

# GRAVIMAGE: A non-parametric method for mass modelling spherical systems

P. Steger<sup>1</sup>\*, J. I. Read<sup>2</sup>

<sup>1</sup>Institute for Astronomy, Department of Physics, ETH Zürich, Wolfgang-Pauli-Strasse 27, CH-8093 Zürich, Switzerland

<sup>2</sup>Department of Physics, University of Surrey, Guildford, GU2 7XH, UK

13 November 2014

## ABSTRACT

We present a new non-parametric method based on Jeans modelling to determine the mass distribution in spherical systems with single or multiple tracer populations. A high dimensional parameter space encoding tracer density, line of sight velocity dispersion and total mass density is sampled with MULTINEST.

Without assumptions on the functional form of any of these profiles, we show that we can successfully recover the radial density profile of spherical mock dwarf galaxies. With two populations with  $\sim 5,000$  tracers in each, we begin to disentangle the degeneracy between dark matter density and tracer velocity anisotropy, determining the dark matter cusp slope at the half light radius at nearly  $2\sigma$  confidence. We also test our method on triaxial mock data for which our method is expected to become biased. We find that with  $\sim 10,000$  tracers, this bias remains smaller than our other uncertainties unless we are staring ‘down the barrel’ of the triaxial figure. This leads to the perhaps counterintuitive result that GRAVIMAGE can be reliably applied to triaxial systems so long as they *do not* appear circularly symmetric on the sky.

**Key words:** galaxies: dwarf – galaxies: fundamental parameters – galaxies: kinematics and dynamics – cosmology: dark matter

## 1 INTRODUCTION

Cosmological  $\Lambda$ CDM simulations predict a hierarchical self-similar assembly of dark matter halos (e.g. White & Rees 1978; Navarro et al. 1996b). Modelling only the dark matter fluid in the absence of ‘baryons’ (stars and gas), Dubinski & Carlberg (1991) found that the density profile of resulting halos are best described by a split power law that diverges as  $r^{-1}$  at the centre. Navarro et al. (1996b) demonstrated that this profile is universal, giving a good match to halos of all mass from those hosting dwarf galaxies to those hosting galaxy clusters; they suggested a fitting function:

$$\rho(r) = \rho_0 \left( \frac{r}{r_0} \right)^{-1} \left( 1 + \frac{r}{r_0} \right)^{-2} \quad (1)$$

that has become known as the ‘NFW’ profile.

While  $\Lambda$ CDM has performed remarkably well on large scales (e.g. Tegmark & Zaldarriaga 2002), the above prediction has long been at tension with observational data from dwarf and low surface brightness (LSB) galaxies. Flores & Primack (1994) and Moore (1994) were the first to show that fits to dwarf galaxy rotation curves favour a dark

matter density profile with a central constant density core:  $d \ln \rho / d \ln r|_{r=0} \equiv \alpha = 0$ , rather than a cusp. Since then, similar results have been reported for a wide range of gas rich dwarf and low surface brightness (LSB) galaxies (e.g. de Blok et al. 2001; McGaugh et al. 2001; de Blok et al. 2008; Hague & Wilkinson 2014); a result that is robust to known observational uncertainties and model systematics (e.g. Kuzio de Naray & Kaufmann 2011). This now long-standing discrepancy between theory and observation has become known as the ‘cusp-core’ problem (for a review see e.g. de Blok 2010).

One proposed solution to the cusp-core problem is to invoke stellar feedback, not modelled in the pure dark matter simulations discussed above. Navarro et al. (1996a) were the first to suggest that impulsive winds driven by supernovae could cause the dark matter halo to expand erasing the central cusp. However, Gnedin & Zhao (2002) demonstrated that, once the angular momentum barrier to gas collapse is taken into account, the maximum effect of a blow out is small. Read & Gilmore (2005) showed that this problem can be overcome if star formation proceeds in multiple bursts, gradually transforming a cusp to a core over several cycles of star formation. This mechanism appears to be what is at play in recent high resolution simulations of dwarf galaxies that model both the dark matter fluid and the baryons

\* E-mail: psteger@phys.ethz.ch

## 2 Steger

(Mashchenko et al. 2008; Governato et al. 2010; Pontzen & Governato 2012; Teyssier et al. 2013). (For an elegant analytic treatment of the effect see Pontzen & Governato 2012; and for a review see Pontzen & Governato 2014.) Teyssier et al. (2013) point to two key observational predictions of such a scenario: (i) star formation should be bursty with a duty cycle of  $\sim$  a dynamical time; and (ii) the stars should be collisionlessly heated along with the dark matter, producing vertically hot stellar discs even in isolated galaxies. Such predictions appear to be supported by the latest data (Leaman et al. 2012; Kauffmann 2014).

If baryons really do transform cusps to cores then this is bad news for constraining dark matter models. Exotic models that produce cores (e.g. Zavala et al. 2013) would become indistinguishable from ‘vanilla’ cold dark matter. This motivates pushing to ever smaller scales where the baryonic effects should diminish. Peñarrubia et al. (2012) have recently suggested that supernova feedback will no longer provide enough energy for cusp–core transformations below some critical stellar mass  $M_c \lesssim 10^6 M_\odot$ , though the precise value of  $M_c$  remains under investigation (e.g. Madau et al. 2014). The dwarf spheroidal galaxies (dSphs) that orbit the Milky Way and Andromeda straddle this critical stellar mass, making them prime targets for measuring their dark matter density profiles. They also have the added advantage that most of their gravitating mass is dark with typically negligible contributions from stars and/or gas. The most massive, Fornax, has some  $\sim 10^7 M_\odot$  in stars (Coleman & de Jong 2008), lying above  $M_c$ ; while the smallest, Segue 1 (if it is indeed a galaxy; Niederste-Ostholt et al. 2009; Martinez et al. 2011), has just  $\sim 1000 M_\odot$  (Belokurov et al. 2007), lying well below  $M_c$ .

Due to their proximity to their host galaxies, dSphs differ from the dwarf and LSB galaxies discussed above in that they are almost completely devoid of gas (e.g. Gatto et al. 2013). This presents a challenge because stellar orbits, unlike gas, can cross. If only radial velocities are available, this leads to a strong degeneracy between the dark matter density profile and the orbit distribution of the stars, typically parameterised by the velocity anisotropy (e.g. Wilkinson et al. 2002):

$$\beta \equiv 1 - \frac{\sigma_\theta^2 + \sigma_\phi^2}{2\sigma_r^2} \quad (2)$$

where  $\sigma_{r,\theta,\phi}$  are the velocity dispersions in spherical polar coordinate directions  $r$ ,  $\theta$  and  $\phi$ , respectively. (An anisotropy  $\beta = 0$  corresponds to isotropic velocity dispersions, while  $\beta = 1$  is pure radial and  $\beta = -\infty$  pure tangential.)

Battaglia et al. (2008) suggested that the above degeneracy could be broken if stars are split into multiple tracer populations by either metallicity or abundance. They performed such an analysis on the Sculptor dSph, weakly favouring a dark matter core. This analysis was further refined and applied to four dSphs by Walker & Peñarrubia (2011) (hereafter WP11). They also favoured cores but with higher statistical significance, supporting earlier timing arguments for UMi and Fornax (Kleyna et al. 2003; Goerdt et al. 2006; Read et al. 2006a; Cole et al. 2012). However, recently the literature has become divided on the issue of cusps or cores in the dSphs. An analysis of four dwarfs using a non-parametric Schwarzschild method (Schwarzschild 1979)

favoured a wide range of density profiles (Jardel & Gebhardt 2013a); simple distribution function models (Amorisco & Evans 2012) or ‘virial parameter’ models (Evans et al. 2011) support the findings of WP11; while single component Schwarzschild (Breddels & Helmi 2013) and higher order Jeans (Richardson & Fairbairn 2013) analyses conclude that the data are simply not sufficient to say one way or the other (see also the discussion in Breddels & Helmi 2014). Finally, Strigari et al. (2014) have recently pointed out that some of the key assumptions in the WP11 analysis may be violated if the density profile of one or more of the tracers is cuspy, leading to central cores being artificially favoured over cusps.

In light of the above discrepancies, in this paper we present a new non-parametric Jeans modelling tool: GRAVIMAGE. Our goal is to assume only equilibrium and spherical symmetry, allowing the model full freedom otherwise. We refer to this approach as non-parametric, though really what we mean is that we have a model with far more parameters than the available data constraints. Since this means that we are necessarily under-constrained, we are then forced to build model ensembles and explore parameter degeneracies. For this, we use the efficient MULTINEST technique (Feroz et al. 2009). We support multiple tracers that can be simultaneously modelled in a single underlying potential (each with its own free form velocity anisotropy); and we allow for velocity moments up to fourth order to be constrained. Similar such approaches have been attempted in the literature before. Ibata et al. (2011) (section 5.1) present a non-parametric spherical Jeans solver for modelling globular clusters; Jardel & Gebhardt (2013b) present a non-parametric Schwarzschild method; and Chakrabarty & Saha (2001) present a non-parametric method for modelling the black hole at the centre of our Galaxy. However, our new method is more general than these previous works. Ibata et al. (2011) – since they are modelling globular star clusters – assume that the mass distribution is known, fitting only for the mass to light ratio and  $\beta(r)$ ; Jardel & Gebhardt (2013b) present a method very different from that discussed here (and therefore complementary); however, they do allow for multiple populations, or test their methodology on mock data. Finally, Chakrabarty & Saha (2001) present an elegant non-parametric distribution function method but it relies on assuming an isotropic velocity distribution function ( $\beta = 0$ ). This may be reasonable for stars orbiting a supermassive black hole, but it likely a poor assumption when modelling dSphs. Using spherical mock data, we set out to determine what type and quality of data is sufficient to constrain the logarithmic cusp slope within dSphs. We also use triaxial mock data to test what happens when our method is pushed beyond its regime of validity. We will present applications of our method to real data in forthcoming papers.

This paper is organised as follows. In §2 we introduce the method. In §4 we test our method on mock data. Finally, in §5 we present our conclusions.

## 2 METHOD

### 2.1 Derivation of the key equations

To a very good approximation, stars in galaxies move as a collisionless fluid that obeys the collisionless Boltzmann equation (e.g. Binney & Tremaine 2008):

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \nabla_{\vec{x}} f \cdot \vec{v} - \nabla_{\vec{v}} f \cdot \nabla_{\vec{x}} \Phi = 0, \quad (3)$$

where  $f(\vec{x}, \vec{v})$  is the distribution function of stars in phase space;  $\vec{x}$  and  $\vec{v}$  are the position and velocity of the stars; and  $\Phi$  is the gravitational potential.

In spherical coordinates  $(r, \theta, \phi)$ , the collisionless Boltzmann equation becomes:

$$\frac{\partial f}{\partial t} + \dot{r} \frac{\partial f}{\partial r} + \dot{\theta} \frac{\partial f}{\partial \theta} + \dot{\phi} \frac{\partial f}{\partial \phi} + v_r \frac{\partial f}{\partial v_r} + v_\theta \frac{\partial f}{\partial v_\theta} + v_\phi \frac{\partial f}{\partial v_\phi} = 0 \quad (4)$$

where  $\dot{r} = v_r$ ,  $\dot{\theta} = v_\theta$  and  $\dot{\phi} = v_\phi/r \sin \theta$  are the velocities in spherical coordinates.

In principle, given the positions and velocities of many stars we could directly solve equation 4 for the force field  $\nabla_{\vec{x}} \Phi$ . In practice, this is impractical as it involves derivatives of  $f$  that is six dimensional and thus very poorly sampled. One way around this problem is to use instead velocity moments of equation 4 – the Jeans equations (e.g. Binney & Tremaine 2008). Assuming a non-rotating steady state ( $\partial f / \partial t = 0$ ,  $\langle v_r \rangle = \langle v_\theta \rangle = \langle v_\phi \rangle = 0$ ); spherical symmetry; and defining the tangential velocity dispersion  $\sigma_\phi^2 \equiv \sigma_\theta^2 = \sigma_t^2$ , the second order moment is given by:

$$\frac{1}{\nu} \frac{\partial}{\partial r} (\nu \sigma_r^2) + \frac{2\beta(r)\sigma_r^2}{r} = -\frac{\partial \Phi}{\partial r} = -\frac{GM(< r)}{r^2}, \quad (5)$$

where  $M(< r)$  is the total cumulative mass;  $\nu$  is the density of a set of tracer star particles moving in the potential  $\Phi(r)$ ;  $G = 6.67398 \cdot 10^{-11} \text{ m}^3/\text{kg s}^2$  is Newton's gravitational constant; and  $\beta$  is the velocity anisotropy as in equation 2.

Integrating both sides of equation 5 gives the radial velocity dispersion as function of radius  $r$ , and correspondingly the fourth order moment:

$$\begin{aligned} \sigma_r^2(r) &= \frac{1}{\nu(r)} \exp \left( -2 \int_{r_{\min}}^r \frac{\beta(s)}{s} ds \right) \cdot \\ &\quad \int_r^\infty \frac{GM(\tilde{r}r)\nu(\tilde{r})}{\tilde{r}^2} \exp \left( 2 \int_{r_{\min}}^{\tilde{r}} \frac{\beta(s)}{s} ds \right) d\tilde{r} \end{aligned} \quad (6)$$

and:

$$\begin{aligned} \langle v_r^4 \rangle(r) &= \frac{3}{\nu(r)} \exp \left( -2 \int_{r_{\min}}^r \frac{\beta'(s)}{s} ds \right) \cdot \\ &\quad \int_r^\infty \frac{GM(\tilde{r})\nu\sigma_r^2}{\tilde{r}^2} \exp \left( 2 \int_{r_{\min}}^{\tilde{r}} \frac{\beta'(s)}{s} ds \right) d\tilde{r} \end{aligned} \quad (7)$$

(Note that  $r_{\min}$  leads to an integration constant that cancels and so the choice of  $r_{\min}$  is arbitrary.)

Typically, only *projected* velocities are observable. Projecting equations 6 and 7 along the line of sight, we obtain:

$$\sigma_{\text{LOS}}^2(R) = \frac{2}{\Sigma(R)} \int_R^\infty \left( 1 - \beta \frac{R^2}{r^2} \right) \frac{\nu(r)\sigma_r^2(r)r}{\sqrt{r^2 - R^2}} dr, \quad (8)$$

where  $\Sigma(R)$  denotes the surface mass density at *projected* radius  $R$ .

From equation 6, it is clear that the velocity anisotropy  $\beta(r)$  trivially degenerates with the cumulative mass distribution  $M(< r)$  that we would like to measure. This is problematic because with only line of sight velocities  $\beta(r)$  is poorly constrained, unless many thousands of velocities are available (e.g. Wilkinson et al. 2002). This is one motivation for considering the higher order moment equations: these may break the  $M(< r) - \beta(r)$  degeneracy. We will explore this later on in this paper, but note that there is cause already for pessimism. For every higher order moment equation, there are new ‘anisotropy parameters’, in this case  $\beta'$  – the hierarchy of Jeans equations are not closed (e.g. Binney & Tremaine 2008). Since each new equation adds yet another new unknown, unless there are strong theoretical reasons for a relationship between  $\beta$ ,  $\beta'$  and similar, the higher order moments are unlikely to yield improved constraints (e.g. Richardson & Fairbairn 2013). More promising are the virial parameters discussed by (Richardson & Fairbairn 2014) that involve fourth order moments but are independent of  $\beta'$ . We will consider these in a separate publication.

### 2.2 The mass distribution

In the following, we present a non-parametric method for the solution of equation 8 for the total gravitating mass density  $\rho(r)$ , given a single or multiple tracer density profiles  $\nu_i(r)$  with corresponding line-of-sight velocity dispersions  $\sigma_{\text{LOS},i}(R)$ . We write the overall density profile  $\rho(r)$  as:

$$\rho(r) = \rho_{\text{DM}}(r) + \frac{M_*}{L} \cdot L_*(r) \quad (9)$$

where  $\rho_{\text{DM}}(r)$  is the dark matter contribution;  $L_*(r)$  is the visible light profile; and  $(M_*/L)$  is a constant mass-to-light ratio (the assumption that this is constant with radius could be relaxed if required by the data). As we will see for our mock data the tracer particles have negligible mass,  $M_*/L \approx 0$ .

The enclosed mass  $M(< r)$  then follows from the density via:

$$M(< r) = \int_0^r \rho(r) r^2 dr, \quad (10)$$

(Note that in principle our method can be generalized to investigate alternative gravity models if the acceleration  $GM(r)/r^2$  is replaced by  $-\partial\Phi/\partial r$  in equation 5.)

The dark matter density  $\rho_{\text{DM}}(r)$  is represented in terms of the logarithmic density slope:

$$n(r_j) = -d \ln \rho(r) / d \ln r |_{r=r_j} \quad (11)$$

for  $j \dots N_{\text{bin}}$  radial bins:  $1 \leq j \leq N_{\text{bin}}$ , as:

$$\rho(r) = \rho_{1/2} \cdot \exp \left[ - \int_{\ln r_{1/2}}^{\ln r} n(s) ds \right],$$

with the density at half-light radius  $\rho(r_{1/2}) = \rho_{1/2}$ , and  $n(r)$  interpolated linearly in between bin radii  $r_{\min,j} < r < r_{\max,j}$ . We prescribe three buffer bins  $n(r_j)$  for  $j \in \{N_{\text{bin}} +$

$1, N_{\text{bin}} + 2, N_{\text{bin}} + 3\}$  outside of the range where data is given to enable sensible extrapolations towards high radii, and two additional slopes  $n_0 < 3, n_\infty > 3$  for the asymptotic density slopes towards  $r = 0$  and  $r = \infty$ , which are reached at half the smallest radius and  $r_\infty = 10r_{\max}$ .

If the  $n(r)$  are allowed to change freely between a minimum value of 0 and a maximum value  $n_{\max}$ , a tendency towards following pattern emerges: If at radius  $n(r_i)$  is chosen too small, and thus the mass  $M(r_i)$  is too high, the next bin corrects for this with a change towards higher  $n(r_{i+1})$ . The next bin after that corrects that one, and the whole  $n(r)$  profile wiggles.

To counteract this behavior, we introduce as a prior the maximal change in  $n(r)$  that is allowed to happen, by sampling – instead of  $n(r)$  directly – its derivative  $dn(r)/d\ln(r)$  with a flat prior in the range  $[-\alpha, \alpha]$ , where  $\alpha$  is chosen such that  $n(r)$  can sweep from 0 to  $n_{\max}$  over the radii where we have data.

To enforce faster convergence we prefer smoother  $n(r)$  profiles by sampling  $dn/d\log r$  from a normal distribution with width  $w > 0$ , not a uniform one in the range  $[-w, w]$ .

### 2.3 The tracer density profile

The tracer density is represented similar to  $\rho_{\text{DM}}$  for each tracer population  $i$ . The 3D light distribution for each tracer  $\nu_i(r)$  is projected to give the surface density:

$$\Sigma_i(R) = 2 \int_0^\infty \frac{\nu_i(r)r}{\sqrt{r^2 - R^2}} dr \quad (12)$$

which is then compared with the observational data (see §2.5).

### 2.4 The velocity anisotropy

The velocity anisotropy  $\beta(r)$  is allowed to vary broadly in the interval  $[-\infty, 1]$  by sampling a modified, symmetric  $\beta^*$  (e.g. Read et al. 2006b):

$$\beta^* = \frac{\sigma_r^2 - \sigma_t^2}{\sigma_r^2 + \sigma_t^2} = \frac{\beta}{2 - \beta} \in [-1, 1] \quad (13)$$

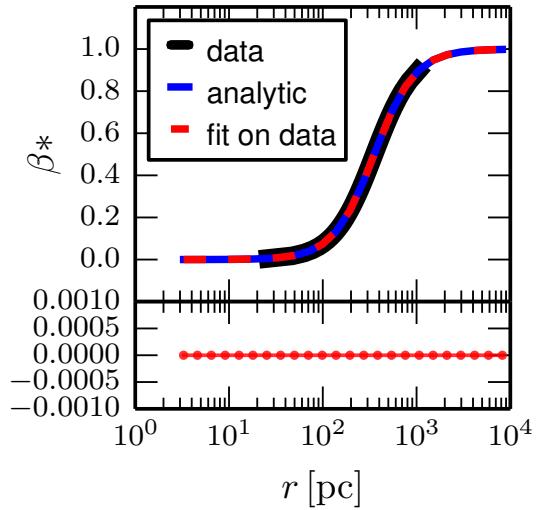
with the following function:

$$\beta^*(r) = \frac{a_0 - a_\infty}{1 + \kappa \exp(\alpha \ln(r/r_s))} + a_\infty \quad (14)$$

$$\kappa = \frac{a_0 - a_\infty}{\beta^*(r_s) - a_\infty} - 1 \quad (15)$$

where  $a_0, a_\infty \in [-1, 1]$  are the inner and outer asymptotic values for  $\beta^*(r \rightarrow 0, \infty)$ ,  $r_s$  is the scale radius of  $a_0 \rightarrow a_\infty$  transition with a  $\beta^*$  at that point of  $\beta^*(r_s)$ ,  $\alpha$  the speed of the transition.

This approach allows us to sample many qualitatively different models – isotropic, radially biased, tangentially biased, and any transitions between those – with very few parameters. In Figure 1, we compare our functional form in equation 15 with a typical Osipkov-Merritt profile taken from our mock data (see §3). With our 5 parameters, we obtain an excellent recovery of the input profile, while being able also to capture constant, and both rising and falling



**Figure 1.** Fitting of the analytic  $\beta$  profile using the form in equation ???. The upper plot shows an analytic Osipkov-Merritt anisotropy profile as dashed, black curve, and our fitted profile. Shown below is the difference between analytic and fitted profile. We see that the analytic profile is fit well over the range of input data, and extrapolated agreeably well to higher radii.

models. The fact that an asymptotic value is reached at high radii allows us to recover the correct  $\int \beta(s)/s$ .

The anisotropy parameters for the fourth order moments ( $\beta'$ ) are assumed to be constant over all radii. A further refinement of our method will include binwise changing  $\beta'$  and the use virial parameters as introduced in Richardson & Fairbairn (2014).

### 2.5 Comparison with data

Given the above functional forms,  $\sigma_{\text{LOS},i}(r)$  is calculated from  $\rho(r)$ ,  $\nu_i(r)$ , and  $\beta_i(r)$  according to equation 8. This is done numerically, involving three integrations, which are performed with polynomial extrapolations of the integrands up to infinity, such that missing contributions from  $r > r_{\max}$  do not lead to an artificial falloff of  $\sigma_{\text{LOS}}$ .

We check our integration routine against an analytically known Hernquist profile, by setting the input parameters to a simple Hernquist profile, in our representation. After calculating all profiles and integrals on 100 points in radius – the same way as is done in all runs – we get a good agreement with the analytic  $\sigma_{\text{LOS}}$  profile.

The last step involves comparison of the projected surface density  $\Sigma_i(r)$  – calculated from the 3D tracer density  $\nu_i(r)$  – as well as  $\sigma_{\text{LOS},i}(r)$ , to the respective 2D data profiles for the tracer populations. We use a likelihood based on the overall goodness of fit:

$$\chi^2 = \sum_{i=1}^N \chi_{\Sigma,i}^2 + \chi_{\sigma,i}^2, \quad (16)$$

$$\chi_{\Sigma,i}^2 = \sum_{j=1}^{N_{\text{bin}}} \left( \frac{\Sigma_{\text{data},i}(r_j) - \Sigma_{\text{model},i}(r_j)}{\varepsilon_{\Sigma}(r_j)} \right)^2, \quad (17)$$

with error  $\varepsilon_{\Sigma}(r_j)$  on the data  $\Sigma_{\text{data},i}(r_j)$ . Analogous expres-

sions hold for  $\chi_{\sigma,i}^2$ . In the absence of a measured  $\beta_i(r)$ , we set  $\chi_{\beta,i}^2 = 0$ .

## 2.6 Priors

We investigate two theoretically motivated priors on the dark matter density profile:

(i) A regularization prior on  $n(r)$ :  $|n(r_{j+1}) - n(r_j)| / (\ln(r_{j+1}) - \ln(r_j)) < f \cdot \frac{\ln(r_{\max})}{N_{\text{bin}}}$  with a tuning factor  $f$  of order unity. This is designed to prevent  $n(r)$  fluctuating wildly from bin to bin.

(ii) An absolute prior on  $n(r)$ :  $n(r < r_{1/2}) < 2.0$ . This corresponds to a prior that the stars in the galaxy lie deep within their host dark matter halo (since at large radii, the dark matter halo is expected to have a slope of  $n = 3$ ).

Both priors can be switched; we explore their effect in §??.

Otherwise, we use very weak flat priors on all other parameters:

- 1)  $0 \leq \rho_{\text{DM},j} \leq 5$
- 2)  $0 \leq \nu_{i,j} \leq 5$
- 3)  $0 \leq b_{i,j} \leq 10$

where  $\rho_{\text{DM},j}$  and  $\nu_{i,j}$  are the dark matter density and tracer density (for population  $i$ ) in bin  $j$ ; and  $b_{i,j}$  is the polynomial coefficient  $b_j$  for tracer population  $i$ .

## 2.7 Parameter Space Sampling

Given the freedom in our function forms for the matter density; tracer densities; and the velocity anisotropy profiles of these, we are left with a large number of free parameters:

$$N_{\text{dim}} = N_{\text{bin}} + N_{\text{pop}} \cdot (N_{\text{bin}} + N_{\text{beta}}) \quad (18)$$

with significantly fewer observational constraints. As discussed in §1, we refer to this as ‘non-parametric’ mass modelling.

To efficiently sample the above high dimensional parameter space, we use the MULTINEST code (Feroz & Hobson 2008), (Feroz et al. 2009), (Feroz et al. 2013), see also <http://ccpforge.cse.rl.ac.uk/gf/project/multinest/>. This is a Bayesian nested sampling algorithm to generate posterior samples from non-trivial distributions in high dimensions. MULTINEST samples the  $n$ -dimensional hypercube  $\kappa = [0, 1]^{N_{\text{dim}}}$ , which needs to be translated into physical prior distributions for each of the parameter profiles.

## 3 MOCK DATA

We apply our method to a set of mock data, consisting of multiple tracers moving in spherical or triaxial potentials with radial or tangential velocity anisotropy. All mock data are available on the GAIA CHALLENGE wiki site<sup>1</sup>. The first set of mocks have cusped or cored dark matter density profiles with radial anisotropy – the ‘Walker’ mocks (WP11), with tracer density:

<sup>1</sup> <http://astrowiki.ph.surrey.ac.uk/dokuwiki/>

**Table 1.** Parameters of the 1-population Gaia challenge mock data.

ID	geometry	$\gamma_{\text{DM}}$	$\gamma_*$	$r_*$
1pop core	sphere	0	0.1	250 pc
1pop cusp	sphere	1	0.1	250 pc

**Table 2.** Parameters of the 2-population Walker mock data.

ID	geometry	$\gamma_{\text{DM}}$	$r_{1/2,\text{DM}}$
2pop core	sphere	0	1000 pc
2pop cusp	sphere	1	1000 pc

$$\nu_*(r) = \nu_0 \left( \frac{r}{r_*} \right)^{-\gamma_*} \left[ 1 + \left( \frac{r}{r_*} \right)^{\alpha_*} \right]^{(\gamma_* - \beta_*)/\alpha_*} \quad (19)$$

inside dark matter halos of the form:

$$\rho_{\text{DM}} = \rho_0 \left( \frac{r}{r_{\text{DM}}} \right)^{-\gamma_{\text{DM}}} \left[ 1 + \left( \frac{r}{r_{\text{DM}}} \right)^{\alpha_{\text{DM}}} \right]^{(\gamma_{\text{DM}} - \beta_{\text{DM}})/\alpha_{\text{DM}}} \quad (20)$$

with scale radii  $r_*$ ,  $r_{\text{DM}}$ ; central slopes of  $\gamma_*$ ,  $\gamma_{\text{DM}}$ ; transition parameters  $\beta_*$ ,  $\beta_{\text{DM}}$ ; and outer slopes  $\alpha_*$ ,  $\alpha_{\text{DM}}$ .

The anisotropy follows the functional form of Osipkov (1979) and Merritt (1985):

$$\beta(r) = 1 - \frac{\sigma_\theta^2}{\sigma_r^2} = \frac{r^2}{r^2 + r_a^2}. \quad (21)$$

with scale radius  $r_a$ , turning over from nearly isotropic at  $r \rightarrow 0$  to radially biased at  $r_* = r_a$ .

Of these distributions, finite samplings are taken, giving our first set of mock data, table 3, and then converted to mock observational data including observational parameters like spectral indices, systemic velocities, proper motions, and binary motions. The full suite of mock data is much larger than that used here. Our particular subset is given in table 3.

In addition to the above Walker mocks, we add also a mock dwarf with cusped triaxial model viewed along three different projection angles: down the minor axis; the major axis; and the intermediate axis. All models are summarised in Table 3.

**Table 3.** Parameters of the 1-population triaxial mock data.

ID	geometry	$\gamma_{\text{DM}}$
Triax01	triaxial, intermediate axis	0
Triax02	triaxial, along x	0
Triax03	triaxial, along y	0
Triax04	triaxial, along z	0
Triax05	triaxial, intermediate axis	1
Triax06	triaxial, along x	1
Triax07	triaxial, along y	1
Triax08	triaxial, along z	1

## 4 RESULTS

### 4.1 Single tracer population

We first apply our method to DM halos hosting a single population of tracer stars. We consider both cored and cusped models. As default, we assume that we have 10,000 tracer stars. We consider poorer sampling in §???. The cusped model has:  $\gamma_{DM} = 1$ , stellar central density slope  $\gamma_{*,1} = 0.1, \gamma_{*,2} = 0.1$ ; stellar turnover slopes  $\beta_{*,1} = \beta_{*,2} = 5$ ; stellar characteristic radii  $r_{*,1} = 100\text{pc}, r_{*,2} = 500\text{pc}$ ; and anisotropy scale radii  $r_{a,1} = r_{a,2} = 1.0$ . All other parameters are as in table 3. Our results for the recovered mass distributions are shown in Figure 2; Figure 3 shows a comparison of the model tracer surface density  $\Sigma(r)$  and projected velocity dispersion  $\sigma_{\text{LOS}}(r)$  with mock data for the cusped case. Notice that in all cases, we successfully recover the input model within our quoted uncertainties.

### 4.2 Two tracer populations

In this section, mock dwarfs with a model where two populations of tracer particles are accounted for are analyzed. This is done in the following manner. Each particle in the mock dataset has a number identifying it as a member of population 1, 2, or background. For fig. 4, we used this information directly.

### 4.3 Triaxial mock data

To test the dependency of GRAVIMAGE on the assumption of spherical symmetry, we employ it on slightly triaxial mock dwarf galaxies.

The models were generated with the Made2Measure algorithm of Dehnen (2009) and are tailored to follow a similar profile to the profiles specified above for the dwarf galaxies. They show a density profile of

$$\rho(r) = \frac{\rho_S}{\left(\frac{r}{r_S}\right)^\gamma \left(1 + \left(\frac{r}{r_S}\right)^{1/\alpha}\right)^{\alpha(\beta-\gamma)}} \quad (22)$$

with radius  $r$ , scale radius  $r_S = 1.5\text{kpc}$ ,  $\alpha = 1, \beta = 4$ . For the cusped profiles we have an inner logarithmic slope of  $\gamma = 1$ , scale density  $\rho_S = 5.522 \cdot 10^7 M_\odot/\text{kpc}^3$ , and  $M_{\text{tot}} = 1.171 \cdot 10^9 M_\odot$ , while for the cored one we have  $\gamma = 0.23$ ,  $\rho_S = 1.177 \cdot 10^8 M_\odot$ ,  $M_{\text{tot}} = 1.802 \cdot 10^9 M_\odot$ . The axis ratios are  $b/a = 0.8$  and  $c/a = 0.6$ . The stars have negligible mass and follow the same functional form in the density profile as dark matter, with  $\alpha = 0.34, \beta = 5.92, \gamma = 0.23, r_S = 0.81\text{kpc}$ .

The velocity anisotropy of the stellar part is calculated via

$$\beta(r) = \frac{r_{s,\beta}^\eta \beta_0 + r^\eta \beta_\infty}{r^\eta + r_{s,\beta}^\eta}, \quad (23)$$

with  $r_{s,\beta} = 0.81\text{kpc}$ ,  $\beta_0 = 0$ ,  $\beta_\infty = 0.5$  and  $\eta = 0.5$ , going from isotropic to radially anisotropic with increasing radius.

The retrieved density profile (fig. 5) recaptures the density inside the half-light radius, but constantly overestimates

it at  $r > r_{1/2}$ . This is partly due to projection effects, as the underlying density profile in blue is calculated for spherically averaged density decrease.

### 4.4 Data quality

How many tracer stars are needed to determine the overall density profile reliably? To address this question, we performed three runs with a restricted set of tracer particles. In the first,  $10^3$  particles were chosen out of the  $10^6$  simulated particles. With  $10^4$  particles, the confidence intervals shrink. These  $10^4$  particles are then split into two populations of each  $5 \cdot 10^3$  particles, with different scalelengths of  $r_S$  and  $r_S/10$ . Most of the second population particles are inside the first two bins, so the overall convergence is not visibly affected above the third bin. However, the models are better constrained around the scalelengths of both tracer populations. This is expected from Walker & Peñarrubia (2011), as any velocity anisotropy sampling yields the same mass constraint there.

## 5 CONCLUSIONS

The new non-parametric method samples the profiles of the overall density bin-wise, and was shown to reconstruct the density of diverse mock data.

The number of bins has to be chosen such that the numerical errors introduced from integrating finite datapoints do not contribute significantly to  $\chi^2$ . Especially the projection of the 3D model density profiles for the tracer densities onto the 2D plane of the sky introduced was erroneous when done with  $N_{\text{bin}} \leq 7$ . Spacing the bins logarithmically instead of setting equal number of tracers per bin helps to circumvent this problem, as does increasing the number of bins. With both approaches, the integral in eq. ?? includes more of the high density region in the center.

## 6 ACKNOWLEDGEMENTS

JIR would like to acknowledge support from SNF grant PP00P2\_128540/1.

## APPENDIX A: APPENDIX

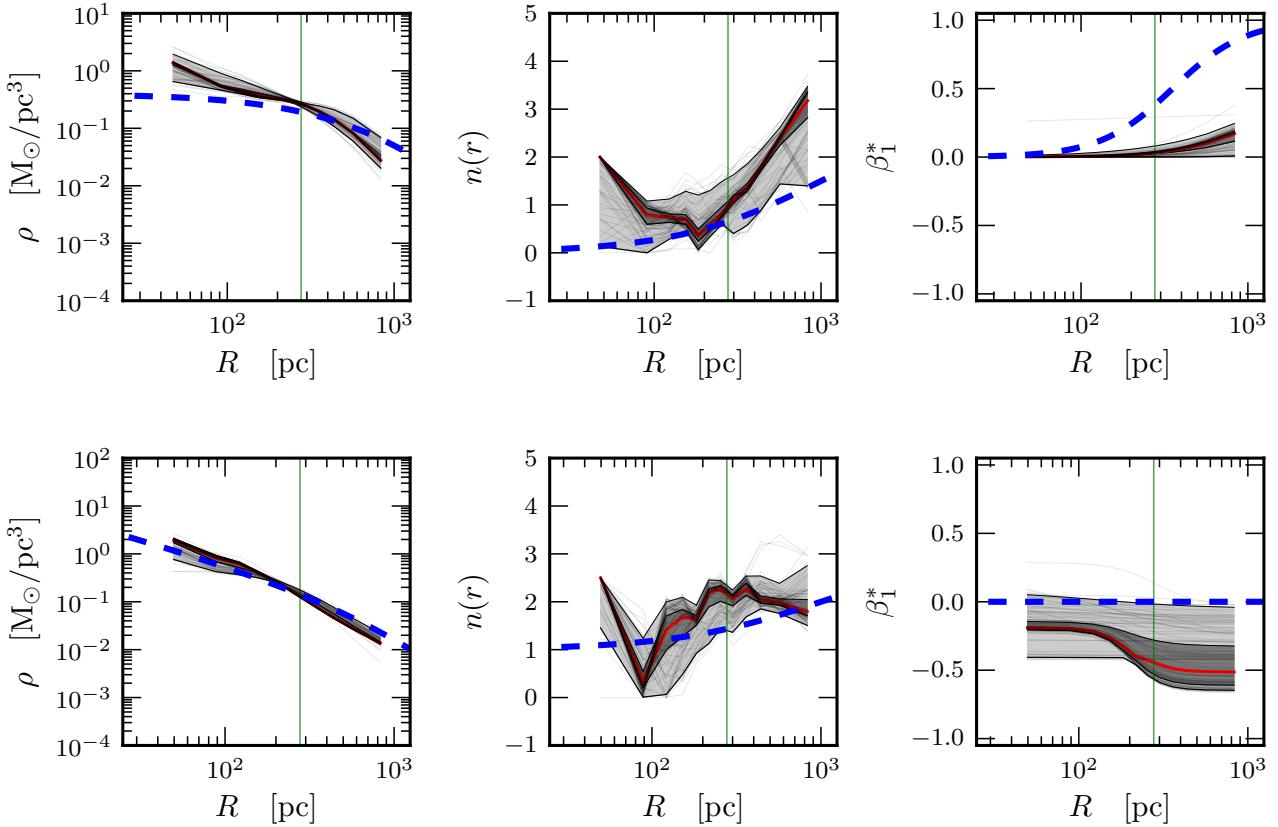
### 1 Effects of priors on $n'(r)$

### 2 Convergence of the MultiNest model ensemble

We check the convergence of the MCMC twofold: first, the range of density profiles swept after 5k, 50k and to the end of our run is increased from 5k to 50k, but stays approximately constant for another 16k runs, thus giving us confidence that we found and fully explored the valid regions of phase space in density. See fig. 2.

## 3 The choice of binning

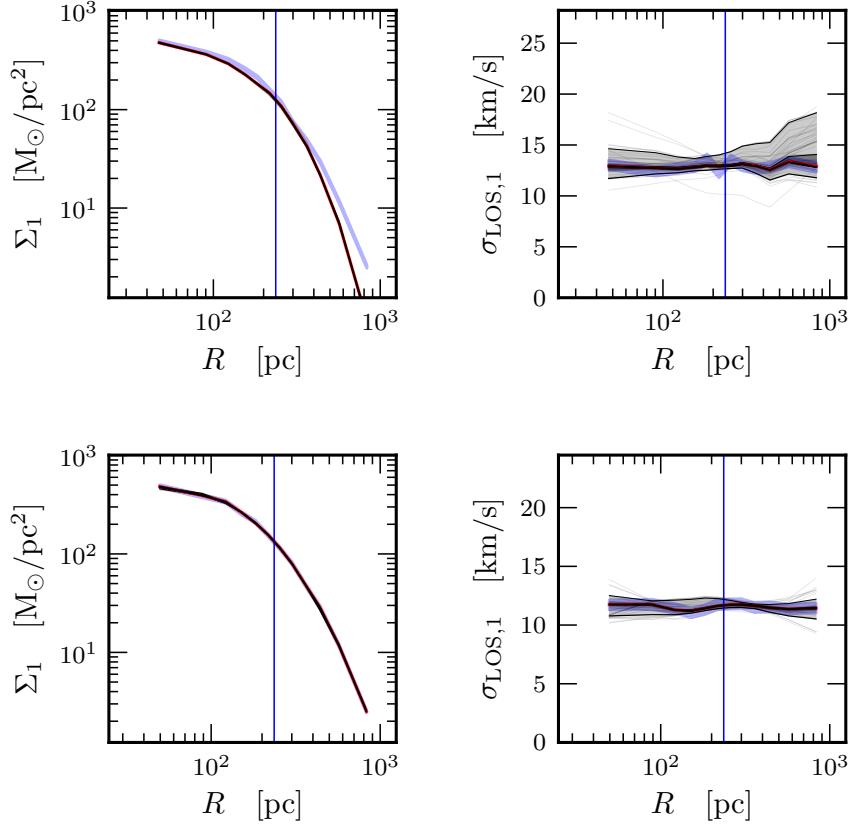
(TODO: Consider the effect of varying the number of particles per bin for a single model.) (TODO: explore different binning and lower sampling in an appendix)



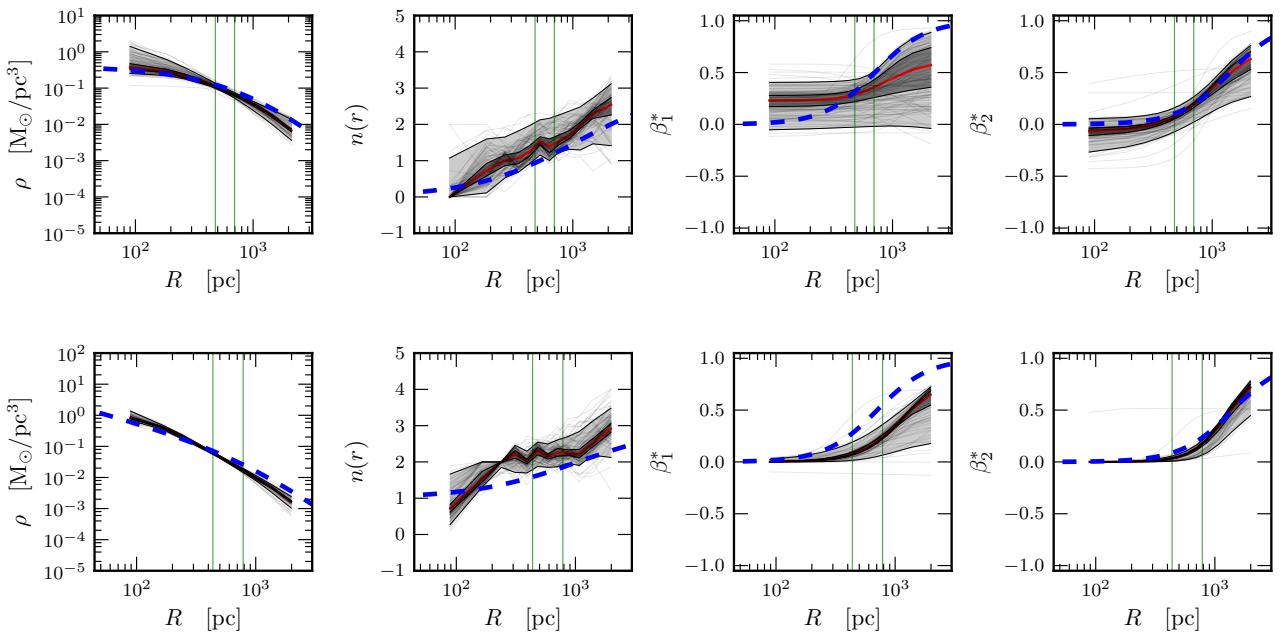
**Figure 2.** Reconstructed density for 1pop core (top), and 1pop cusp (bottom); logarithmic density slope; and velocity anisotropy profile, using all tracer stars, on the order of 3000. The input model profile is marked by the blue dashed line; the red line and grey contours show the median, 68% and 95% confidence intervals for our chains, respectively; the vertical green line marks the 3D half-light radius of the stars; and the gray lines show a sub-set of individual models. The full ensemble shown samples of accepted models in total.

## REFERENCES

- Amorisco N. C., Evans N. W., 2012, MNRAS, 419, 184  
 Battaglia G., Helmi A., Tolstoy E., Irwin M., Hill V., Jablonka P., 2008, ApJ, 681, L13  
 Belokurov V. et al., 2007, ApJ, 654, 897  
 Binney J., Tremaine S., 2008, Galactic Dynamics: Second Edition. Princeton University Press  
 Breddels M. A., Helmi A., 2013, A&A, 558, A35  
 Breddels M. A., Helmi A., 2014, ArXiv e-prints  
 Chakrabarty D., Saha P., 2001, AJ, 122, 232  
 Cole D. R., Dehnen W., Read J. I., Wilkinson M. I., 2012, MNRAS, 426, 601  
 Coleman M. G., de Jong J. T. A., 2008, ApJ, 685, 933  
 de Blok W. J. G., 2010, Advances in Astronomy, 2010  
 de Blok W. J. G., McGaugh S. S., Rubin V. C., 2001, AJ, 122, 2396  
 de Blok W. J. G., Walter F., Brinks E., Trachternach C., Oh S.-H., Kennicutt, Jr. R. C., 2008, AJ, 136, 2648  
 Dehnen W., 2009, MNRAS, 395, 1079  
 Dubinski J., Carlberg R. G., 1991, ApJ, 378, 496  
 Evans N. W., An J., Deason A. J., 2011, ApJ, 730, L26  
 Feroz F., Hobson M. P., 2008, MNRAS, 384, 449  
 Feroz F., Hobson M. P., Bridges M., 2009, MNRAS, 398, 1601  
 Feroz F., Hobson M. P., Cameron E., Pettitt A. N., 2013, ArXiv e-prints  
 Flores R. A., Primack J. R., 1994, ApJ, 427, L1  
 Gatto A., Fraternali F., Read J. I., Marinacci F., Lux H., Walch S., 2013, MNRAS, 433, 2749  
 Gnedin O. Y., Zhao H., 2002, MNRAS, 333, 299  
 Goerdt T., Moore B., Read J. I., Stadel J., Zemp M., 2006, MNRAS, 368, 1073  
 Governato F. et al., 2010, Nature, 463, 203  
 Hague P. R., Wilkinson M. I., 2014, ArXiv e-prints  
 Ibata R., Sollima A., Nipoti C., Bellazzini M., Chapman S. C., Dalessandro E., 2011, ApJ, 738, 186  
 Jardel J. R., Gebhardt K., 2013a, ApJ, 775, L30  
 Jardel J. R., Gebhardt K., 2013b, ApJ, 775, L30  
 Kauffmann G., 2014, MNRAS, 441, 2717  
 Kleyna J. T., Wilkinson M. I., Gilmore G., Evans N. W., 2003, ApJ, 588, L21  
 Kuzio de Naray R., Kauffmann T., 2011, MNRAS, 414, 3617  
 Leaman R. et al., 2012, ApJ, 750, 33  
 Madau P., Shen S., Governato F., 2014, ApJ, 789, L17  
 Martinez G. D., Minor Q. E., Bullock J., Kaplinghat M., Simon J. D., Geha M., 2011, ApJ, 738, 55  
 Mashchenko S., Wadsley J., Couchman H. M. P., 2008, Science, 319, 174  
 McGaugh S. S., Rubin V. C., de Blok W. J. G., 2001, AJ,



**Figure 3.** Tracer surface density profile  $\Sigma(r)$ , and projected velocity dispersion profile  $\sigma_{\text{LOS}}(r)$  (right) for the stars in the single component cusped profile of Figure 2. The vertical green lines show the 3D projected half-light radius.

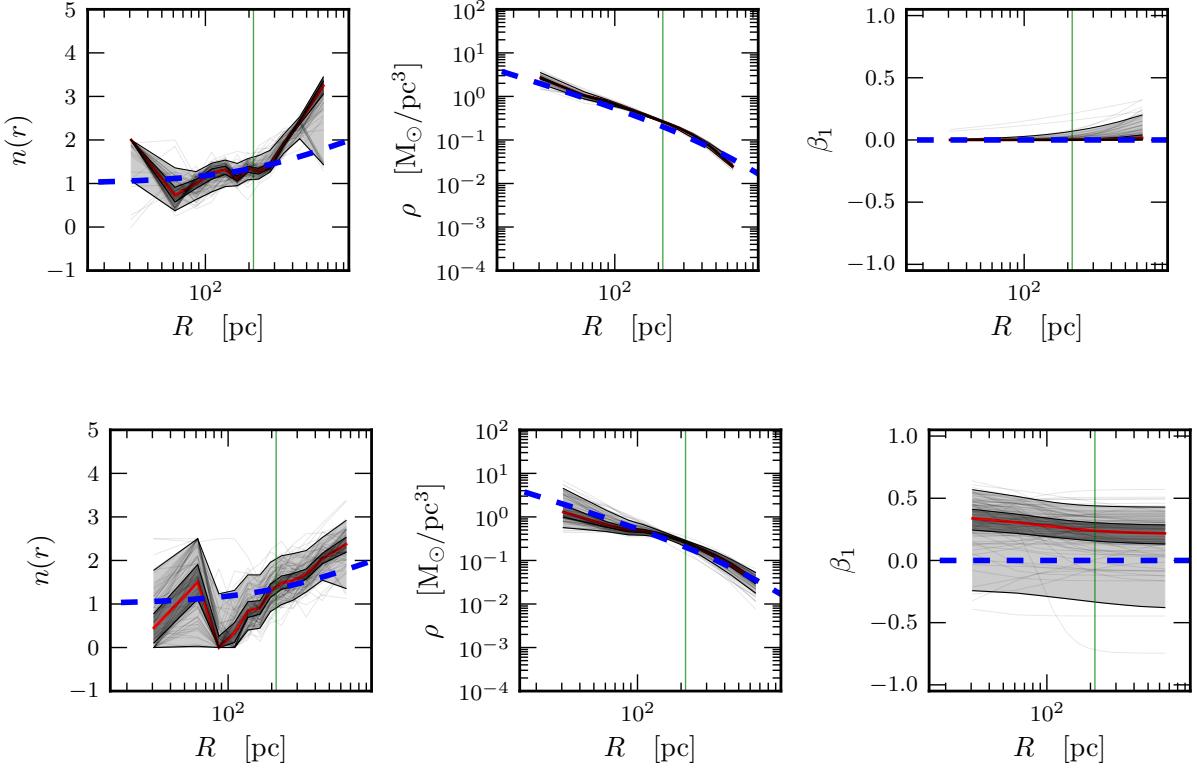


**Figure 4.** Reconstructed density and density slope for the 2pop core model (top) and 2pop cusp model (bottom), with two tracer populations with half light radii marked by the vertical green lines. Lines are as in Figure 2. We incorporated a  $\beta^*(r) \geq 0$  prior here to speed up convergence.

**Figure 5.** Density profile of a cusped (**TODO: triaxial**) mock dwarf, for which the line of sight is inclined with 45 degrees with respect to all axes. The vertical line indicates the half-light radius at 640pc.

**Figure 6.** (**TODO: Hernquist**) profile found by MCMC model (red) for  $10^3$ ,  $10^4$  and two times  $5 \cdot 10^3$  tracer particles. The black curve shows the enclosed mass derived from the theoretical model.

- 122, 2381  
Merritt D., 1985, MNRAS, 214, 25P  
Moore B., 1994, Nature, 370, 629  
Navarro J. F., Eke V. R., Frenk C. S., 1996a, MNRAS, 283, L72  
Navarro J. F., Frenk C. S., White S. D. M., 1996b, ApJ, 462, 563  
Niederste-Ostholt M., Belokurov V., Evans N. W., Gilmore G., Wyse R. F. G., Norris J. E., 2009, MNRAS, 398, 1771  
Osipkov L. P., 1979, Pisma v Astronomicheskii Zhurnal, 5, 77  
Peñarrubia J., Pontzen A., Walker M. G., Koposov S. E., 2012, ApJ, 759, L42  
Pontzen A., Governato F., 2012, MNRAS, 421, 3464  
Pontzen A., Governato F., 2014, Nature, 506, 171  
Read J. I., Gilmore G., 2005, MNRAS, 356, 107  
Read J. I., Pontzen A. P., Viel M., 2006a, MNRAS, 371, 885  
Read J. I., Wilkinson M. I., Evans N. W., Gilmore G., Kleyna J. T., 2006b, MNRAS, 367, 387  
Richardson T., Fairbairn M., 2013, MNRAS, 432, 3361  
Richardson T., Fairbairn M., 2014, MNRAS, 441, 1584  
Schwarzschild M., 1979, ApJ, 232, 236  
Strigari L. E., Frenk C. S., White S. D. M., 2014, ArXiv e-prints  
Tegmark M., Zaldarriaga M., 2002, Phys. Rev. D, 66, 103508  
Teyssier R., Pontzen A., Dubois Y., Read J. I., 2013, MNRAS, 429, 3068  
Walker M. G., Peñarrubia J., 2011, ApJ, 742, 20  
White S. D. M., Rees M. J., 1978, MNRAS, 183, 341  
Wilkinson M. I., Kleyna J., Evans N. W., Gilmore G., 2002, MNRAS, 330, 778  
Zavala J., Vogelsberger M., Walker M. G., 2013, MNRAS, 431, L20



**Figure 1.** Influence of  $dn(r)/d \log r$  prior. The top panels show  $n(r)$ ,  $\rho(r)$ , and  $\beta^*(r)$  for Gaia07 with a moderate prior of  $|dn(r)/d \log r| < 1.5/(8/N_{\text{ipol}})$ , the lower panels show the same profiles with two times this value. We see that a tighter prior on  $n(r)$  yields tighter constraints on  $\beta^*$ .

**Figure 2.** (TODO: Convergence) of the density profile after (3k,30k,300k) iterations (left to right) for Gaia01.