Approximate Bayesian Computation for Sculptor Dwarf Galaxy's Dark Matter Density Profile

Brendan S. McVeigh

Advisers:
Jessi Cisewski*
Matthew Walker†

February 2, 2016

*Yale Department Department of Statistics, jessica.cisewski@yale.edu
†Carnegie Mellon Department of Physics, mgwalker@andrew.cmu.edu

1 Introduction

The standard model for the evolution of cosmic structure is know as the ΛCDM (Lambda cold dark matter) model. ΛCDM has made successful predictions about the temperature of the cosmic microwave background (CMB) radiation and features of galaxy formation [1]. A prediction of the ΛCDM model is that a galaxy's dark matter profile of should approximately follow the form of a Navarro-Frenk-White (NFW) profile [2, 3]. The model predicts that in the absence of baryonic effects the dark matter profile with diverge as r^{-1} forming a cusp [4]. However, the evidence is mixed on the accuracy of the model's predictions for dark matter profiles on small scales, where the distribution of dark matter is strongly non-linear [5]. Another possible arrangement for the dark matter profile would be for it to follow a core structure. Under this arrangement a dark matter profile increases in density as one moves closer to the center of the galaxy up to some core radius. The density profile then remains roughly constant anywhere within this radius.

For a galaxy it is possible to estimate the total mass using two different methods. One method attempts to sum the mass of all observed stars by counting the number of stars in the galaxy and multiplying by the mass of a typical star. While this calculation discounts the mass of non-luminous objects such as planets the mass of such objects make up only a small portion of the total mass contained in a galaxy and is thus not a major source of error. A second method involves using the observed speeds and positions of the stars to and solving the gravitational equations for mass to infer the mass of the galaxy. In theory these two calculations should match, but in practice the laws of gravity predict a mass that is several times what is predicted by summing the mass of observed stars. Astronomers refer to the discrepancy as the mass of the dark matter in the galaxy. Physicists have never been able to observe dark matter directly and it is not known with any certainty what type of particles make up this additional mass. Determining where within a galaxy the dark matter principally resides will aide astronomers in there efforts to learn more about dark matter.

In this context the Sculptor dSph (dwarf spheroidal) and other dwarf galaxies are particularly interesting because they contain a higher ratio of dark matter to normal matter than larger galaxies. Sculptor is of sufficient mass that baryonic effects should not alter the shape of its dark matter profile and the ΛCDM model would predict that the dark matter density profile would follow a dark matter cusp [4]. Yet previous analysis have been unable to rule out parameter values of the NFW profile that are consistent with both a dark matter cusp and a dark matter core [4]. We hope to help contribute to the ongoing debate about appropriate parameter values for the NFW profile function using Approximate Bayesian Computing (ABC) to estimate credible intervals for the relevant parameters. The complexity of relationship between the dark matter profile and the observational data, positions and velocities of stars, makes the use of a likelihood-free technique to estimate credible intervals particularly attractive. A further complication is provided by the two distinct populations of stars, a metal rich population concentrated near the core of the galaxy and a metal poor population found primarily farther from the center, that make up the Sculptor dSph. Attempting to model these two populations simultaneously while also incorporating the random processes by which stars are observed would further complicate the estimation of a likelihood function. If we are able to rule out parameter values consistent with either a dark matter cusp or a dark matter core this would help bring clarity to the debate and help to either further validate ΛCDM or indicate areas where the theory can be further refined.

2 Data

The data we analyze was collected from the e ESO 2.2 m Wide Field Imager (WFI) and the VLT/FLAMES spectrograph[6, 7]. Observations of 636 individual stars in Sculptor dwarf spheroidal are included. The characteristics observed for each star are shown in Table 1. Figure 1 shows the positions of the observed stars as they would appear in the sky, projected onto a sphere. The velocity of stars with a high probability of being in Sculptor appears with a normal distribution as predicted by theory. A crucial step in analyzing the data is determining which stars in the observed region belong to Sculptor and which are members of the milky way galaxy that can be seen in the same region of the sky. For this we rely on previous analysis by Walker et. al [8]. We confirm, as shown in Figure 2, that the stars identified as metal-rich in Sculptor appear to be clustered in the center of the galaxy while the metal-poor stars are found primarily outside of

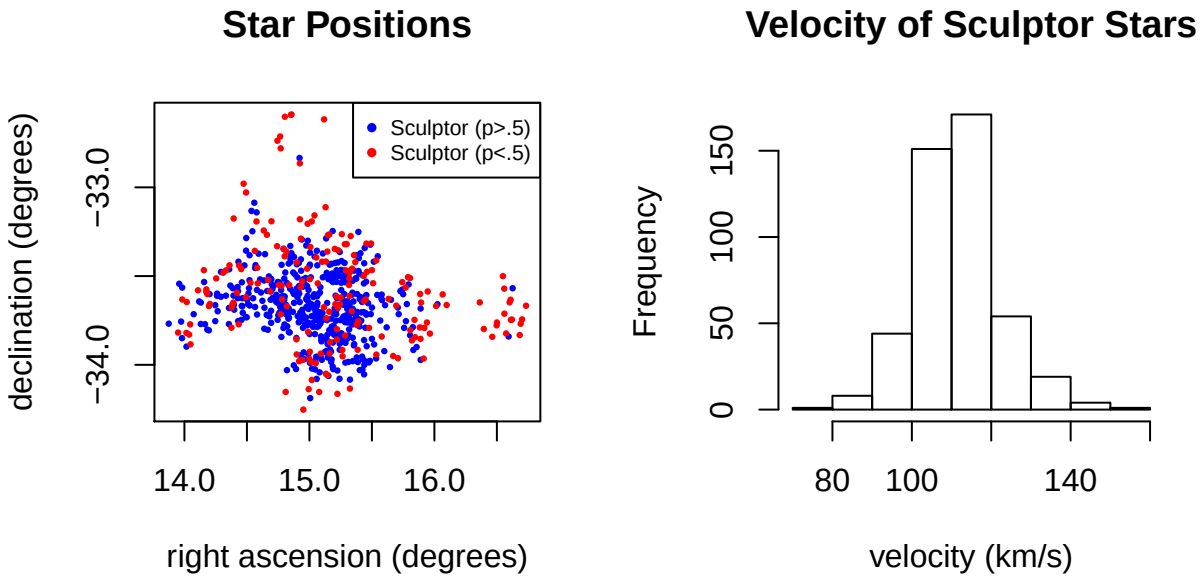

Figure 1: Position of observed stars and distribution of velocities.

the galaxy core, this is consistent with previous findings [4]. Once stars that are unlikely to be located in Sculptor, those with $p < .05$ are and stars are indeterminate metallicity are both excluded the remaining population contains metal-rich 93 and 294 metal-poor stars.

A final step that must be taken before the data can be used in our ABC algorithm is to transform the data to euclidean coordinates comparable to those generated by our galaxy simulation method presented in Section 3.4. The reported measurements, in degrees, place the stars on a celestial sphere which mas be mapped to x and y euclidean coordinates. For right ascension and declination pair (α, δ) the conversion to euclidean coordinate pair (x, y) is given by 1 and 2 where (α_0, δ_0) is mapped to $(0, 0)$.

$$x = \frac{\cos(\delta)\sin(\alpha - \alpha_0)}{\sin(\delta_0)\sin(\delta) + \cos(\delta_0)\cos(\delta)\cos(\alpha - \alpha_0)} \tag{1}$$

Variable	Units
right ascension	angle hours
declination	degrees
velocity	km/s
error in velocity	km/s
metallicity	Fe/H
error in metallicity	Fe/H
magnesium index	Angstroms
error in magnesium index	Angstroms
probability of Sculptor membership from EM algorithm	
metallicity class	0=excluded, 1=metal-rich, 2=metal-poor

Table 1: Star Characteristics

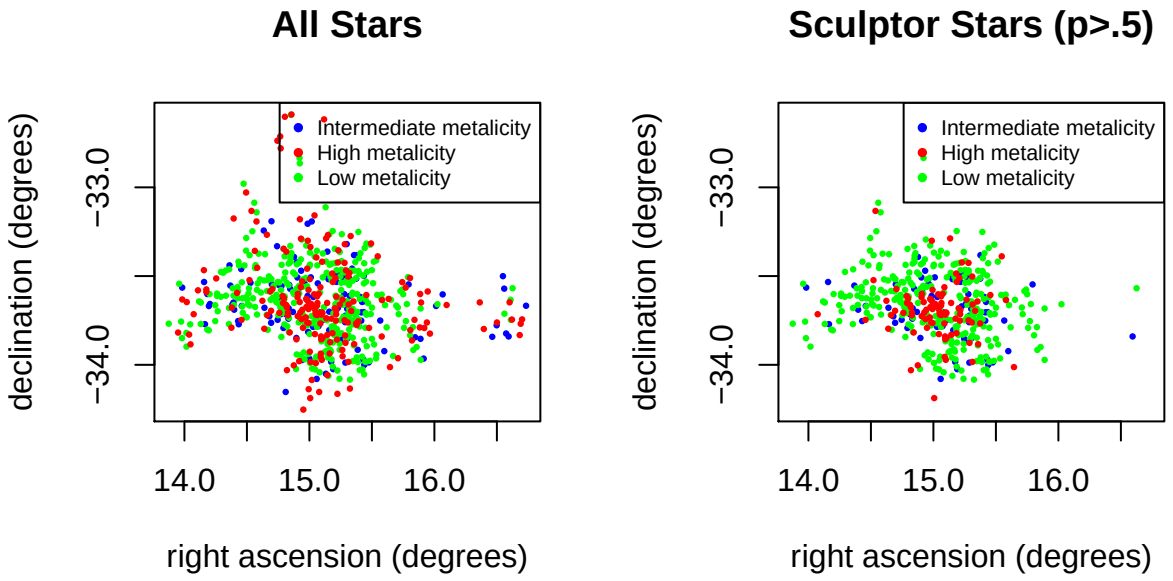

Figure 2: More metallic stars in Sculptor are clustered near the center of the dwarf galaxy.

$$y = \frac{\cos(\delta_0)\sin(\delta) - \sin(\delta_0)\cos(\delta)\cos(\alpha - \alpha_0)}{\sin(\delta_0)\sin(\delta) + \cos(\delta_0)\cos(\delta)\cos(\alpha - \alpha_0)} \tag{2}$$

3 Methodology

3.1 Approximate Bayesian Computing

In a standard Bayesian framework the observed data is used to condition the prior distribution $\pi(\theta)$, which contains a researcher's beliefs about the parameter of interest before examining the data. The result is a posterior distribution $\pi(\theta|x^n)$:

$$\pi(\theta|x^n) = \frac{\mathcal{L}(x^n|\theta)\pi(\theta)}{\int \mathcal{L}(x^n|\theta)\pi(\theta)d\theta} \tag{3}$$

Implementing this approach requires both that the likelihood function $\mathcal{L}(x^n|\theta)$ be known and that the integral, $\int \mathcal{L}(x^n|\theta)\pi(\theta)d\theta$ be tractable. For complex models a likelihood function may not be available. Even when the likelood function is known the integral in the denominator of 3 may not be tractable. Employing ABC allows us to sidestep these common barriers be approximating the posterior without the use of a likelihood function. ABC, although more computationally intensive, has been applied in areas such as cosmology[9], genetics[10], and psychology[11] where data can display complicated dependencies and a likelihood is thus particularly difficult to compute.

While ABC does not require a likelihood function it does require that the researcher specify a prior distribution, $\pi(\theta)$, a forward model $f(x|\theta)$ and a distance function $\rho(x^n, y^n)$. The prior distribution, as in other Bayesian settings, simply needs to specify a distribution for the parameter. The forward model must be a way to simulate data based on the parameter of interest. The distance function must provided a way to compare the observed dataset x^n with y^n, a dataset generated by the forward model. The distance function is generally computed from a summary statistic of the data but care must be taken in choosing an appropriate statistic. If an appropriate statistic is not chosen then the resulting ABC posterior will be a approximation of the true posterior.

Once the prior, forward model, and distance function have been specified the ABC algorithm can be run to sample N paremeter values from an approximate posterior. The closeness of the approximation to the true posterior varies with a specified tolerance ϵ with the ABC posterior converging to the true posterior as $\epsilon \to 0$. The steps of the algorithm are as follows:

1. Sample a proposed parameter θ^* from prior, $\pi(\theta)$
2. Simulate a dataset $(x_1^*, \ldots, x_n^*)$ from forward process, $f(x|\theta^*)$. The simulated dataset must be the same size as the observed dataset.
3. Accept θ^* if $\rho(x^n, x^{n*}) \leq \epsilon$ where ϵ is the chosen tolerance level.
4. Repeat 1-3 until N values have been accepted

As with other computation methods there is a trade-off between the speed of the algorithm and the closeness of the approximation as determined by the acceptance threshold ϵ. In practice a sufficiently small level of ϵ can be determined by repeatedly running the algorithm for lower and lower threshold levels until further decreases in ϵ do not appear to change the result. In this case the approximate posterior can be said to have "converged".

3.2 Population Monte Carlo Sampling

We implement a modified version of the ABC algorithm known as ABC population Monte Carlo sampling (ABC PMC) as described in [11]. The ABC PMC algorithm takes an iterative approach, improving the posterior approximation over the course of multiple time steps. In the standard ABC algorithm observations are repeatedly drawn from the prior until a sufficient number of proposals have been accepted. ABC PMC algorithm instead uses importance sampling to draw from sets of previously accepted proposals parameters. The previously accepted parameters provide an approximation to the unknown true posterior distribution. As a result sampling method can produce a greater proportion of proposal parameters that are likely to generate data that is close, as measured by the distance function, to the observed data.

Implementing the ABC PMC algorithm requires that we specify a transition kernel which is used in the algorithm avoid simply resampling proposal parameter values identical to those of previously accpeted parameters. When running the ABC PMC algorithm an initial approximate posterior is calculated

1. Sample a M proposed parameter values $\theta_1^*, \ldots, \theta_M^*$ from prior, $\pi(\theta)$
2. For each proposed parameter value simulate a dataset $(x_1^*, \ldots, x_n^*)$ from forward process, $f(x|\theta_i^*)$.
3. Compute $d_i = \rho(x^n, x_i^{n*})$ for each of the M generated datasets.
4. Accept the proposed parameters corresponding $d_{(1)}, \ldots, d_{(N)}$ the smallest N distances and let $\epsilon_1 = d_{(N)}$.
5. Define weights $\{w\}_1$ where $w_1, \ldots, w_n = 1/N$

An approximate posterior for time step t is computed from the approximate posterior at time $t-1$ as follows

1. Select a threshold $\epsilon_t < \epsilon_{t-1}$
2. Sample a proposal θ' from $\{\theta^*\}_{t-1}$ using multinomial sampling with weights $\{w\}_{t-1}$
3. Use the transition kernel $q(\dot{|}\theta')$ to generate θ^*.
4. Simulate a dataset $(x_1^*, \ldots, x_n^*)$ from forward process, $f(x|\theta^*)$.
5. Accept θ^* if $\rho(x^n, x^{n*}) \leq \epsilon_t$.
6. Repeat 2-5 until N values have been accepted

7. Calculate $\{w\}_t$, the set of weights associated with $\{\theta^*\}_t$ using 4

In our algorithm we use a truncated Gaussian kernel for our noise function. We truncate the distribution so that it will not generate values in regions of parameter space where the density of the prior distribution is zero. After the desired number of parameter values have been accept in a given time step they are then assigned weights according to equation 4 where $q(\theta_{i,t}|\theta_{j,t-1}, \sigma_{t-1})$ is the density of the kernel evaluated at the accepted parameter value $\theta_{i,t}$. For this approach to work the sampling from the previously accepted particles must be done with appropriate weights with appropriate weights.

$$w_{i,t} = \frac{\pi(\theta_{i,t})}{\sum_{j=1}^{N} w_{j,t-1} q(\theta_{i,t}|\theta_{j,t-1}, \sigma_{t-1})} \tag{4}$$

This contrasts with the straight forward approach of the standard ABC algorithm. In successive time steps the weights are updated allowing from a more intensive sampling of regions of parameter space that often generate datasets with smaller distances. As a result tthis algorithm can yield faster convergence rates.

3.3 Navaro-Frenk-White Model

Strigari, Frenk, and White model the mass density profile of dark matter as a function of the specific energy, E, and the specific angular momentum of a star J [4]. Where both E and J are specified by the position and velocity vectors of a star

$$E = v^2/2 + \Phi(r) \tag{5}$$

$$J = vr \sin\theta \tag{6}$$

and the function $\Phi(r)$ is defined as

$$\Phi(r) = \Phi_s \left(1 - \frac{\ln(1 + r/r_s)}{r/r_s}\right) \tag{7}$$

The constants Φ_s and r_s are specified by model parameters.

$$f(E, J) = g(J)h(E) \tag{8}$$

$$g(J) = \begin{cases} [1 + (J/J_\beta)^{-b}]^{-1}, & \text{for } b \leq 0 \\ 1 + (J/J_\beta)^b, & \text{for } b > 0 \end{cases} \tag{9}$$

$$h(E) = \begin{cases} N E^a (E^q + E_c^q)^{d/q} (\Phi_{lim} - E)^e, & \text{for } E < \Phi_{lim} \\ 0, & \text{for } E \geq \Phi_{lim} \end{cases} \tag{10}$$

where we select N, the normalizing constant such that such that

$$\int_0^{r_{lim}} 2\pi \int_0^{\pi} \sin\theta d\theta \int_0^{v_{esc}} f(E, J) v^2 r^2 dv dr = 1 \tag{11}$$

Where r_{lim} is specified by the model parameters, $\Phi_{lim} = \Phi(r_{lim})$, and v_{esc} varies with the radius according to the function

$$v_{esc}(r) = \sqrt{2(\Phi_{lim} - \Phi(r))} \tag{12}$$

3.4 Galaxy Simulation

The implementation of and ABC algorithm requires the construction of a forward model, for our problem this means that we must be able to simulate a galaxy consistent with the NFW density profile described by a set of parameters that we can compare with our observed data. Thus, we must be able to sample from the marginal distributions of the positions and velocities of the stars, which are unknown. Based on the model in the previous section we use a rejection sampling algorithm to generate observations from these distributions conditional on given parameter set even though we do not have a closed form for them. The basic rejection sampling algorithm requires the following steps:[1]

1. Generate a proposal but sampling a radius, tangential velocity, and radial velocity uniformly between 0 and the maximum physical value allowed by the parameter set.
2. Sample a threshold uniformly from 0 to 1.
3. Compute the ratio between the density at the sampled radius and velocities and the maximum density point in the parameter set scaled up by a factor M.
4. Accept the proposal if the ratio is greater than the sampled threshold
5. Reject the proposal and repeat the procedure if the ratio is less than the sampled threshold

We use a factor of 1.1 for M. The resulting points contain a radius, as well as tangential and radial velocities. There are then converted to euclidean coordinates by assuming spherical symmetry in position and circular symmetry for the tangential component of velocity.

We also experimented with a modified rejection sampling algorithm in which we approximated our density function with a step function. This greatly reduced the time necessary to generate a sample galaxy as checking the proposal density relative to a step function is a much more efficient calculation. However, we ultimately abandoned this approach due to the failure of the resulting ABC posteriors to converge to the true posterior distribution in simulated examples as discussed further in Section 3.6.

3.5 Reduced Model Posterior

We chose to implement and ABC algorithm specifically because we were unable to calculate an exact posterior distribution due to the complexity of our model. However, as is commonly done we were able to calculated an exact posterior for a reduced version of our model. In the reduced version we vary only a single parameter, E_c the critical energy, and modify our density function to include only $h(E)$ as defined in 10. For our simulated galaxy used in the posterior calculation we use a parameter set reported by Strigari, Frenk, and White for the metal-rich star population in as shown in Table 2.

We use a simulated galaxy with a population of 93 stars, the same number as identified in the metal rich population of Sculptor, and adopt a prior function for E_c with a density of Uniform(.01, 1). We bound the prior away from 0 to avoid potentially huge computation times due to inefficient sampling for very small

[1] The code for this was written by Mao Sheng (Terrence) Liu of the Carnegie Mellon Department of Physics

Star Population	a	d	e	E_c	Φ_{lim}	r_{lim}	b	q	J_β	V_{max}	r_{max}
Metal Rich	2.0	-5.3	2.5	0.16	0.45	1.5	-9.0	6.9	8.6×10^{-2}	21	1.5
Metal Poor	2.4	-7.9	1.1	0.17	0.60	3.0	0.0	8.2	8.6×10^{-2}	21	1.5

Table 2: Parameters values for which Strigari, Frenk, and White were unable to reject the hypothesis that the parameters were consistent with the Sculptor data [4]

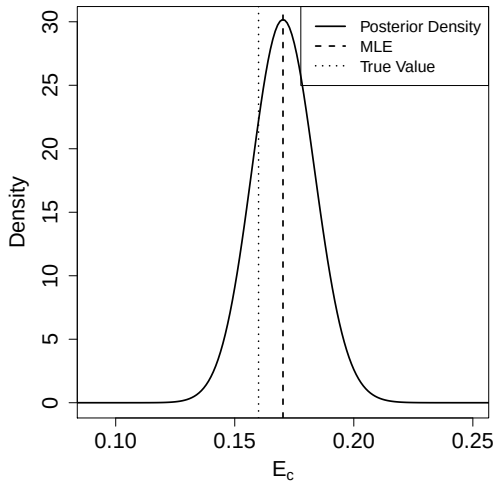

Figure 3: Posterior distribution for E_c based on simulated galaxy with 93 stars, a $E_c = 0.16$, and Uniform(.01,1) prior

parameter values. Setting the other parameters to the values listed for the metal-rich population in Table 2 we then calculate the posterior density

$$\pi(E_c|x^n) = \frac{\prod_{i=1}^{93} \left[h(E_i) r_i^2 v_i^2\right]}{\int_{.01}^{1} \prod_{i=1}^{93} \left[h(E_i) r_i^2 v_i^2\right] dEc} \tag{13}$$

The results of our calculation are shown in Figure 3. We note that with a population comparable in size to what we actually observe the posterior distribution is centered on the maximum likelihood estimate and is close to the true parameter value.
Since there is no guarantee that the true parameter value is in fact the one that was reported in Table 2 we also examine the posterior densities for galaxies simulated using different parameter values of $\mathrm{E_c}$ as shown in Figure 4. Interestingly we note that the posterior density flattens substantially for larger values of $\mathrm{E_c}$. In particular if the density function is noticeably different from zero for values above 0.40 we see that the tail flattens almost completely indicating that the parameter has very little effect on galaxy shape above this value. This is consistent with findings that indicated $\mathrm{E_c}$ having very little effect on galaxy shape when $E_c > \Phi_{\mathrm{lim}}$, which is about 0.45 in this case.

3.6 Selection of Summary Statistics and Distance Functions

Appropriate summary statistics are necessary to distinguish between data sets similar to the observed data and data sets that are substantively different. Frequently, the means of the observed data, position or velocity coordinate values for our data, would be tested as summary statistics before considering more complicated statistics. However, in the case of our model the mean of position and velocity will be zero in expectation regardless of the parameter values. Thus, originally considered distance functions based on the mean radius of simulated stars as well as the second moment of both the velocities and radii. However, ABC posteriors calculated using distance functions based on these summaries did not appear to contain enough information for the ABC posteriors to approximate the true posterior as well as desired.

We next tried several nonparametric statistics in an attempt to capture the information contained such as distances based on Kolmogorov-Smirnov statistic between smoothed empirical CDFs for star radii and velocities. We also considered using an L^2 norm. While both of these statistics appeared to strong differentiate

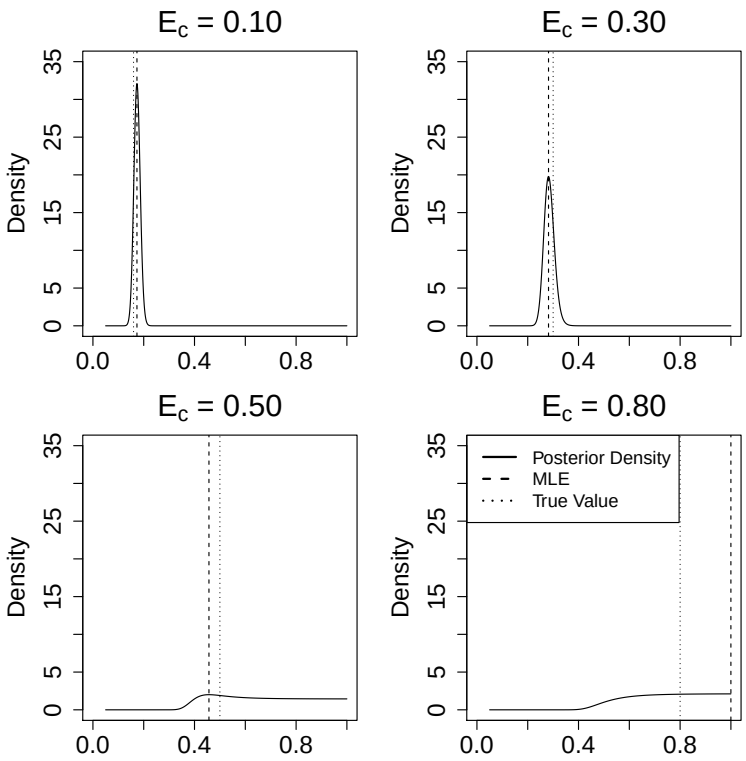

Figure 4: Posterior distribution for E_c based on simulated galaxy with 93 stars, E_c values of 0.10, 0.30, 0.50, and 0.80 using a Uniform(.01,1) prior

between parameter values we still were unable to achieve convergence for the ABC posteriors.

A flaw with the distance metrics discussed so far is that they assume independence between a star's radius and velocity. This is not a valid assumption and thus it is preferable to consider a distance which does not assume independence. So we developed two additional distance functions, one based on a 2D kernel density estimate across the velocity and radii or the stars. The second is based on the quantity rv^2, an appearing quantity because it has the same units as mass and radius seemed an appropriate weighting to apply to velocity. To produce our final statistic we compute an L^2 norm between the estimates density estimates for two different galaxies and between the empirical CDFs where we use the rv^2 statistic. As shown in Figure 5 these both also appear to be able to distinguish between two galaxies with different E_c parameters. We note that the dark upper square in both plots indicates that there is little power to determine between galaxies with larger E_c parameter values. However, this is consistent with the findings of broader posterior distributions for parameter values and in range. Indeed the nearly flat posterior distributions shown in Figure 4 indicate that no set of summary statistics would be able to detect the difference between these distributions.

Despite using the two final distances adopting different thresholds for each distance which we shrink independently. To be accepted a proposed parameter value must be below both thresholds. To evaluate if these statistics would allow our ABC algorithm to correctly estimate posterior densities we ran several test simulations the results of which indicated that the statistics did not provide enough information and thus were unable to generate densities with appropriate maximum densities as shown in the left panel of Figure 6. Here we show successive iterations of the ABC PMC algorithm which suggest that the algorithm has converged.

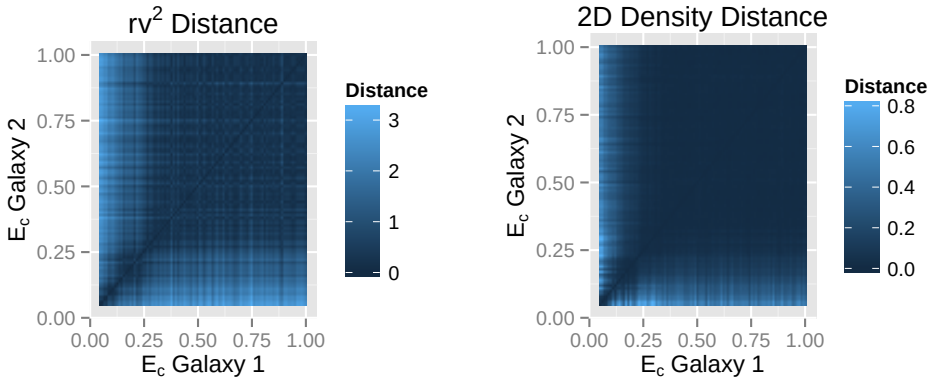

Figure 5: Sample distances varying E_c parameter using simulated galaxies with 500 stars. All pairwise distances.

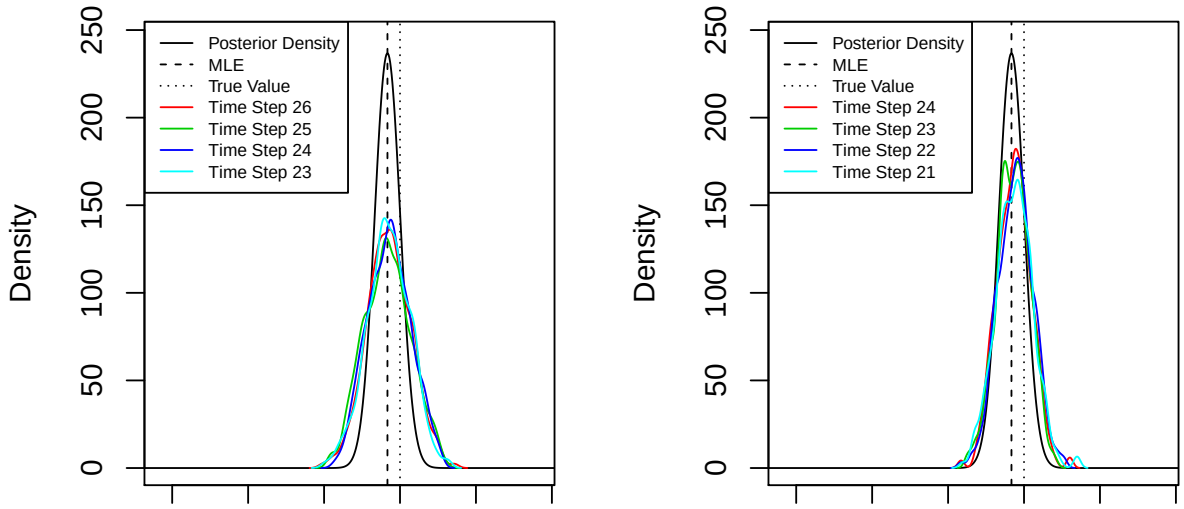

Figure 6: Comparison between ABC Posteriors and actual posteriors. The left panel was calculated with the step-function rejection sampler to simulate galaxies while the right panel was computed using the slower standard rejection sampler simulation.

Star Population	Strigari Value	ABC Maximum Value	95% Lower Bound	95% Upper Bound
Metal Rich	0.16	0.28	0.22	0.96
Metal Poor	0.17	0.3	0.26	0.34

Table 3: Estimates and 95% Credible Interval for ABC posterior for Sculptor E_c parameters

4 Results

4.1 Simulated Data

Using the selected distance metric we run our ABC algorithm using the simplified density model described in the previous section the same simulated dataset with known parameter values as our "observed" data. This allowed us to compare the posterior distribution generated by our ABC algorithm with the true posterior computed from the simulated data and the simplified model. The results are shown in the right panel of Figure 6. We can see that the posterior generated by successive iterations of the ABC PMC algorithm approximate the true posterior with increasing accuracy. While the maximum density achieved by the ABC algorithm is somewhat lower than that of the true posterior this is not surprising given that the posterior is only approximate. Running additional steps in the algorithm would result in an improved approximation however, the additional computation time would be considerable and thus was not implemented. We also note that our ABC posteriors are centered on the true posterior with the difference being that they are somewhat wider. Thus, we can expect that any credible intervals generated using our ABC posteriors would contain more values and an equivalent credible interval using the unknown true posterior.

4.2 Sculptor Estimates

We then ran our ABC PMC algorithm, using the full galaxy simulation model, on the metal-rich and metal-poor stars observed in Sculptor separately. These populations contain only 93 and 294 stars respectively so as we expect the posterior distributions are considerably less concentrated than those for larger simulated galaxies. We once again estimate only the critical energy relying on previous estimates to set the other parameter values. Our ABC posterior densities are shown in Figure 7 and maximum density estimates along with a 95% credible interval are shown in Table 3. In both cases we find that the previously estimated value for E_c lies outside of our 95% credible interval. In the case of the metal-poor population the posterior density puts almost no weight of the value of 0.17 for E_c estimated by Strigari, Frenk, and White. We also note that the metal-rich posterior distribution has a low and relatively flat region for $E_c > \Phi_{\text{lim}} = 0.45$. This is consistent with the findings of our distance functions as shown in Figure 5 that galaxies with E_c values in that range appear very similar.

5 Discussion

For the single parameter we estimated our results are inconsistent with those reported previously [4]. We construct our 95% credible intervals by taking range of the middle 95% of parameter weights generated by the last step of our ABC simulations. It is important to note our estimate is conditional upon the other parameter values. It is possible that if credible intervals were calculated for all parameters simultaneously that our joint intervals would not be inconsistent with the previously reported values. The simplest explanation may be that different methods of likely we lead to different conclusions however, it is also possible that our ABC algorithm is able to extract more information out of the data. The previous parameter values reported were only found not to be rejected in a test for consistency with the data. The test performed, a chi-squared test on binned data likely discarded much of the information contained in the data.[4]

Future work will include attempts to estimate the posterior distributions of other parameters both individually and collectively. In particular it would be useful to create a more sophisticated galaxy simulation model.

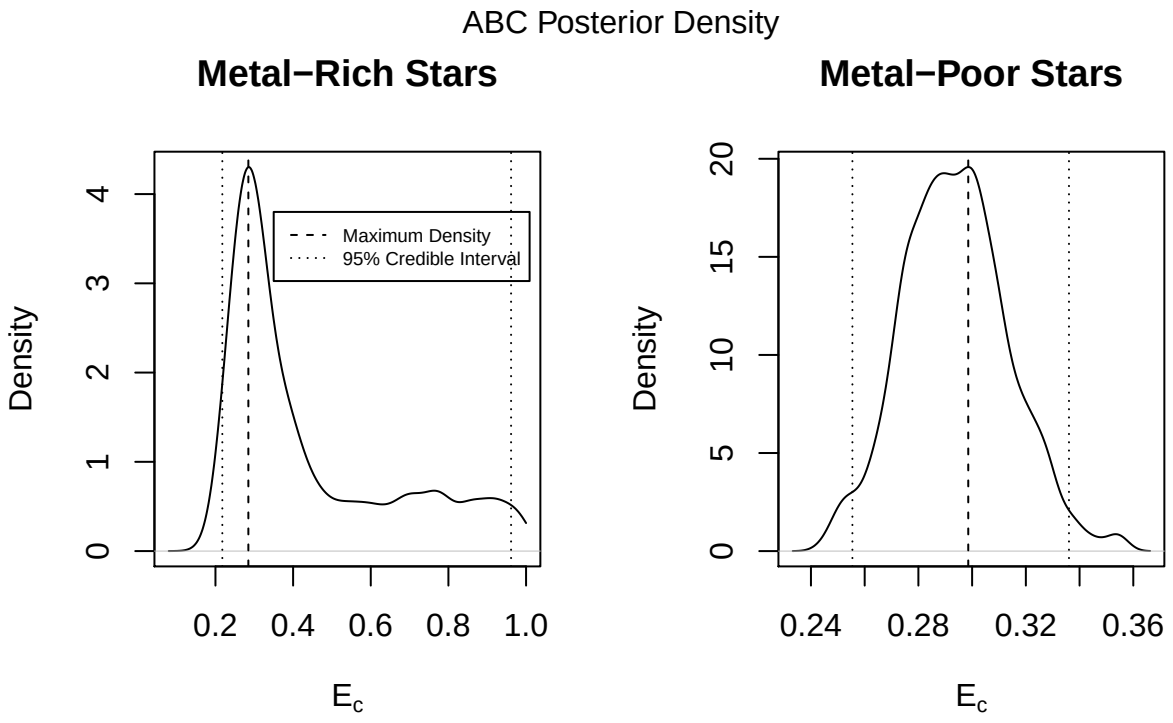

Figure 7: Posterior distribution for E_c for metal-rich and metal-poor star populations in Sculptor.

An improved model would simulate not just observed and classified stars but would be capable of modeling the entire galaxy, including unobserved stars and stars between Sculptor and Earth that must be removed from the population as well as the probabilities of classification as metal-rich and metal-poor. Modeling both populations simultaneously as well as capturing other random would allow for greater power when estimating parameter values that affect both metal-rich and metal-poor populations as well as providing a more accurate model of the data generating process. We could also investigate other methods of likelihood free inference such as the Spectral Series Method developed by Izbicki, Lee, and Schafer[12]. .

References

[1] W. S. D. M. Frenk C. S., "Dark matter and cosmic structure," *Annalen der Physik*, pp. 507–534, 2012.

[2] W. S. D. M. Navarro J. F., Frenk C. S., "The structure of cold dark matter halos," *ApJ*, p. 563, 1996.

[3] ——, "A universal density profile from hierarchical clustering," *ApJ*, p. 493, 1997.

[4] C. S. F. Lous E. Strigari and S. D. M. White, "Dynamical models for the sculptor dwarf spheroidal in a lambda cdm universe," *arXiv:1406.6079 [astro-ph.GA]*, 2014.

[5] M. G. Walker and J. Penarrubia, "A method for measuring (slopes of) the mass profiles of dwarf spheroidal galaxies," *ApJ*, p. 19, 2011.

[6] E. Tolstoy, M. J. Irwin, A. Helmi, G. Battaglia, P. Jablonka, V. Hill, K. A. Venn, M. D. Shetrone, B. Letarte, A. A. Cole, F. Primas, P. Francois, N. Arimoto, K. Sadakane, A. Kaufer, T. Szeifert, and T. Abel, "Two Distinct Ancient Components in the Sculptor Dwarf Spheroidal Galaxy: First Results from the Dwarf Abundances and Radial Velocities Team," *apjl*, vol. 617, pp. L119–L122, Dec. 2004.

[7] G. Battaglia, A. Helmi, E. Tolstoy, M. Irwin, V. Hill, and P. Jablonka, "The Kinematic Status and Mass Content of the Sculptor Dwarf Spheroidal Galaxy," *apjl*, vol. 681, pp. L13–L16, Jul. 2008.

[8] M. G. Walker, M. Mateo, E. W. Olszewski, B. Sen, and M. Woodroofe, "Clean Kinematic Samples in Dwarf Spheroidals: An Algorithm for Evaluating Membership and Estimating Distribution Parameters When Contamination is Present," *aj*, vol. 137, pp. 3109–3138, Feb. 2009.

[9] W. W.-V. A. Weyant, C. Schafer, "Likelihood-free cosmological inference with type ia supernovae: Approximate bayesian computation for a complete treatment of uncertainty." *The Astrophysical Journal*, p. 116, 2013.

[10] D. B. M. Beaumont, W. Zhang, "Approximate bayesian computation in population genetics," *Genetics*, pp. 2025–35, 2002.

[11] T. V. Z. Brandon M. Turner, "A tutorial on approximate bayesian computation," *Journal of Mathematical Psychology*, pp. 69–85, 2012.

[12] C. S. R. Izbicki, A. Lee, "High-dimensional density ratio estimation with extensions to approximate likelihood computation," *Artificial Intelligence and Statistics Conference*, 2014.