N_p^{N_s + 1/3}(N_s - 1)^{N_s} e^{-N_a}}$$

$$= \left(\frac{3(N_a - 1)}{4\pi n e}\right)^{1/3} \left(\frac{N_s + 1/3}{N_a}\right)^{N_s + 1/3}. \quad (B15)$$

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

$$E(\hat{r}_{N_s, \min}) \approx \left(\frac{3(N_s - 1)}{4\pi n}\right)^{1/3}. \quad (B16)$$

The minimum expected distance to the $N_s^{th}$ particle is roughly equal to the average expected distance to the $(N_s - 1)$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 (B7) 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 $\gamma$ depends slightly on $N_p$. We confirm that $E(\hat{r}_{N_s, \min})$ scales with $N_p$ as predicted (Figure 18).

**APPENDIX C**

**AN ANALYTIC ONE-DIMENSIONAL CAUSTIC**

We would like to test the bias and variance of three common density estimation schemes in the context of high density contrast. The three-dimensional test distribution has a one-dimensional caustic in $(x)$ and is uniform in the other two...
dimensions ($y$ and $z$). Below we present the derivation of Equation (16) and discuss the calculation of $\Gamma$.

C.1. Density Profile

Following the notation of Shandarin & Zeldovich (1989), 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:

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

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

Conservation of mass states that

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

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 $\rho_0$, the local density $\rho$ of the distribution then evolves in time as

$$\rho(\vec{x}, t) = \sum_{\vec{q} \rightarrow \vec{x}} \rho_0 \left| \frac{d\vec{x}}{d\vec{q}} \right| = \sum_{\vec{q} \rightarrow \vec{x}} \rho_0 \left| \frac{\delta_{ik} + \gamma_{\rho\rho}}{\delta_{\rho\rho}} \right|.$$

(C3)

Caustics arise when the determinant in the denominator is zero.

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

$$y = q_y,$$

(C4)

$$y = q_y,$$

(C5)

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:

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

(C6)

$v_{\text{th}}(q_x)$ 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 $\Gamma$ will not diverge at the caustic surface. Equation (C6) is not invertible since the $v_{\text{th}}$ 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):

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

(C7)

where $v(q) \equiv v(q_0) + v_{\text{th}}(q)$ 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 $\sigma$:

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

(C8)

$v_{\text{th}}(q)$ 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 $\sigma$ could easily be time dependent, but again we take the simplest case and assume it is a constant. We refer to the $v_{\text{th}}$ 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:

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

(C9)

Combining, we find

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

(C10)

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

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

(C11)
We reintroduce \(v(q) \equiv v_0(q) + v_{th}(q)\) and pull out the parts of the exponent containing \(v_{th}\):

\[
\rho(x, t) = \rho_0 \int_{-\infty}^{\infty} dq \lim_{\epsilon \to 0} \frac{1}{\sqrt{2\pi \epsilon^2}} e^{-\frac{(x-q-v_0(q)t)^2}{2\epsilon^2}} \\
\times \int_{-\infty}^{\infty} dv_{th} \exp \left\{ - \frac{v_{th}^2}{2\sigma^2} + \frac{t^2}{2\epsilon^2} - \frac{t \Delta}{\epsilon^2} v_{th} \right\}.
\]

(C12)

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:

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

(C13)

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

\[
a^2 \equiv \left( \frac{1}{2\sigma^2} + \frac{t^2}{2\epsilon^2} \right)
\]

we find that the inner integral becomes

\[
I = \int_{-\infty}^{\infty} dv_{th} \exp \left[ -a^2 \left( v_{th} - \frac{t \Delta}{2\epsilon^2} \right)^2 + \frac{t^2 \Delta^2}{4\epsilon^4 a^2} \right].
\]

(C15)

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

\[
I(x, t) = \exp \left( \frac{t^2 \Delta^2}{4\epsilon^2 a^2} \right) \sqrt{\frac{\pi}{a^2}}.
\]

(C16)

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

\[
\rho(x, t) = \rho_0 \int_{-\infty}^{\infty} dq \lim_{\epsilon \to 0} \frac{1}{\sqrt{2\pi \sigma^2 \epsilon^2}} e^{-\Delta^2/2\epsilon^2} \\
\times \exp \left[ -\Delta^2 \left( \frac{1}{2\epsilon^2} - \frac{t^2}{4\epsilon^4 a^2} \right) \right].
\]

(C17)

Replacing \(a^2\) with its definition and simplifying makes the limit easy to take

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

(C18)

and finally we are down to the last integral:

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

(C19)

As \(\sigma\) 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 \(v_0(q) = -\alpha q^2\):

\[
\rho(x, t) = \rho_0 \frac{1}{\sqrt{2\pi \sigma^2 t^2}} \int_{-\infty}^{\infty} dq e^{-\frac{(x-q+\alpha q^2 t^2)}{2\sigma^2 t^2}}.
\]

(C20)

For this form of \(v_0\) we can evaluate the integral

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

(C21)

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

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

(C22)

Then we make the substitution

\[
u \equiv q - \frac{1}{2\alpha t}.
\]

(C23)

This change of variables is certainly suspect as \(t \to 0\), 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 \(\rho_0\) everywhere.

With this substitution

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

(C24)

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 (C3) to zero and solving for \(x_c\), with the help of Equation (C1). 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

\[
\frac{x_c}{\alpha t} - \frac{1}{4\alpha^2 t^2} = 0
\]

(C25)

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

\[
\frac{x}{\alpha t} - \frac{1}{4\alpha^2 t^2} = \frac{1}{\alpha t} (x - x_c) \equiv \tilde{x}
\]

(C26)

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

\[
I(x, t) = \int_{-\infty}^{\infty} du \exp \left[ -\frac{a^2}{2\sigma^2} (\tilde{x} + u^2)^2 \right].
\]

(C27)

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

\[
I(\tilde{x}, t) = e^{-\alpha^2 \tilde{x}^2/2\sigma^2} \int_{-\infty}^{\infty} du \exp \left( -\frac{a^2}{2\sigma^2} u^4 - \frac{\tilde{x}^2 \alpha^2}{\sigma^2} u^2 \right).
\]

(C28)

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

\[
\rho(x, t) = \frac{\rho_0}{\sqrt{2\pi \sigma^2 t^2}} \int_{-\infty}^{\infty} dq e^{-\frac{(x-x_c)^2}{2\sigma^2 t^2}} B \left( \frac{(x-x_c)^2}{4\sigma^2 t^2} \right)
\]

(C29)

with

\[
B(u) = \begin{cases} 
\frac{\pi}{2} [I_{-1/4}(u) + I_{1/4}(u)] & x \leq x_c \\
\frac{\pi}{2} [I_{-1/4}(u) - I_{1/4}(u)] & x > x_c,
\end{cases}
\]

(C30)
where $I$ is a modified Bessel function of the first kind. Taking the limit $t \to 0$ recovers $\rho_0$ everywhere and validates our change of variables. This form is the same as that obtained by Mohayaee & Shandarin (2006) for the density profile of a caustic with Gaussian velocity dispersion, with the substitutions $\sigma_0, \sigma_1 \to \sigma t$, $\Delta x \to x - x_1$, and $A_0 = \rho_0/\sqrt{4t}$. 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. In fact, this density profile is, with a few slight changes, universal for any initial population with a Gaussian velocity dispersion, 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 $\kappa$. For the equation of motion used to derive Equation (C29), $\kappa = \alpha t^3$, for motion in a gravitational potential, $\kappa$ depends on the radial gravitational force at the caustic:

$$\kappa = \frac{1}{\frac{1}{2} \frac{d^2 \rho}{d v_r^2}} = -\frac{1}{2g(r_c)},$$

(C31)

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

Inspection of Equation (C29) shows that the curvature $\kappa$ influences the height of the caustic relative to the initial density, while the product $\sigma t$ 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 $\sigma t$ can be replaced everywhere in the density profile by a generic parameter representing the width of the caustic, $\delta r$. The quantity $\rho_0/\sigma$ 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, $\sigma$ remains the one-dimensional velocity dispersion while $\rho_0$ 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 $\delta r$ can be estimated using energy conservation. Although each particle turns around at apocenter where $\sigma_0, \sigma_1 \to \sigma t$, $\Delta x \to x - x_1$, and $A_0 = \rho_0/\sqrt{4t}$. 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. In fact, this density profile is, with a few slight changes, universal for any initial population with a Gaussian velocity dispersion, 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 $\kappa$. For the equation of motion used to derive Equation (C29), $\kappa = \alpha t^3$, for motion in a gravitational potential, $\kappa$ depends on the radial gravitational force at the caustic:

$$\kappa = \frac{1}{\frac{1}{2} \frac{d^2 \rho}{d v_r^2}} = -\frac{1}{2g(r_c)},$$

(C31)

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

Inspecting Equation (C29) shows that the curvature $\kappa$ influences the height of the caustic relative to the initial density, while the product $\sigma t$ 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 $\sigma t$ can be replaced everywhere in the density profile by a generic parameter representing the width of the caustic, $\delta r$. The quantity $\rho_0/\sigma$ 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, $\sigma$ remains the one-dimensional velocity dispersion while $\rho_0$ 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 $\delta r$ 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

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

(C32)

For two particles on neighboring orbits, their energy difference $\Delta E$ is found by subtracting their energies, so

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

(C33)

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 $\Delta v/v_1 << 1$, 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):

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

(C34)

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 $g(r)/\Delta r \gg dg/dr$:

$$\sigma_E^2 = \frac{1}{2} \sigma_v^2 - g(r)\sigma_v^2 + g(r)^2 \sigma_r^2.$$  

(C35)

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 $\delta v \ll \delta r/\sqrt{\kappa}$. Then the covariance term can be ignored and the equation solved to give an estimate of the width $\sigma_r$:

$$\sigma_r = \sqrt{\frac{\sigma_E^2 - \frac{1}{2} \sigma_v^2}{g(r_c)}}.$$  

(C36)

With these physically motivated generalizations, the density profile

$$\rho(x, t) = \frac{f_0}{\sqrt{2\pi}} \frac{1}{\kappa} e^{-(x-x_c)^2/4\delta r^2} B \frac{(x-x_c)^2}{4\delta r^2}$$

(C37)

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 $dp/dr = 0$. The equation reduces to a linear combination of various Bessel functions, which can be solved numerically to find $r_{\text{max}}$. 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,

$$r_{\text{max}} = r_c - 0.7649508674\delta r.$$  

(C38)

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

$$\rho_{\text{max}} = 1.0211365847 f_0 \sqrt{\frac{\delta r}{\kappa}}.$$  

(C39)

The peak density occurs about one characteristic width behind the nominal caustic surface $r_c$ (Figure 19).

C.2. Rate

Armed with an analytic expression for the density, we can now calculate the rate $\Gamma$ for comparison with estimates from random realizations. We must integrate the density with finite limits on $q$ and $x$, since we will use the density estimator on a finite-sized sample of the distribution. Since the caustic is only in the $x$ dimension,

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

(C40)

with $x_\pm$ being the lower and upper edges of the box in the $x$-direction, and $\pm L_x$ and $\pm L_z$ the box 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 (C27), over a finite range:

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

(C41)
The exponential and Bessel function terms are functions of the quantity in the form of Equation (C29). The exponential and Bessel function instead if possible. We can see that this is sufficient by examining the endpoints of the rate integral about the caustic.

$$\Gamma(t) = \int_{\bar{x}_-}^{\bar{x}_+} e^{-u^2} \left[ I_{-1/4}(u) + I_{1/4}(u) \right] du,$$

where

$$u_{\pm} = q_{\pm} - \frac{1}{2\sigma t}.$$  \hspace{1cm} (C42)

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

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

with

$$\bar{x}_\pm = \frac{x_{\pm} - x_c}{\sigma t}.$$  \hspace{1cm} (C44)

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 (C29), instead if possible. We can see that this is sufficient by examining the form of Equation (C29). The exponential and Bessel function terms are functions of the quantity

$$\frac{(x - x_c)}{2\sigma t} = \bar{x}_d$$  \hspace{1cm} (C45)

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

$$K_{1/4}(u) = \frac{\sqrt{\pi}}{\sqrt{u}} e^{-u^2} \left[ \frac{1}{u} - \frac{3}{32 u^3} + O \left( \frac{1}{u^5} \right) \right].$$  \hspace{1cm} (C46)

The sum of modified Bessel functions $I_{\pm 1/4}$ has an exponentially increasing term, shown in its asymptotic series:

$$I_{-1/4}(u) + I_{1/4}(u) = e^{u^2} \sqrt{\frac{\pi}{2}} \left[ \frac{1}{u} + \frac{3}{32 u^3} + O \left( \frac{1}{u^5} \right) \right] + e^{-u^2} i \sqrt{\pi} \left[ \frac{1}{u} - \frac{3}{32 u^3} + O \left( \frac{1}{u^5} \right) \right].$$  \hspace{1cm} (C47)

that is canceled by the exponential piece of the equation, leaving power-law convergence to zero for $(x - x_c) \gg \sigma t$ in the asymptotic series:

$$e^{-u^2} (I_{-1/4}(u) + I_{1/4}(u)) = \sqrt{\frac{2}{\pi}} \left[ \frac{1}{u} + \frac{3}{32 u^3} + O \left( \frac{1}{u^5} \right) \right] + e^{-u^2} i \sqrt{\pi} \left( \frac{1}{u} \right).$$  \hspace{1cm} (C48)

From examining these limits, we find that for $x > x_c$ the density is completely negligible for distances more than a few times $\sigma t$ 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

$$\Gamma(t) = 4L_y L_z \rho_0^2 \frac{1}{2\pi \sigma^2 t^2} \int_{\bar{x}_-}^{\bar{x}_+} \left( x - x_c \right) e^{-2(x - x_c)^2/4\sigma^2 t^2} B^2 \left[ \left( x - x_c \right) e^{-2(x - x_c)^2/4\sigma^2 t^2} \right] dx.$$  \hspace{1cm} (C49)

Practically speaking, the approximation is sufficient for our needs as long as $x > 500$ at the endpoints. If this condition is not satisfied we must do the two numerical integrals in Equation (C43).

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

$$\Gamma(t) = \frac{4L_y L_z \rho_0^2}{t} \left[ \int_{\bar{x}_d}^{\bar{x}_d^+} |\bar{x}_d| e^{-2\bar{x}_d^2} \left[ I_{-1/4} \left( \bar{x}_d \right) + I_{1/4} \left( \bar{x}_d \right) \right]^2 d\bar{x}_d \right. + \left. \int_{\bar{x}_d^-}^{\bar{x}_d} |\bar{x}_d| e^{-2\bar{x}_d^2} \left[ I_{-1/4} \left( \bar{x}_d \right) - I_{1/4} \left( \bar{x}_d \right) \right]^2 d\bar{x}_d \right],$$  \hspace{1cm} (C50)

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

$$\bar{x}_d \equiv \frac{x_{\pm} - x_c}{2\sigma t},$$  \hspace{1cm} (C51)

are such that $\bar{x}_d^- < 0$ and $\bar{x}_d^+ > 0$. This integral must be performed numerically, but it is a single integral rather than a two-step process.

REFERENCES

Abdo, A. A., et al. (Fermi–LAT Collaboration) 2009a, Phys. Rev. Lett., 102, 181101
Abdo, A. A., et al. (Fermi–LAT Collaboration) 2009b, ApJS, 183, 46
Abdo, A. A., et al. (Fermi–LAT Collaboration) 2010, ApJ, 712, 147
Adriani, O., et al. 2009a, Nature, 458, 607
Adriani, O., et al. 2009b, Phys. Rev. Lett., 102, 051101
Afshordi, N., Mohayaee, R., & Bertschinger, E. 2009, Phys. Rev. D, 79, 083526
Ahmed, Z., et al. 2009b, Phys. Rev. Lett., 102, 011101
Ahmed, Z., et al. 2009b, Phys. Rev. Lett., 103, 141802
Angle, J., et al. 2009, Phys. Rev. D, 80, 115005
Ascasibar, Y., & Binney, J. 2005, MNRAS, 356, 872
Asztalos, S. J., et al. 2004, Phys. Rev. D, 69, 011101
Baltz, E. A., Battaglia, M., Peskin, M. E., & Wizansky, T. 2008, Phys. Rev. D, 74, 103521
Battaglia, M., et al. 2004, Eur. Phys. J. C, 33, 273
Bergström, L., Ullio, P., & Backley, J. H. 1998, Astropart. Phys., 9, 137
Bertone, G., Hooper, D., & Silk, J. 2005, Phys. Rep., 405, 279
