Measuring the Mass of a Galaxy: An evaluation of the performance of Bayesian mass estimates using statistical simulation

Gwendolyn Eadie*

Abstract

A common way to estimate the mass of a galaxy is to observe the kinematics of objects orbiting it, and then infer the mass by assuming a model for the gravitational potential and mass distribution. We use a simulation approach to study biases that may occur when kinematic data are used in this way. Data sets are simulated from two different velocity distributions and then each set is analysed, either with complete velocity vectors (complete data) or without complete velocity vectors (incomplete data), using a Bayesian isotropic Hernquist (1990) model. We investigate three scenarios: (1) the model and data come from the same probability distribution function (PDF), (2) the model and data come from the same PDF and the data is incomplete, and (3) the model and data come from different PDFs and the data is incomplete. No biases were found in scenarios 1 and 2, but a positive bias was found in scenario 3. The cause of the bias in the latter scenario appears to be caused by the isotropic Hernquist model incorrectly interpolating for the incomplete data.

Key Words: galaxy - mass profile - Bayesian - incomplete data - astrophysics - dark matter

1. Introduction

Mass estimates and mass profiles of galaxies place important constraints on models and theories in many fields of astronomy. Most notably, masses and mass profiles of galaxies are important for testing both cosmological models and simulations that predict the spatial distribution of dark matter in galaxies. The mass estimates also play a role in developing galactic evolution theories, star formation theories, and mass-luminosity relationships. Measuring the mass and mass profile, however, is not easy.

Two notable ways to measure the mass of a galaxy are to measure 1) rotation curves, or 2) the velocity distribution of satellites orbiting the galaxy. The former is only applicable to spiral galaxies, whereas the latter is applicable to elliptical galaxies. There is another method, however, that uses the positions and velocities of tracers to estimate the mass of the dark matter halo. Tracers orbit the host galaxy, and are objects such as dwarf galaxies, globular clusters (GCs), halo stars, and planetary nebulae. We can measure the distance of these objects, r, from the center of the Milky Way (Galaxy) via standard candles (i.e. objects with known intrinsic luminosities). The velocities of tracers in our line of sight can be measured by observing the Doppler shift of known spectral lines in the object's electromagnetic spectrum. The velocities of the tracers across the plane of the sky (also known as the proper motion) are more difficult to measure, but measurements can be achieved through parallax and/or many years of observation. Accordingly, not all tracers orbiting the Milky Way have proper motion measurements. Both line of sight *and* proper motion measurements are necessary to obtain the complete three-dimensional velocity vector of a tracer and to convert the velocity into the galaxy-centered reference frame for analysis.

We explore the Bayesian method that uses kinematic data $(r, \boldsymbol{v})$ of tracers to estimate the mass and mass profile of a galaxy, which was first introduced by Little and Tremaine (1987). We also propose a new method for estimating the mass of a galaxy with incomplete

*McMaster University, Physics & Astronomy Department, Hamilton, Ontario, Canada, L8S 4M1. Email: eadiegm@mcmaster.ca

data. We search for biases that may occur when an incorrect model is assumed and when the data is incomplete.

In the Bayesian method, one assumes a model for the gravitational potential of a galaxy, and derives a probability distribution function for the satellites' positions and velocities. Bayesian inference is performed to obtain parameter estimates for the model, given the data and prior assumptions. In order to investigate biases, we simulate data sets and perform an analysis as if we were an observer at the center of the galaxy. The simulated data sets represent tracers seen orbiting galaxies in nature, such as dwarf galaxies, globular clusters, planetary nebulae, and halo stars.

For our work, we use the Hernquist (1990) model, which is both analytic and ubiquitous in the literature. Although the model assumes spherical symmetry and no net rotation, and a typical disc galaxy has rotation and a non-spherical shape, the model is still of interest for estimating the large-scale mass distribution of the dark matter halo, which goes far beyond the extent of the visible matter in a galaxy.

2. Models and Notation

We are interested in the cumulative mass profile of the Hernquist model, which is given by

$$M(r) = M_{tot}\frac{r^2}{(r+a)^2}. \tag{1}$$

M_{tot} is the total mass of the system and a is the scale radius, which determines the steepness of the mass profile $M(r)$.

We use a Galactocentric coordinate system, which is a non-rotating, inertial reference frame in the center of the Galaxy. From this point of view, the velocity vector $\boldsymbol{v}$ for a satellite can be split up into two components: 1) the radial velocity, in the line of sight of an observer sitting at the center of the Galaxy, and 2) the tangential velocity, which is the velocity projected onto the plane of the sky. These components are orthogonal to each other and give the total speed of a tracer by

$$v = \sqrt{v_r^2 + v_t^2} \tag{2}$$

where v_r and v_t are the radial and tangential velocity components respectively.

The velocity distribution of satellites orbiting a galaxy may be isotropic or anisotropic, depending on the orbits of the tracers. If there is a bias towards elliptical orbits, then the system is radially anisotropic, and if there is a bias towards circular orbits, then it is tangentially anisotropic. The anisotropy parameter β describes the degree of anisotropy in the system by the equation

$$\beta = 1 - \frac{\sigma_t^2}{2\sigma_r^2} \tag{3}$$

where σ_r^2 and σ_t^2 are the radial and tangetial velocity variances (Binney and Tremaine 2008). The three extremes for β are

1. $\beta \to -\infty$, tangential anisotropy - only circular orbits
2. $\beta = 0$, isotropic velocity distribution
3. $\beta = 1$, perfectly radial anisotropy - plunging orbits

Osipkov (1979) and Merritt (1985) independently introduced an anisotropic radius parameter, r_a, to allow β to vary with r, such that

$$\beta(r) = \frac{r^2}{r^2 + r_a^2}. \tag{4}$$

As $r_a \to \infty$, eq. 4 goes to the isotropic case ($\beta = 0$). Hereafter, the velocity anisotropy described by eq. 4 will be referred to as OM-type.

Assuming a velocity anisotropy of the satellite population, and given the kinematic information of the satellites, it is possible to learn about the gravitational potential of the whole galaxy, and thus its mass. Little and Tremaine (1987) were the first to apply Bayesian inference to this problem, and showed how to obtain the likelihood from a physical model so that it can be used in Bayes' theorem.

In astrophysics, the distribution function (DF) is the probability of a satellite having a particular position $\boldsymbol{r}$ and velocity $\boldsymbol{v}$ in an infinitesimal phase space $d^3\mathbf{r}d^3\mathbf{v}$ (Binney and Tremaine 2008). The DF is a probability density function (PDF) and integrates to one,

$$\int f(\mathbf{r}, \mathbf{v}) d^3\mathbf{r} d^3\mathbf{v} = 1. \tag{5}$$

The DF can be derived from the mass density profile $\rho(r)$ and the gravitational potential $\Phi(r)$ of a model. For example, the Hernquist model has a mass density profile given by

$$\rho(r) = \frac{aM_{tot}}{2\pi r(r+a)^3} \tag{6}$$

and a gravitational potential

$$\Phi(r) = -\frac{M_{tot}}{r+a} \tag{7}$$

where the gravitational constant G has been set to 1. If the mass density profile and the *relative* gravitational potential, $\Psi(r)$, are an analytic potential-density pair, then the mass density can be written in terms of the relative potential, $\rho(\Psi)$, which can then be used in an Abel transform to solve for the DF (see Binney and Tremaine 2008, for more details),

$$f(\mathcal{E}) = \frac{\sqrt{2}}{4\pi^4} \int_0^{\mathcal{E}} \frac{1}{\sqrt{\Psi - \mathcal{E}}} \left(\frac{d\rho}{d\Psi} \right) d\Psi \tag{8}$$

where $\mathcal{E}$ is the relative energy (per unit mass) of the satellite, given by

$$\mathcal{E}(r, v) = -\frac{v^2}{2} + \Psi(r). \tag{9}$$

When many satellites orbit the galaxy, the product of the DFs gives the probability of finding all of the satellites with their positions and velocities (assuming independence),

$$p(y|\boldsymbol{\theta}) = \prod_{i=1}^{n} f(r_i, \boldsymbol{v}_i|\boldsymbol{\theta}). \tag{10}$$

Eq. 10 is the likelihood in the numerator of Bayes' rule:

$$p\left(\boldsymbol{\theta}|y\right) \propto \prod_{i=1}^{n} f(r_i, \boldsymbol{v}_i|\boldsymbol{\theta}) p\left(\boldsymbol{\theta}\right). \tag{11}$$

In the case of the Hernquist model, $\Psi(r) = -\Phi(r)$, and for an isotropic velocity distribution the DF becomes,

$$f(q) = \frac{M_{tot}}{8\sqrt{2}\pi^3 a^3 v_g^3 \left(1-q^2\right)^{5/2}} \left[3\arcsin(q) + q\sqrt{(1-q^2)}\left(1-2q^2\right)\left(8q^4-8q^2-3\right)\right] \tag{12}$$

where q and v_g are given by,

$$q = \sqrt{\frac{a\mathcal{E}}{M_{tot}}}\,; v_g = \sqrt{\frac{M_{tot}}{a}}. \tag{13}$$

Hernquist (1990) also derived the DF for their model when the system has OM-type anisotropy. The DF for a system with an anisotropic velocity distribution is no longer a function of $\mathcal{E}$ alone— it becomes a function of both $\mathcal{E}$ and the specific angular momentum $L = rv_t$. For notational simplicity, the Hernquist OM-type DF is written in terms of a new variable Q that takes into account the specific angular momentum:

$$Q = \mathcal{E} - \frac{L^2}{2r_a^2} \tag{14}$$

The OM-type Hernquist DF is then,

$$f(Q) = f(\tilde{q}) + \frac{M}{\sqrt{2}\left(\pi a v_g\right)^3}\frac{a^2}{r_a^2}\tilde{q}\left(1-2\tilde{q}\right) \tag{15}$$

where $f(\tilde{q})$ refers to equation 12, with q replaced by $\tilde{q}$:

$$\tilde{q} = \sqrt{\frac{aQ}{M}}. \tag{16}$$

As $r_a \to \infty$, $Q \to \mathcal{E}$, and $f(Q) \to f(q)$. In other words, as $r_a \to \infty$, the OM-type DF becomes the isotropic DF given by eq. 12.

3. Methods

3.1 Simulated Data

We use two types of simulated data sets in our exploratory analysis. The data sets consist of tracer particles that have the basic kinematic elements (r, v_r, v_t). All data sets come from the Hernquist (1990) model, but the first type of data have an isotropic velocity distribution (eq. 12) and the second type have an OM-type velocity distribution (eq. 15). The data sets we explore for scientific purposes have 25 tracers each, whereas the data sets used to test code have 100 tracers each. There are 500 independent data sets of each type, so that the Bayesian analysis can be repeated 500 times for each data-model combination. By repeating the analysis many times on the same type of data, it is possible to discover biases that may occur when the model assumed is different from the true model the data follow. In this preliminary study, we do not introduce observational uncertainties. We analyse the following three scenarios, under the assumption of an isotropic Hernquist model:

1. Isotropic data sets of 100 tracers each with complete velocity vectors
2. Isotropic data sets of 25 tracers each with incomplete velocity vectors (no v_t)
3. OM-type data sets of 25 tracers each with incomplete velocity vectors (no v_t)

The first scenario is mainly for testing our method and code. We expect that when the model and the data share the same DF, then on average the posterior probability will return the true parameter values. Furthermore, we expect the credible regions for the $M(r)$ profile will not be over- or underconfident.

The second scenario introduces incomplete data into the analysis, treating all the unknown v_t values as parameters in the model. Because the model and data still share the same DF, we expect that the posterior probabilities will return the correct parameter values on average, but that the variance in the results will be larger. We might also expect that the unknown v_t's can be estimated from the resulting Markov chain.

The third scenario not only has incomplete data, but also assumes an incorrect model. This scenario represents the most realistic situation of the three cases.

3.2 Sampling and Convergence of the Posterior Distribution

When sampling eq. 2, we apply a Metropolis algorithm (Metropolis and Ulam 1949; Metropolis et al. 1953) to obtain a Markov chain whose stationary distribution is proportional to the Bayesian posterior probability. In scenarios 2 and 3, when the v_t's are treated as parameters in the model, we also employ a Gibbs sampler (Geman and Geman 1984). The priors on the parameters are uninformative; $p(M_{tot})$ and $p(a)$ are uniform in space, and the priors on the v_t parameters are uniform in v_t^2. The latter prior is chosen because v_t is a two-dimensional vector on the plane of the sky. We also assume that all tracers are bound to the galaxy, i.e. $\mathcal{E} > 0$. Thus, the maximum tangential velocity allowed in the isotropic model is

$$v_{t,max} = \sqrt{2\Psi(r) - v_r^2}, \tag{17}$$

which defines the upper bound in the prior for v_t.

For each set of data, we create three parallel Markov chains that have initial values far apart in parameter space. We assess convergence of the three chains with the convergence diagnostic $\widehat{R}$, as suggested by Gelman and Rubin (1992). The code we use to calculate $\widehat{R}$ is available in the R Software Statistical Computing Language in the CODA package (Plummer et al. 2006), and we use the SNOW package (Tierney et al. (2013)) for parallel computing.

4. Results

Scenario 1

The distribution of the mean parameter values from the 500 Markov chains in scenario 1 are shown in Fig. 1. The true parameter values are shown with a red dashed line, and the mean of each parameter estimate is shown as a black dot. As expected, on average the parameter estimates for M_{tot} and a returned the true parameter value. The standard deviation (SD) of each estimate from a Markov chain is comparable to the SD of all the estimates from all the chains (shown in the legend of Fig. 1).

We also check the $M(r)$ credible intervals at different r values, and find that they are completely reliable (see Eadie 2013, for more details). An $M(r)$ credible region is calculated from the stationary distribution of one of the Markov chains and is presented in Fig. 2. The true $M(r)$ profile is shown in red, and three credible regions are shown in different shades of teal.

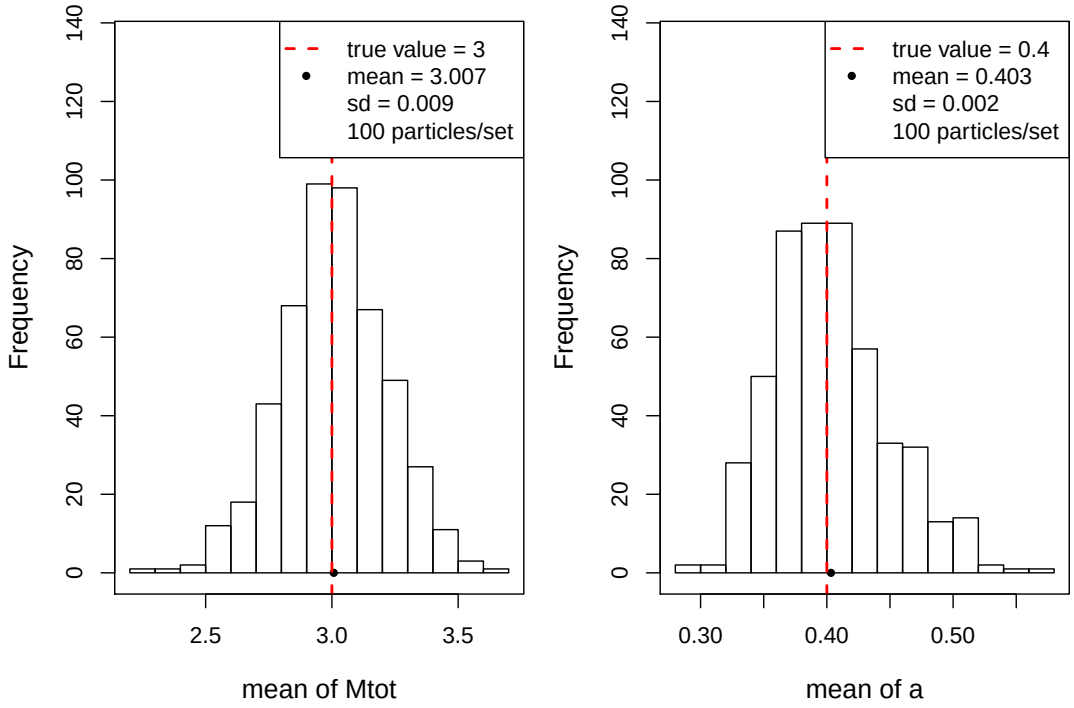

Figure 1: Distribution of the estimates of M_{tot} and a from scenario 1. M_{tot} is in units of $2.325 \times 10^9 M_\odot$, and a is in units of kpc.

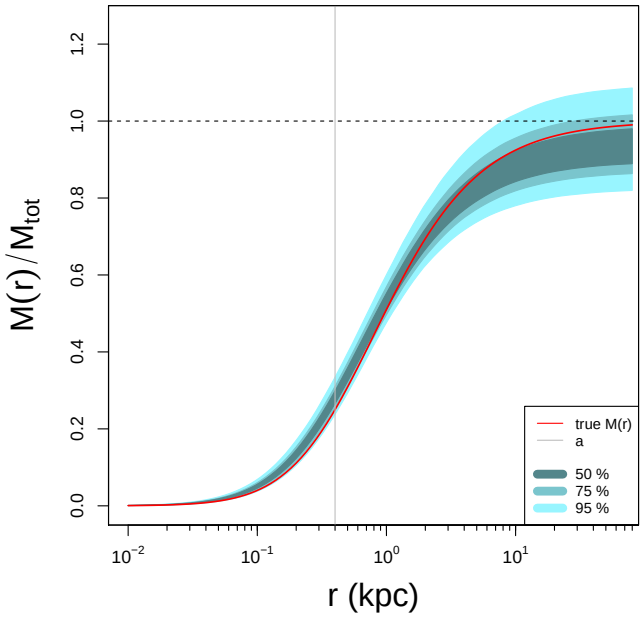

Figure 2: An example of $M(r)$ credible regions from scenario 1. The mass profile is scaled by the true parameter value of the total mass, $M_{tot} = 3$. The red line is the true $M(r)$ profile, and the shaded regions show the 50, 75, and 95% credible regions.

Scenario 2

In scenario 2 we treat the v_t values of the data as unknown, and sample them as parameters in the model via a Gibbs sampler. The distribution of the parameter estimates obtained in 500 independent analyses are shown in Fig. 3. In constrast to scenario 1, there are only 25 tracers in each data set. The decrease in data points causes the standard deviation of the parameter estimates to increase, as shown in the legends of Fig. 3

Fig. 4 shows an $M(r)$ credible region plot from one data set. Note that the width of the credible regions has increased compared to Fig. 2 as a result of the unknown v_t's and the smaller data sets. The true $M(r)$ curve fell within the 50% credible regions no less than 47% of the time, and fell within the 75% and 95% credible regions no less than 72% and 93% of the time.

Scenario 3

The final scenario considers data that is not only incomplete, but also follows a different DF from that of the assumed model. When incomplete anisotropic data is assumed to follow as isotropic model, we find a positive bias in both parameter estimates, as shown in Fig. 5. An example $M(r)$ credible region from scenario 3 is presented in Fig. 6. The true cumulative mass profile (red curve) falls within the 75% credible region for all r, but does not follow the overall shape predicted by the isotropic model. In particular, the predicted profile rises more steeply than the true profile, and reaches a higher overall total mass.

In scenario 3, the $M(r)$ credible regions are overconfident on average, as shown in Table 1. For example, over the 500 analyses in scenario 3, the true $M(r)$ profile falls within the 50% credible region only 32.6% of the time at $r = 4$kpc. The overconfidence in $M(r)$ credible regions is true for all values of r, with the overconfidence becoming worse at larger r. The innermost regions ($r < 0.20$kpc) are described reasonably well by the isotropic model, but as r increases, the credible regions become quite overconfident. At $r = 80$kpc, the true cumulative mass profile falls within the 95% credible region only 78% of the time.

The reason for a biased mass estimate in scenario 3 is most likely caused by the assumptions made by the isotropic model. The isotropic model assumes that there is an even mixture of orbit types at all r values. However, in OM-type anisotropy the orbits become radially biased at larger r, effectively lowering the tracers' v_t values. Without knowing tangential velocities or the true anisotropy of the system, the isotropic model estimates a high v_t parameter at large r. This behavior is shown in Fig. 7, where the true v_t values of the tracers are orange diamonds, and the estimates of v_t and their 95% credible intervals are shown as blue dots with error bars. The light blue curve is the maximum allowed v_t for a bound particle, calculated using the isotropic model parameter estimates from the analysis and eq. 17. The orange curve is the maximum allowed v_t for a bound particle given the true model parameters and the OM-type Hernquist model. In both models, the height of the curves are determined mainly by the total mass parameter M_{tot}. At large r, the acceptable tangential velocities in the OM-type Hernquist model are much lower than those in the isotropic model. Thus, when an isotropic model is assumed, v_t parameter values that would normally cause tracers to be unbound in an OM-type model are accepted, driving up the parameter estimate for the total mass.

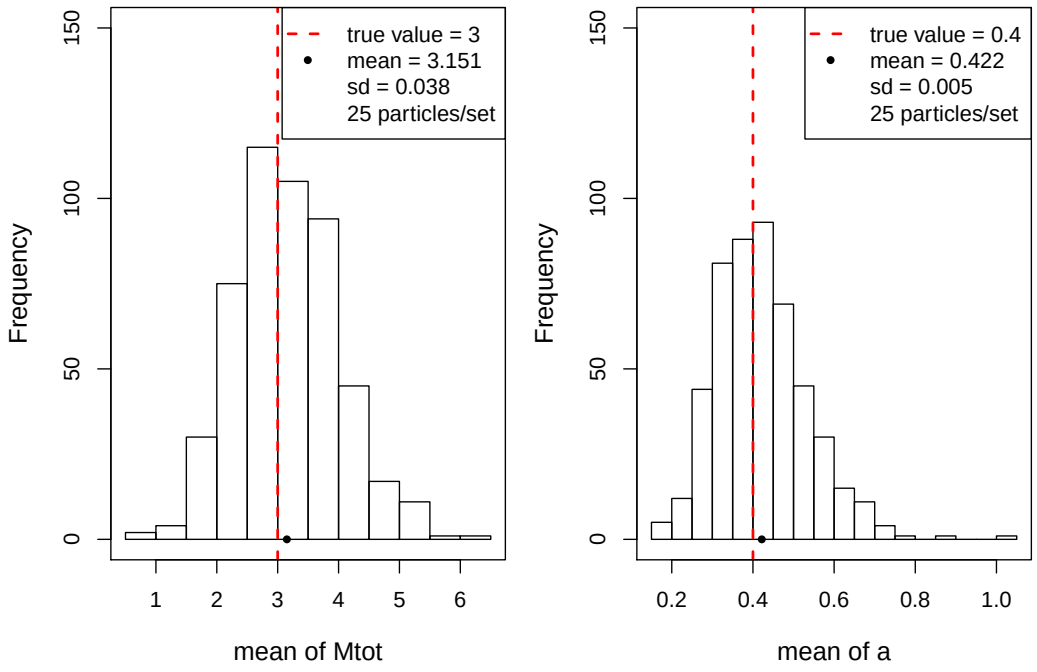

Figure 3: Distribution of parameter estimates for 500 data sets in scenario 2. The true parameter values are shown as red dashed lines, and the mean of the parameter estimates are black dots. M_{tot} is in units of $2.325 \times 10^9 M_\odot$, and a is in units of kpc.

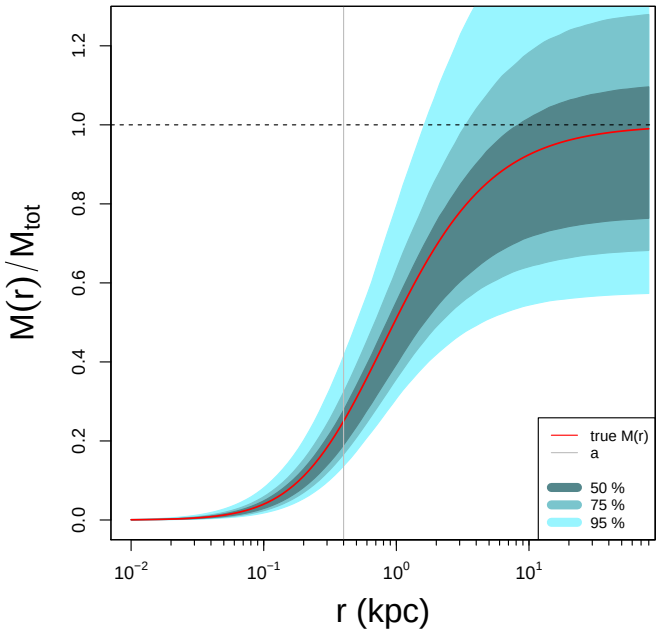

Figure 4: Example credible regions for one data set in screnario 2.

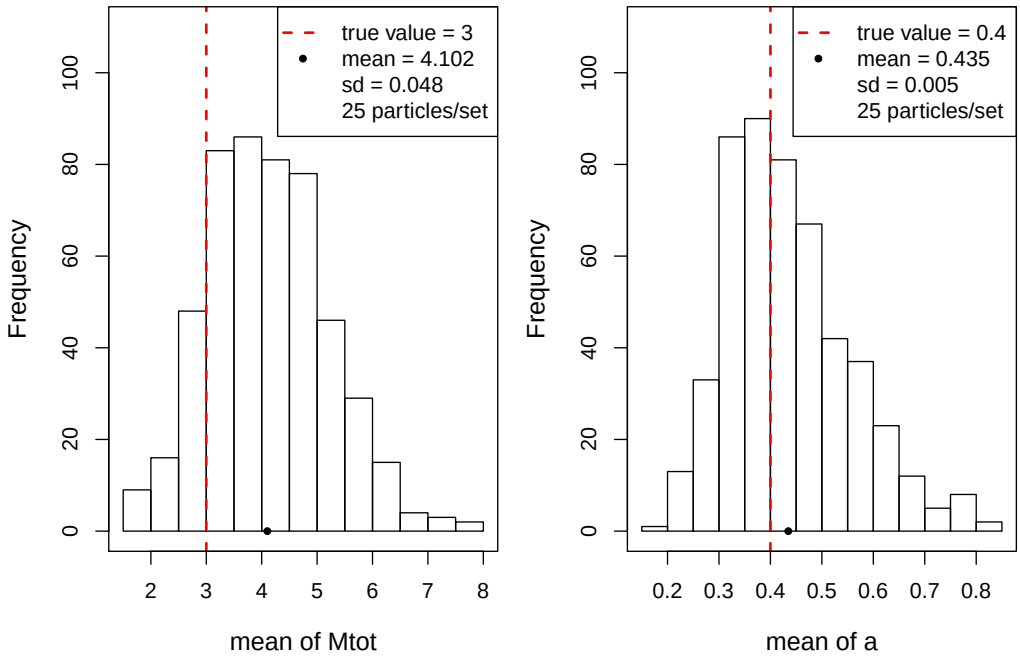

Figure 5: The distribution of parameter estimates from 500 analyses in scenario 3. M_{tot} is in units of $2.325 \times 10^9 M_\odot$, and a is in units of kpc.

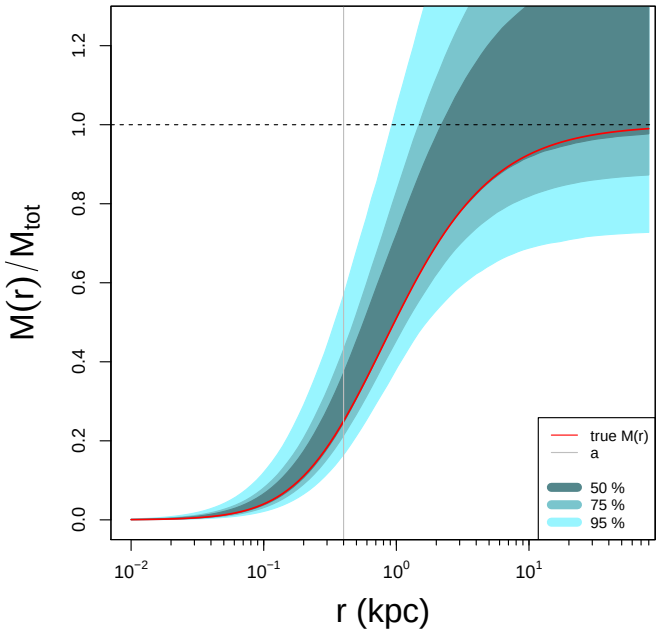

Figure 6: Example $M(r)$ credible regions for scenario 3. The dashed line at $M(r)/M_{tot} = 1$ corresponds to the true total mass, and the vertical grey line is the true scale radius a.

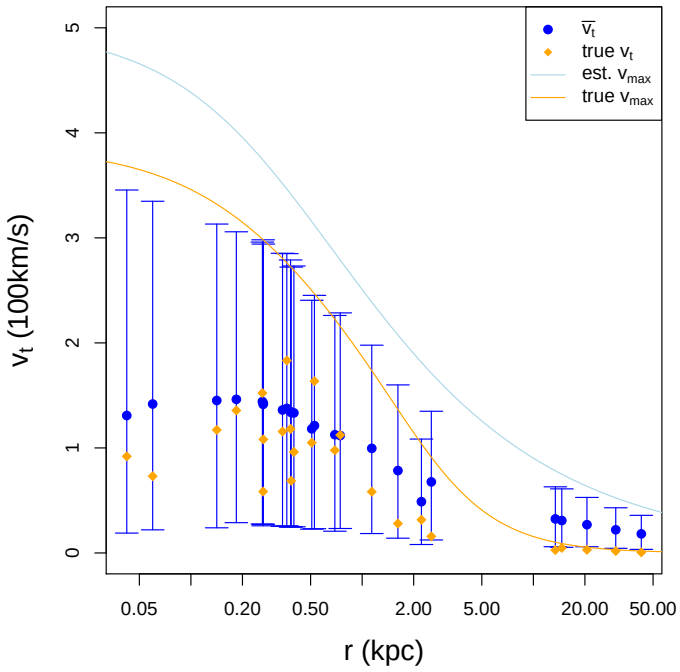

Figure 7: Tangential velocity estimates as a function of radius for one data set in scenario 3. The light blue curve shows the maximum speed for a bound particle given the parameter estimates from the posterior, while the orange curve shows the maximum speed given by the true anisotropic model. The blue dots are the estimates of v_t and the error bars represent 95% credible intervals. The orange diamonds are the true tangential velocities.

r **(kpc)**	within 50%	within 75%	within 95%
0.01	46.6	70.8	93.0
0.03	46.4	70.6	92.4
0.07	45.6	70.0	92.2
0.20	42.2	67.8	91.6
0.54	37.4	61.8	90.0
1.47	31.8	55.2	83.0
4.00	32.6	52.2	78.4
10.86	32.6	50.6	78.4
29.47	32.6	52.0	78.0
80.00	32.6	52.0	78.0

Table 1: Reliability of the credible regions in scenario 3 at different r values. The numbers in the second to fourth columns show the percentage of times that the true $M(r)$ profile fell within the corresponding credible region.

5. Discussion

The analyses carried out for scenario 1 were mainly for testing purposes, and they demonstrated that the parameter estimates for M_{tot} and a can be well estimated when the data and the model have the same DF. The test also showed that the credible regions for $M(r)$ in the isotropic Hernquist model are reliable when the data comes from the same distribution.

We found in both scenarios 2 and 3 that the lower number of particles and the incomplete data caused the credible regions for $M(r)$ to widen significantly, and that the distribution of parameter estimates will have a larger variance. While the estimates in scenario 2 remained unbiased, the estimates in scenario 3 showed a positive bias. Furthermore, the $M(r)$ credible regions in scenario 3 became somewhat unreliable at large r.

The analyses performed here do not represent a realistic scenario for a number of reasons. For one, our analyses looked at incomplete data or complete data— not both simultaneously. When using tracer objects to estimate the mass of the Milky Way, at least half of the tracers *do* have tangential velocity measurements. This information is invaluable to the analysis and we expect that it would help reduce biases such as those seen in scenario 3. Second, we did not include any measurement errors or observational uncertainties in the analysis. In astronomy, the technical capabilities of telescopes and optical equipment is believed to be very well understood, and thus measurement errors are good indicators of data quality. Therefore, they could be included in the analysis via a hierarchical model and used to weight the data. Third, recent evidence suggests that the dark matter halo may not have spherical symmetry, and instead may be triaxial in shape (e.g. Law et al. 2009). Fourth, all of the simulated tracers in our analysis are bound to the system. In nature, we don't always know if satellites and tracers are bound to the Milky Way. The dwarf galaxy Leo I was a contentious object in this regard for many years, and only after its proper motion was measured and taken into account has the issue been somewhat settled (it's likely to be bound, see Boylan-Kolchin et al. 2013). The boundedness of some other dwarf galaxies that are often used as tracers are still in question, however, such as the Magellanic Clouds (Besla et al. 2007). It would be useful to investigate the effects of bound and unbound tracers with more simulations, to see how strongly they affect the parameter estimates.

From Fig. 7 it is clear that the v_t estimates are not well constrained, as the estimates of v_t tend to go to an average value between $v_t = 0$ and v_{max} for all r. However, the v_t estimates provide insight into the reasons for biases in the M_{tot} and a parameter estimates.

The results from scenario 3 seem to imply that the isotropic Hernquist model might not do well with a real astrophysical problem if the velocity distribution of the system is anisotropic. However, for the Milky Way we do have tangential velocity measurements for some satellites, which provides some constraint on the velocity anisotropy of the tracer population. By incorporating both complete and incomplete data simultaneously, the biases may reduce or increase. Thus, the next step is to analyse simulated complete and incomplete data simultaneously, and perform tests similar to those outlined here.

6. Conclusion

We have explored the reliability of the Bayesian method to uncover the true cumulative mass profiles of simulated data in three different situations, assuming the Hernquist model and an isotropic velocity distribution of tracer particles. The method is reliable when the simulated data and the model share the same DF, even when the data are incomplete. The parameter estimates are biased, however, when the simulated data and the model have different DFs, and the data are incomplete. The assumptions about the unknown tangential velocities made by the isotropic Hernquist model appear to be the cause of these biases,

but it is unclear if the bias will occur when a mixture of complete and incomplete data is analysed. Future simulations will need to be performed to investigate ths issue.

Despite the biases observed in our simulations of scenario 3, the true $M(r)$ profiles were still reasonably well estimated by the posterior distributions. If the methods presented here were used in an analysis of Milky Way tracer data, then it would be advisable to report results with at least 95% credible regions.

7. Acknowledgments

The author would like to thank her PhD supervisor W.E. Harris (McMaster University, Department of Physics & Astronomy Department, Hamilton, Ontario, Canada) and her MSc supervisor L.M. Widrow (Queen's University at Kingston, Department of Physics, Engineering Physics & Astronomy, Kingston, Ontario, Canada).

References

Besla, G., Kallivayalil, N., Hernquist, L., Robertson, B., Cox, T. J., van der Marel, R. P., and Alcock, C. (2007), "Are the Magellanic Clouds on Their First Passage about the Milky Way," *The Astrophysical Journal*, 668, 949–967.

Binney, J. and Tremaine, S. (2008), *Galactic Dynamics*, Princeton, 2nd ed.

Boylan-Kolchin, M., Bullock, J. S., Sohn, S. T., Besla, G., and van der Marel, R. P. (2013), "The Space Motion of Leo I: The Mass of the Milky Way's Dark Matter Halo," *The Astrophysical Journal*, 768, 140.

Eadie, G. M. (2013), "Measuring the Mass of a Galaxy: An evaluation of the performance of Bayesian mass estimates using statistical simulation," Master's thesis, Queen's University, Kingston, Ontario, Canada.

Gelman, A. and Rubin, D. (1992), "Inference from iterative simulation using multiple sequences," *Statistical science*, 7, 457–472.

Geman, S. and Geman, D. (1984), "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 6, 721–741.

Hernquist, L. (1990), "An analytical model for spherical galaxies and bulges," *The Astrophysical Journal*, 356, 359–364.

Law, D. R., Majewski, S. R., and Johnston, K. V. (2009), "Evidence for a Triaxial Milky Way Dark Matter Halo from the Sagittarius Stellar Tidal Stream," *The Astrophysical Journal Letters*, 703, L67–L71.

Little, B. and Tremaine, S. (1987), "Distant satellites as probes of our Galaxy's mass distribution," *The Astrophysical Journal*, 320, 493–501.

Merritt, D. (1985), "Spherical stellar systems with spheroidal velocity distributions," *Astronomical Journal*, 90, 1027–1037.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E. (1953), "Equation of state calculations by fast computing machines," *The journal of chemical physics*, 21, 1087.

Metropolis, N. and Ulam, S. (1949), "The Monte Carlo method," *Journal of the American Statistical Association*, 44, 335–341.

Osipkov, L. P. (1979), "Spherical systems of gravitating bodies with an ellipsoidal velocity distribution," *Soviet Astron. Letters*, 5, 42–44.

Plummer, M., Best, N., Cowles, K., and Vines, K. (2006), "CODA: Convergence Diagnosis and Output Analysis for MCMC," *R News*, 6, 7–11.

Tierney, L., Rossini, A. J., Li, N., and Sevcikova, H. (2013), *snow: Simple Network of Workstations*, r package version 0.3-13.