

# Dark matter density profiles in dwarf galaxies: linking Jeans modelling systematics and observation

Laura J. Chang<sup>1</sup>★ and Lina Necib<sup>2,3,4</sup>

<sup>1</sup>*Department of Physics, Princeton University, Princeton, NJ 08544, USA*

<sup>2</sup>*Walter Burke Institute for Theoretical Physics, California Institute of Technology, Pasadena, CA 91125, USA*

<sup>3</sup>*Center for Cosmology, Department of Physics and Astronomy, University of California, Irvine, CA 92697, USA*

<sup>4</sup>*Observatories of the Carnegie Institution for Science, 813 Santa Barbara St, Pasadena, CA 91101, USA*

Accepted 2021 August 20. Received 2021 August 13; in original form 2020 September 28

## ABSTRACT

The distribution of dark matter in dwarf galaxies can have important implications on our understanding of galaxy formation as well as the particle physics properties of dark matter. However, accurately characterizing the dark matter content of dwarf galaxies is challenging due to limited data and complex dynamics that are difficult to accurately model. In this paper, we apply spherical Jeans modelling to simulated stellar kinematic data of spherical, isotropic dwarf galaxies with the goal of identifying the future observational directions that can improve the accuracy of the inferred dark matter distributions in the Milky Way dwarf galaxies. We explore how the dark matter inference is affected by the location and number of observed stars as well as the line-of-sight velocity measurement errors. We use mock observation to demonstrate the difficulty in constraining the inner core/cusp of the dark matter distribution with data sets of fewer than 10 000 stars. We also demonstrate the need for additional measurements to make robust estimates of the expected dark matter annihilation signal strength. For the purpose of deriving robust indirect detection constraints, we identify Ursa Major II, Ursa Minor, and Draco as the systems that would most benefit from additional stars being observed.

**Key words:** galaxies: dwarf – (*cosmology*:) dark matter – gamma-rays: galaxies.

## 1 INTRODUCTION

The standard  $\Lambda$ CDM model, consisting of the cosmological constant  $\Lambda$  and cold dark matter (CDM), has had remarkable success at predicting physics on large scales, e.g. the cosmic microwave background (Aghanim et al. 2020) and the large-scale distribution of matter in the Universe (Tegmark & Zaldarriaga 2002, 2009; Hlozek et al. 2012), but faces several small-scale challenges (Bullock & Boylan-Kolchin 2017). Among these challenges is the ‘core-cusp problem’ (Flores & Primack 1994; Moore 1994) –  $\Lambda$ CDM predicts that, in the absence of baryonic physics, dark matter (DM) haloes universally follow a Navarro–Frenk–White (NFW) density profile (Navarro, Frenk & White 1997), which steeply rises as  $\rho \propto r^{-1}$  towards central regions. However, a number of measurements of rotation curves and stellar dynamics have suggested that the DM distribution in the centres of dwarf galaxies may be more consistent with having a constant-density core (e.g. Flores & Primack 1994; Moore 1994; Salucci & Burkert 2000; Swaters et al. 2003; Spekkens, Giovanelli & Haynes 2005; Walter et al. 2008; Oh et al. 2011, 2015). In this paper, we apply spherical Jeans modelling to individual stars in simulated dwarf galaxies to characterize the observational regimes in which the method can robustly distinguish a cored halo from a cuspy one.

If the DM haloes of dwarf galaxies truly are cored, one potential way to explain the apparent discrepancy is through baryonic physics. During baryonic contraction, the central density of a galaxy increases due to the infall of dissipative baryons, deepening the potential well and dragging DM into the central region, which leads to the formation of a DM core – this happens primarily in Milky Way-sized galaxies (Blumenthal et al. 1986). On smaller scales, stellar feedback can lead to core formation due to the ejection of baryons (Navarro, Eke & Frenk 1996; Read & Gilmore 2005; Mashchenko, Couchman & Wadsley 2006; Pontzen & Governato 2012), as can dynamical friction arising from baryonic clumps interacting with the DM (see e.g. the reviews Del Popolo & Pace 2016; Del Popolo & Le Delliou 2017, and references within).

While there is qualitative agreement in the simulation literature surrounding the formation of cores in dwarf galaxy-sized DM haloes, there is considerable scatter in the quantitative results from various works. Recent studies of hydrodynamic simulations have shown that lower mass dwarfs ( $M_* \lesssim 10^6 M_\odot$ ) have cuspy DM haloes, while efficient core formation from stellar feedback turns on around  $M_* \sim 10^9 M_\odot$ ; for galaxies slightly more massive than the Milky Way, the DM halo reverts back to a cuspy distribution (see e.g. Di Cintio et al. 2014; Tollet et al. 2016; Lazar et al. 2020). Read, Agertz & Collins (2016) correlated the presence of cores to an active stellar formation history in isolated simulated dwarf galaxies. Similarly, simulations with a lower density threshold for star formation, for example Auriga (Grand et al. 2017) and APOSTLE (Sawala et al. 2016), find that cores do not form at dwarf galaxy sizes (Bose

\* E-mail: ljchang@alumni.princeton.edu

et al. 2019). Core formation thus depends on the baryonic feedback model, and while present observations are inconsistent with low star formation thresholds (Benítez-Llambay et al. 2019; Dutton et al. 2019), reliable observational evidence for cusps or cusps in dwarf galaxies has important implications for understanding stellar feedback and galaxy formation.

A different approach to resolving the core-cusp problem is to modify the particle model of DM itself – for example, models of self-interacting dark matter (SIDM) notably predict the formation of central cores in the DM density profiles of low-mass galaxies (Spergel & Steinhardt 2000). There has been extensive work in the literature studying halo formation in SIDM (e.g. Tulin & Yu 2018; Despali et al. 2019; Fitts et al. 2019; Robles et al. 2019). In addition to SIDM, other theories of DM can also predict different halo properties from the  $\Lambda$ CDM prediction, e.g. theories of dissipative DM have been shown to lead to the formation of haloes with inner density profiles that are more steeply cusped than NFW haloes (Shen, Hopkins & Necib 2021). The inner profiles of dwarf galaxy DM haloes can therefore encode information about the particle physics that governs the DM.

Whether the Milky Way dwarf galaxies truly all reside in cored or cuspy haloes, or there is a large scatter in the inner density profile shapes, there would be important consequences for our understanding of the underlying baryonic and DM physics. At present, there is a lack of consensus in the dwarf galaxy literature on whether the stellar data favour cuspy or cored DM distributions. One specific example is the case of Sculptor, one of the more extensively analysed dwarf galaxies in the mass modelling literature. Sculptor has been observed to have two chemo-dynamically distinct subpopulations of stars with different half-light radii, which can be leveraged to constrain the DM density at two different radii. Battaglia et al. (2008) applied separate Jeans analyses to the two stellar components and found that either a cored halo or an NFW halo were statistically consistent with their data. Walker & Peñarrubia (2011) applied a mass estimator to the data for the two components and concluded that their analysis ruled out an NFW profile at  $\gtrsim 99$  per cent significance. Amorisco & Evans (2012) used a separable distribution function method and found strong statistical preference for a cored DM profile, while Strigari, Frenk & White (2017) found that with a more flexible distribution function model, the statistical preference went away and the data were consistent with an NFW halo.

Aside from addressing the core-cusp problem, robustly inferring the DM density distribution in dwarf galaxies is also important in the context of DM indirect detection. Indirect detection is the process in which DM annihilates or decays into Standard Model (SM) particles, and the resulting SM particles are subsequently detected. The probability of detecting such a signal is maximized in regions of the sky with high DM density, such as the centres of dwarf galaxies or the Milky Way Galactic Center (GC). Indeed, an excess of  $\sim$ GeV photons was detected near the GC by the *Fermi* Large Area Telescope (Atwood et al. 2009), which could be interpreted as a signal of DM annihilation (e.g. Goodenough & Hooper 2009; Calore, Cholis & Weniger 2015b; Ajello et al. 2016; Daylan et al. 2016). However, DM analyses near the GC are complicated by bright and complex astrophysical backgrounds, and it is important to have complementary search targets, some of which have excluded or placed the DM interpretation of the excess under tension (e.g. Ackermann et al. 2015; Albert et al. 2017; Lisanti et al. 2018a; Calore, Serpico & Zaldívar 2018; Chang, Lisanti & Mishra-Sharma 2018; Hoof, Geringer-Sameth & Trotta 2020; Di Mauro et al. 2019).

Some of the complementary targets studied in the indirect detection literature have been the Milky Way halo at high latitudes (Ackermann et al. 2012; Chang et al. 2018; Zechlin, Manconi & Donato

2018), galaxy groups (Lisanti et al. 2018a,b), Andromeda (Di Mauro et al. 2019), and stacked dwarf galaxies (e.g. Ackermann et al. 2011, 2015; Geringer-Sameth & Koushiappas 2011; Albert et al. 2017; Calore et al. 2018; Hoof et al. 2018). In particular, dwarf galaxies are generally considered to be the most robust search targets within the indirect detection literature, because they are expected to have little astrophysical background emission (Gallagher et al. 2003; Grcevich & Putman 2009).

In general, the expected signal flux from DM annihilation is proportional to the so-called astrophysical  $J$ -factor, which is defined as the integrals along the line of sight  $s$  and over the solid angle  $\Omega$  of the DM density squared,

$$J = \int d\Omega \int ds \rho^2(s, \Omega), \quad (1)$$

where  $\rho$  is the DM density. The robustness of any dwarf galaxy-based indirect detection constraint on DM annihilation is dependent on accurately estimating the  $J$ -factors of the analysed dwarf galaxies, and therefore dependent on accurately inferring their DM density distributions.

Finally, reliably reconstructing the total DM mass in dwarf galaxies also has important scientific ramifications. As we demonstrate in this paper, this is related to – but can be separate from – accurately inferring the full DM density distribution, because while the density and enclosed mass distributions are directly related, there can be cases where the total mass is accurately estimated even if the shape of the density distribution is not fully reconstructed. Obtaining accurate estimates of the total DM mass in dwarf galaxies plays a key role in determining the low-mass end of the stellar-to-halo mass relation (SHMR) (see Wechsler & Tinker 2018 for a review of the galaxy-halo connection). Studies have found that galaxy formation is significantly suppressed in DM haloes with virial mass below  $\sim 10^8 M_\odot$  (e.g. Read et al. 2017; Benítez-Llambay & Frenk 2020) and that scatter at the low-mass end of the SHMR may be required to fit the data well (Behroozi, Wechsler & Conroy 2013; Garrison-Kimmel et al. 2017). A more accurate determination of the DM halo mass in the smallest dwarf galaxies would help empirically anchor the SHMR for the smallest systems, for which the uncertainty on the relation between galaxies and their DM haloes is the largest.

In this paper, we apply spherical Jeans modelling (Jeans 1915; Binney 1980; Merritt 1985; Dejonghe & Merritt 1992) to simulated dwarf galaxy kinematic data sets, varying over properties of the mock observations such as the total number of observed stars, the measurement error on line-of-sight velocities, as well as the locations of the observed stars (e.g. whether they are primarily in the central region of the dwarf or farther out). We choose to focus on spherical isotropic dwarf galaxies in equilibrium. By studying the limitations of the Jeans analysis method even in this simplified scenario, we are able to identify which observational advancements are more likely to make an impact on our ability to accurately reconstruct the properties of dwarf galaxy DM haloes in the near future.

This paper is organized as follows. In Section 2, we present details on the Jeans modelling method and simulated data sets used in this study. In Section 3, we explore the effects of the number of observed stars (Section 3.1), the measurement errors of the line-of-sight velocities (Section 3.2), the locations of the observed stars (Section 3.3), and the level of ‘embeddedness’ of the stars (Section 3.4) on the DM inference. Section 4 recasts our results into the context of  $J$ -factors for indirect detection, where we emphasize the dependence of the indirect detection results on the priors chosen in the Jeans analysis and discuss recommendations for future observations. We summarize our main conclusions in Section 5.

## 2 METHODS

In this section, we describe the Jeans modelling procedure we employ (Section 2.1), the models we use to parametrize the distributions of the DM (Section 2.2) and the stars (Section 2.3), the specifics of the mock data we generate (Section 2.4), and the priors we assume for the model parameters throughout our analysis (Section 2.5). We use the public code STARSAMPLER (Liu 2019) to generate our mock data.

### 2.1 Jeans modelling

We summarize the standard procedure for inferring the velocity dispersion profile of the stars in a dwarf galaxy from measurements of their line-of-sight velocities, following the derivations of Binney & Mamon (1982) and Binney & Tremaine (2008). We start with the collisionless Boltzmann equation,

$$\frac{\partial f}{\partial t} + \vec{v} \cdot \frac{\partial f}{\partial \vec{x}} - \frac{\partial \Phi}{\partial \vec{x}} \cdot \frac{\partial f}{\partial \vec{v}} = 0, \quad (2)$$

where  $f$  is the phase-space density of a stellar tracer population, a function of the position  $\vec{x}$  and velocity  $\vec{v}$  of each star, and  $\Phi$  is the gravitational potential of the dwarf galaxy. Multiplying equation (2) by velocity component  $v_j$  and integrating over all velocities, we have

$$\frac{\partial}{\partial t} (v \bar{v}_j) + \frac{\partial}{\partial x_i} (v \bar{v}_i v_j) + v \frac{\partial \Phi}{\partial x_j} = 0, \quad (3)$$

where we have defined  $v = \int d\vec{v}^3 f(\vec{x}, \vec{v})$ , the spatial number density of the tracer stars. Assuming the system is spherically symmetric and is in steady state (and therefore the  $\partial/\partial t$  term is negligible), we have

$$\frac{\partial}{\partial r} (v \sigma_r^2) + v \left( \frac{\partial \Phi}{\partial r} + \frac{2\sigma_r^2 - \sigma_\theta^2 - \sigma_\phi^2}{r} \right) = 0, \quad (4)$$

where  $\sigma_i^2$  is the square of the  $i$ th component of the velocity dispersion, i.e.  $\sigma_i^2 = \langle v_i^2 \rangle - \langle v_i \rangle^2$ , for  $i \in \{r, \theta, \phi\}$ .

We can then define the velocity anisotropy,

$$\beta(r) = 1 - \frac{\sigma_\theta^2 + \sigma_\phi^2}{2\sigma_r^2}, \quad (5)$$

and explicitly write the potential as

$$\Phi = -\frac{GM(<r)}{r}, \quad (6)$$

where  $G$  is the gravitational constant and  $M(<r)$  is the enclosed mass within radius  $r$ . Plugging these quantities back into equation (4), we end up with the following first-order differential equation for  $v\sigma_r^2$ :

$$\frac{1}{v} \left[ \frac{\partial}{\partial r} (v\sigma_r^2) + \frac{2\beta(r)}{r} (v\sigma_r^2) \right] = -\frac{GM(<r)}{r^2}. \quad (7)$$

The generic solution to equation (7) takes the form

$$v(r)\sigma_r^2(r) = \frac{1}{g(r)} \int_r^\infty \frac{GM(<\tilde{r})v(\tilde{r})}{\tilde{r}^2} g(\tilde{r}) d\tilde{r}, \quad (8)$$

where the new function  $g(r)$  is defined as

$$g(r) = \exp \left( 2 \int \frac{\beta(r)}{r} dr \right). \quad (9)$$

The enclosed mass  $M(<r)$  in equation (8) can be related to the overall density distribution by

$$M(<r) = 4\pi \int_0^r \rho(s) s^2 ds, \quad (10)$$

where, again, we have assumed spherical symmetry of the system. While both the stars and DM contribute to the mass density

distribution, i.e.  $\rho = \rho_{\text{DM}} + \rho_{\text{stars}}$ , we expect the density of DM to dominate, and therefore make the approximation  $\rho \approx \rho_{\text{DM}}$ . This is a valid approximation due to the large mass-to-light ratios of dwarf galaxies,  $M_{\text{halo}}/M_* \approx 10^2-10^5$  (Garrison-Kimmel et al. 2017).

In practice, typically only projected radii and line-of-sight velocities are measured, and therefore equation (8) needs to be projected along the line of sight. To do so, we use the Abel transform, defined for a spherically symmetric function as

$$S(R) = 2 \int_R^{+\infty} \frac{s(r)r dr}{\sqrt{r^2 - R^2}}, \quad (11)$$

where  $s(r)$  is the function in three-dimensional spherical coordinates,  $R$  is the projected radius, and  $S(R)$  is the resulting projected function. Projecting equation (8) along the line of sight leads to the equation (Binney & Mamon 1982; Mamon & Boué 2010)

$$\sigma_p^2(R)I(R) = 2 \int_R^\infty \left( 1 - \beta(r) \frac{R^2}{r^2} \right) \frac{v(r)\sigma_r^2(r)r}{\sqrt{r^2 - R^2}} dr, \quad (12)$$

where  $\sigma_p$  is the projected velocity dispersion profile and  $I(R)$  is the projected number density distribution of the tracer stars, given by

$$I(R) = 2 \int_R^\infty \frac{v(r)r dr}{\sqrt{r^2 - R^2}}. \quad (13)$$

Throughout the remainder of this paper,  $I(R)$  is referred to as the surface brightness profile or light profile.

Using equation (12), we build a likelihood function to fit the observed data and extract information on the dark matter distribution. In the literature, the analysis has been performed in either a binned (e.g. Strigari et al. 2007b; Charbonnier et al. 2011) or unbinned (e.g. Strigari et al. 2008) fashion. In this work, we will focus on the unbinned analysis. The unbinned Gaussian likelihood function is given by (Strigari et al. 2008)

$$\mathcal{L} = \prod_{i=1}^{N_{\text{stars}}} \frac{(2\pi)^{-1/2}}{\sqrt{\sigma_p^2(R_i) + \Delta_{v_i}^2}} \exp \left[ -\frac{1}{2} \left( \frac{v_i - \bar{v}}{\sigma_p^2(R_i) + \Delta_{v_i}^2} \right)^2 \right], \quad (14)$$

where  $\bar{v}$  is the mean velocity for the population of tracer stars, and for star  $i$ ,  $v_i$  is the measured line-of-sight velocity,  $\sigma_p(R_i)$  is the intrinsic velocity dispersion at the projected radius  $R_i$ , and  $\Delta_{v_i}$  is the velocity measurement error. In our analysis, we choose closed-form parametrizations for the stellar and dark matter distributions, thereby reducing the number of integrals that need to be performed when calculating the likelihood.

It is important to emphasize the interplay between the intrinsic velocity dispersion and the measurement error in equation (14) – if the measurement errors are subdominant to the intrinsic velocity dispersion of the system, it is not expected that improvements to the line-of-sight velocity measurements would drastically improve the quality of the fit. This will be further discussed in Section 3.2.

The degeneracy between the velocity anisotropy,  $\beta(r)$ , and the enclosed mass profile,  $M(<r)$ , is a known complication in Jeans modelling (e.g. Merrifield & Kent 1990; Wilkinson et al. 2002; Lokas & Mamon 2003; De Lorenzi et al. 2009; Read & Steger 2017; Genina et al. 2020). It can be seen from equation (12) that  $\beta(r)$  and  $\sigma_p^2(r)$  are degenerate with each other, which, combined with equation (8), implies that  $\beta(r)$  is degenerate with  $M(<r)$ . Unfortunately,  $\beta(r)$  can only be measured with full 3D velocity information, which is not yet available for the majority of the stars in dwarf galaxies. It is therefore common in Jeans analyses to assume a parametric model for  $\beta(r)$  and fit for it in conjunction with fitting for  $M(<r)$  (e.g. Mamon & Łokas 2005; Mamon, Biviano & Boué 2013; Bonnard et al. 2015b; Geringer-Sameth, Koushiappas & Walker

2015b; Mashchenko 2015). The effect on dynamical mass modelling estimates when the assumed  $\beta(r)$  model does not match the true velocity anisotropy distribution has been studied in El-Badry et al. (2017). In this work, we choose to focus entirely on isotropic data sets and models in order to understand the limitations of the Jeans modelling procedure even in the absence of additional complications due to velocity anisotropy, i.e. we assume

$$\beta(r) = 0. \quad (15)$$

While this was chosen for demonstrative purposes, for CDM haloes,  $\beta$  has been found to be consistent with 0 for dwarf galaxies in the APOSTLE simulations (Genina et al. 2020) and in the central regions of isolated dwarfs from the FIRE simulations (González-Samaniego et al. 2017), further motivating our choice.

## 2.2 Dark matter profile

Using STARSAMPLER, we generate the tracer stars in a DM potential which follows the Hernquist/Zhao profile (Hernquist 1990; Zhao 1996)

$$\rho_{\text{DM}}^{\text{Zhao}}(r) = \rho_0 \left( \frac{r}{r_s} \right)^{-\gamma} \left[ 1 + \left( \frac{r}{r_s} \right)^\alpha \right]^{(\gamma-\beta)/\alpha}, \quad (16)$$

where  $\alpha$ ,  $\beta$ ,  $\gamma$  are the slopes of the distribution,  $\rho_0$  is the overall normalization of the density profile, and  $r_s$  is the scale radius – in particular,  $\gamma$  sets the asymptotic inner slope of the distribution. This model has five free parameters, which introduces too many degenerate degrees of freedom into the model to effectively constrain the parameters (we discuss the role of degeneracies in the Supplementary Material). We therefore simplify the DM profile by setting  $\alpha = 1$  and  $\beta = 3$ , which reduces equation (16) to a generalized Navarro–Frenk–White (gNFW) distribution with inner slope parameter  $\gamma$ , defined as (Navarro et al. 1997)

$$\rho_{\text{DM}}^{\text{gNFW}}(r) = \rho_0 \left( \frac{r}{r_s} \right)^{-\gamma} \left( 1 + \frac{r}{r_s} \right)^{-(3-\gamma)}. \quad (17)$$

While we use the gNFW distribution to model the DM profile in our fiducial analysis set-up, we additionally consider the special cases where the inner slope  $\gamma = 0$  or 1. The case of  $\gamma = 1$  corresponds to the standard, cuspy Navarro–Frenk–White (NFW) profile

$$\rho_{\text{DM}}^{\text{NFW}}(r) = \rho_0 \left( \frac{r}{r_s} \right)^{-1} \left( 1 + \frac{r}{r_s} \right)^{-2}, \quad (18)$$

whereas the case of  $\gamma = 0$  leads to a constant-density central core. We refer to this distribution as the cored NFW (NFWc) distribution, given by

$$\rho_{\text{DM}}^{\text{NFWc}}(r) = \rho_0 \left( 1 + \frac{r}{r_s} \right)^{-3}. \quad (19)$$

The profiles defined by equations (17)–(19) give rise to closed-form enclosed mass distributions, which we list in the Supplementary Material for reference.

## 2.3 Light profile

Using STARSAMPLER, we can model the stellar density distribution also as a Hernquist/Zhao profile

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

**Table 1.** DM halo parameters and properties of the data sets generated in this work.  $\rho_0$ ,  $r_s$ , and  $\gamma$  are the true values of the normalization, scale length, and inner slope input into equation (17).  $M_{200}$  is defined as the enclosed mass at  $r_{200}$ , the radius within which the average density is equal to 200 times the critical density of the Universe at redshift  $z = 0$ , derived from the true density distribution. We adopt a generalized definition of the concentration  $c_{200} \equiv r_{200}/r_s$  for all of our parameter sets.  $\sigma_p$  is the median line-of-sight velocity dispersion across all the data sets generated for each set of parameters (10 realizations each for sample sizes of 20, 100, 1000, and 10 000 stars, resulting in a total of 40 data sets).

		$\rho_0$ [ $M_\odot/\text{kpc}^3$ ]	$r_s$ [kpc]	$\gamma$	$M_{200}$ [ $M_\odot$ ]	$c_{200}$	$\sigma_p$ [ $\text{km s}^{-1}$ ]
Cusp	I	$6.4 \times 10^7$	1	1	$1.9 \times 10^9$	25.8	14.6
	II	$6.4 \times 10^7$	0.2	1	$1.5 \times 10^7$	25.8	2.9
Core	III	$6.4 \times 10^7$	1	0	$1.4 \times 10^9$	23.6	9.5
	IV	$6.4 \times 10^7$	0.2	0	$1.1 \times 10^7$	23.6	1.9

In this paper, we generate stars following a Plummer profile, which is a specific case of equation (20).<sup>1</sup> For ease of comparison across different samples, the stellar mocks are all generated with the same level of embeddedness in their respective DM haloes by setting the scale radius of the tracers,  $r_*$ , to be equal to the scale radius of the DM distribution,  $r_s$ .<sup>2</sup>

Correspondingly, in our Jeans analysis, we model the stellar density  $\nu(r)$  as a 3d Plummer profile (Plummer 1911), defined as

$$\nu(r) = \frac{3L}{4\pi r_*^3} \left( 1 + \frac{r^2}{r_*^2} \right)^{-5/2}, \quad (21)$$

where  $L$  is the total luminosity and  $r_*$  is scale length of the distribution. equation (21) has the same form as equation (20), with  $\alpha_* = 2$ ,  $\beta_* = 5$ ,  $\gamma_* = 0$ , and  $\rho_* = 3L/(4\pi r_*^3)$ .<sup>3</sup> The surface brightness profile (or light profile), which is the projection of  $\nu(r)$  along the line of sight, is then given by the closed-form expression

$$I(R) = \frac{L}{\pi r_*^2} \left( 1 + \frac{R^2}{r_*^2} \right)^{-2}. \quad (22)$$

Because we have assumed the contribution of the stellar tracers to the gravitational potential is negligible, changing the value of  $L$  in equations (21) and (22) does not meaningfully affect the result of the Jeans modelling.

## 2.4 Mock data

We generate data sets with four different sets of DM halo parameters (summarized in Table 1). Our parameter choices span different halo masses and either an inner cusp or inner core in the DM density profile

<sup>1</sup>The general form of the Hernquist/Zhao profile spans cored and cuspy stellar profiles. While more flexible, the extra degrees of freedom introduced by this stellar model further exacerbate the effects of parameter degeneracies on the fit (see Supplementary Materials for more discussion on the effect of parameter degeneracies).

<sup>2</sup>The choice of  $r_* = r_s$  was made for demonstrative purposes and ease of comparison between models. Note that for realistic choices of mass-to-light ratio, halo model, and halo concentration, this assumption may not be well motivated. We explore the effects of higher degrees of embeddedness in Section 3.4.

<sup>3</sup>In practice, when we generate our samples, we set  $\gamma_* = 0.1$  as setting  $\gamma_* = 0$  results in an unphysical distribution function (Liu 2019). This corresponds to an approximate Plummer profile; we do not expect the small deviation from a true Plummer profile to affect our results.

while maintaining approximately the same halo concentration. Due to the large amount of scatter in the theoretical predictions for the subhalo mass–concentration relation, we choose not to focus on a specific mass–concentration model; however, the concentrations of our simulated haloes are broadly consistent within uncertainty with theoretical predictions in the literature for the relevant mass range (Pieri et al. 2011; Sánchez-Conde & Prada 2014; Moliné et al. 2017; Wang et al. 2020).

Parameter sets I and III correspond to  $M_{200} \sim 10^9 M_{\odot}$  haloes, while sets II and IV correspond to smaller haloes with mass  $M_{200} \sim 10^7 M_{\odot}$ . We emphasize that we have chosen to study  $M_{200} \sim 10^7 M_{\odot}$  haloes for demonstrative purposes, to study how the effect of the measurement error on the line-of-sight velocities impacts less massive haloes differently from more massive ones. We have adopted a generalized definition of the halo concentration,  $c_{200} \equiv r_{200}/r_s$ , for all of the parameter sets that we generate, where  $r_{200}$  is the radius within which the average density is 200 times the critical density of the Universe at redshift  $z = 0$ . The virial mass  $M_{200}$  is subsequently defined as the enclosed mass at  $r_{200}$ .

For each set of DM parameters, we generate 10 realizations each of data sets with 20, 100, 1000, and 10 000 stars, respectively. The chosen sample sizes are meant to provide comparison with current measurements of ultrafaint dwarfs and classical dwarfs (see the table of observed dwarf galaxies in the Supplementary Material for comparison), as well as projections for how future measurements might improve the quality of the DM inference. For our fiducial analyses, we assume a measurement error of  $\Delta v = 2 \text{ km s}^{-1}$  on the line-of-sight velocity. This is comparable to the typical errors in current measurements. We explore the effect of increasing or decreasing the measurement error relative to our fiducial value of  $2 \text{ km s}^{-1}$  in Section 3.2. Throughout this paper, we will use  $R$  to denote the projected radius and  $r$  to denote the 3d galactocentric radius. We show representative distributions of the projected and 3d galactocentric radii in our generated stellar data sets in the Supplementary Material.

## 2.5 Parameters and priors

We perform our Jeans modelling procedure in two stages. First, we perform a fit to only the positions of the stars. We describe this light profile fitting procedure in the Supplementary Material. We do so because the light profile is generally much better constrained than the stellar kinematics. We can then use the results from the initial fit to set the prior range on the light profile parameters in our full Jeans fit. We conservatively set the prior ranges on the light profile parameters in the full scan to be the middle 95 per cent containment range of the posterior probability distributions output from the initial fit. In both stages, we use the PYMULTINEST module (introduced in Buchner et al. 2014), which interfaces with the nested sampling Monte Carlo library MULTINEST (Feroz, Hobson & Bridges 2009), to sample the relevant likelihood.<sup>4</sup>

We summarize the priors for all of the parameters in our model in Table 2. We choose a wide prior of  $[-1, 5]$  for the parameter  $\gamma$ , which sets the inner slope of the gNFW distribution. The lower edge is chosen to be at  $-1$  such that there is sufficient range for

<sup>4</sup>We use  $n_{\text{live}} = 100$  live points in the nested sampling procedure throughout this paper, but have verified that increasing to  $n_{\text{live}} = 1000$  does not affect the results. Convergence is controlled by the evidence tolerance factor. The recommended value for the tolerance factor is 0.5, while in this work, it has been conservatively set to 0.25.

**Table 2.** Prior ranges for the stellar and DM parameters used in our analysis. We implement uniform priors within each of the listed prior ranges. The ranges listed here for  $\log_{10}(r_s)$  and  $\log_{10}(L)$  are used in the initial light profile fit; in the full Jeans scan, we set the prior ranges for  $\log_{10}(r_s)$  and  $\log_{10}(L)$  to be the middle 95 per cent containment range of the posterior for each parameter from the initial fit (see Supplementary Material for more discussion on the light profile fit).

Parameter	Prior
$\log_{10}(r_s/\text{kpc})$	$[-3, 3]$
$\log_{10}(L/L_{\odot})$	$[-2, 5]$
$\ln(\rho_0/(M_{\odot} \text{ kpc}^{-3}))$	$[5, 30]$
$\ln(r_s/\text{kpc})$	$[-10, 10]$
$\bar{v}/(\text{km s}^{-1})$	$[-100, 100]$
$\gamma$	$[-1, 5]$

convergence at  $\gamma = 0$  while not allowing for larger negative values of  $\gamma$ , which are unphysical. We note that values of  $\gamma \geq 3$  are also unphysical, as they lead to divergent enclosed mass at finite radius; we have verified that excluding these values from our prior range has negligible effect on our results (see Supplementary Material for more detailed discussion on varying priors).

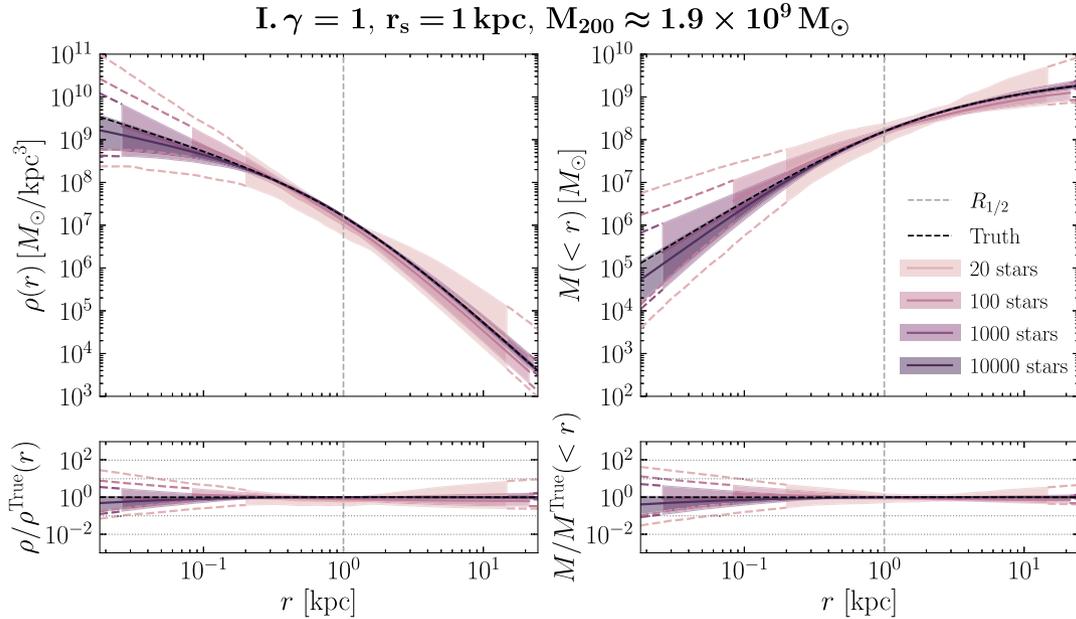
In our fiducial model, there are a total of six free parameters: two for the light profile, three for the DM density distribution parametrized as a gNFW profile, and one for the mean stellar velocity. In our discussion on characterizing the inner slope of the DM distribution, we additionally perform fits assuming either an NFW or cored NFW distribution, and compare the Bayesian evidence between the two models – in these fits, there are a total of five free parameters.

## 3 RESULTS

We now apply the analysis pipeline described in Section 2 to the simulated stellar samples described in Section 2.4 and summarized in Table 1. Our main figures of merit for evaluating the success or limitations of our analyses are: (i) the overall recovered DM density profile, (ii) the recovered enclosed DM mass, which we quantify as the recovered virial mass  $M_{200}$ ,<sup>5</sup> and (iii) the recovered inner slope of the DM density profile, i.e. the parameter  $\gamma$  in equation (17). Of the figures of merit, (i) has important implications on the inferred astrophysical  $J$ -factors (equation 1) which are used in indirect DM searches, (ii) is crucial for empirically probing the SHMR down to low halo masses, while (iii) can shed light on the particle physics properties of the DM as well as baryonic feedback and galaxy formation mechanisms.

We explore how several factors in the analysis influence the accuracy of the inferred DM profiles, focusing primarily on the effects of variations on the specifics of the analysed data sets. In Section 3.1, we study how the total number of observed stars influences the inferred DM profile. In Section 3.2, we study the role of the line-of-sight velocity measurement errors; we explore how the magnitude of the error differently impacts the DM inference in

<sup>5</sup>We note that in practice, due to tidal stripping,  $M_{200}$  is not a well-defined metric in dwarf galaxies. While we include it here as one measure of the quality of the fit, we also show the comparisons of the full recovered and true enclosed mass profiles  $M(<r)$ .



**Figure 1.** Inferred DM density profiles  $\rho(r)$  (left-hand panels) and corresponding enclosed mass profiles  $M(<r)$  (right-hand panels) for parameter set I – from lightest to darkest colour, we plot the results for samples with 20, 100, 1000, and 10000 stars. We show the full reconstructed distributions in the top panels as well as the fractional (relative to truth) distributions in the bottom panels. For each sample size, the solid line denotes the median (across our 10 independent realizations) of the median recovered profiles, while the shaded band shows the median of the 68 per cent containment regions, plotted from the innermost to outermost star across all 10 data sets for that sample size. For the samples with fewer than 10000 stars, the median 68 per cent containment regions over the full radial range are shown bracketed by each pair of dashed lines in the colour corresponding to the sample size. Across all sample sizes, the typical inferred density profile and enclosed mass profile are consistent within uncertainty with the true distributions. Increasing the observed sample size reduces the uncertainty on the recovered profiles, as expected.

dwarf galaxies with different halo masses. In Section 3.3, we study the effect of the locations of observed stars on the inferred DM profile. In Section 3.4, we study how the degree of embeddedness of the stars within the DM halo affects the DM inference.

### 3.1 Increase in sample size

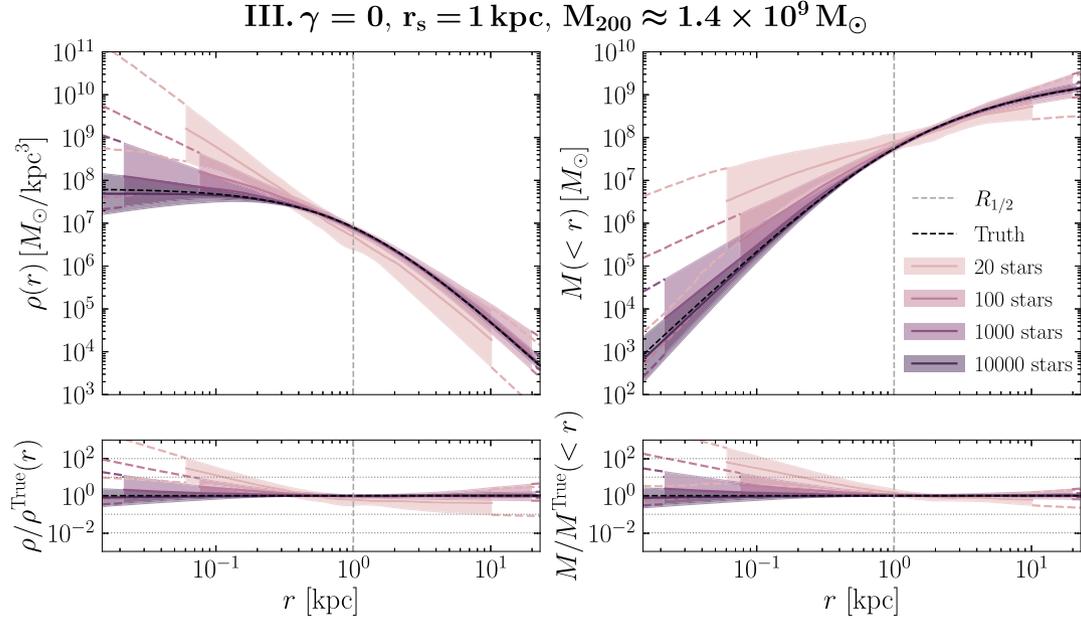
Our first question of interest is how the number of observed stars in a dwarf galaxy affects the DM inference. In Fig. 1, we show the inferred DM density profiles  $\rho(r)$  and corresponding enclosed mass profiles  $M(<r)$  for parameter set I (which has  $\gamma = 1$ ), for the four different sample sizes – from lightest to darkest colour, we plot the results for 20, 100, 1000, and 10000 stars. For a given sample size, we run each of our 10 realizations through the analysis pipeline and obtain the resulting posterior density and enclosed mass profiles. Each solid line in Fig. 1 shows the median across the 10 realizations of the median recovered profiles, while the shaded band depicts the median of the 68 per cent containment regions across the realizations. The solid line and shaded band for each sample size are plotted from the innermost to outermost star across the 10 generated data sets for that sample size; outside of the radial range for the smaller sample sizes, we outline just the 68 per cent containment region with dashed lines in the colour corresponding to each sample size. The extension down to smaller radii is particularly important in understanding the implications for indirect detection, which we discuss in Section 4. The vertical dashed grey line indicates the projected half-light radius,  $R_{1/2}$ , which for a Plummer profile is equal to the scale radius  $r_*$ .

We find that, for all sample sizes in parameter set I ( $\gamma = 1$ ), the typical inferred density profile and enclosed mass profile are consistent within uncertainty with the true distributions over the full range of measured radii. This can be seen from the fact that the dashed

black lines in the top panels of Fig. 1, indicating the true distributions, are contained within the bands for all of the sample sizes, as well as the fact that all the bands in the bottom panels overlap with the horizontal dashed black line. Additionally, we find that increasing the observed sample size reduces the uncertainty on the inferred density and enclosed mass profiles, as is to be expected. For all sample sizes and parameter sets, we tabulate the median across our 10 realizations of the median and  $\pm 1\sigma$  values of the inferred virial mass,  $M_{200}$ , in the Supplementary Material.

We show the analogous results for parameter set III ( $\gamma = 0$ ) in Fig. 2. In this case, for sample sizes of 20 stars and 100 stars, the inferred density distribution is typically biased towards a steeper inner profile than the true distribution, which has an inner slope of  $\gamma = 0$ , while for the data sets with 1000 and 10000 stars, the typical inferred density profiles are consistent with the true distribution within uncertainty. Importantly, across all of the sample sizes, we obtain an accurate estimate for the total mass of the system, with the uncertainties on the estimate reduced as the sample size is increased (values listed in the fourth column of the corresponding table in the Supplementary Material).

This suggests that while the inferred density distribution may not always accurately represent the true underlying distribution, the virial mass estimate remains fairly robust. Namely, if the inferred density profile is biased high in the inner region of the dwarf (as seen in the  $r \lesssim R_{1/2}$  region for the smaller samples from parameter set III), this is compensated for by the density profile being biased low in the outer region. We note that because the outer slope of the density profile is not a free parameter in the fit, the outer profile is uniquely determined by the scale radius and overall normalization. Our likelihood (equation 14) depends directly on the enclosed mass distribution of the system rather than the density distribution, and



**Figure 2.** Same as Fig. 1, except for parameter set III. For sample sizes of 20 and 100 stars, the inferred density distribution is typically biased towards a steeper inner profile than the true distribution; however, the inferred virial mass is still consistent with the true virial mass (values tabulated in the Supplementary Material). For sample sizes with 1000 and 10 000 stars, both the inferred density distribution and enclosed mass profile are consistent within uncertainty with the true distributions across the measured radial range.

therefore it is not surprising that the fit is successful at recovering the total mass of the system even when it fails to accurately reproduce the inner density profile.

Figs 1–2 demonstrate that the inner regions of the inferred DM density profiles can be biased and/or poorly constrained, especially for the smaller data sets. We can further assess how well the inner density profile is recovered by directly examining the posterior probability distribution of the parameter in our model which sets the asymptotic inner slope,  $\gamma$ . In the top row of Fig. 3, we show histograms of the posterior  $\gamma$  values corresponding to the scans shown in Fig. 1 (parameter set I), i.e. for a true inner slope of  $\gamma = 1$ . The lines (bands) show the median (middle 68 per cent) in each bin across the 10 realizations. The inner slope is generally poorly constrained for the smaller samples, with the median posterior distribution only peaking around the true value of  $\gamma = 1$  for the largest sample size of 10 000 stars – notably, even in this case, there is typically non-negligible posterior probability at  $\gamma = 0$ , so we would not be able to exclude an incorrect inner slope value of 0 at high significance. We also draw attention to the fact that, for the samples with 100 and 1000 stars, although the posterior distributions are fairly flat and poorly constrained, the posterior probability sharply drops off above  $\gamma \sim 2$ . This is important because the enclosed mass for a gNFW profile (equation provided in the Supplementary Material) diverges at finite  $r$  for  $\gamma \geq 3$ . For the most statistics-limited samples containing only 20 stars, the fit cannot fully exclude unphysical values of  $\gamma \geq 3$ .

In the bottom row of Fig. 3, we show the results for parameter set III, which has a true inner slope of  $\gamma = 0$ . The results are qualitatively similar: the posterior distributions of  $\gamma$  tend to be poorly constrained for the smaller sample sizes, and we are only able to recover the true value of the inner slope for the 10 000-star samples. In this case, for the largest sample size, we would be able to exclude an incorrect inner slope value of 1 at high significance. However, for data sets with  $\leq 1000$  stars from both parameter sets – on par with the existing dwarf galaxy measurements – we cannot determine whether the underlying

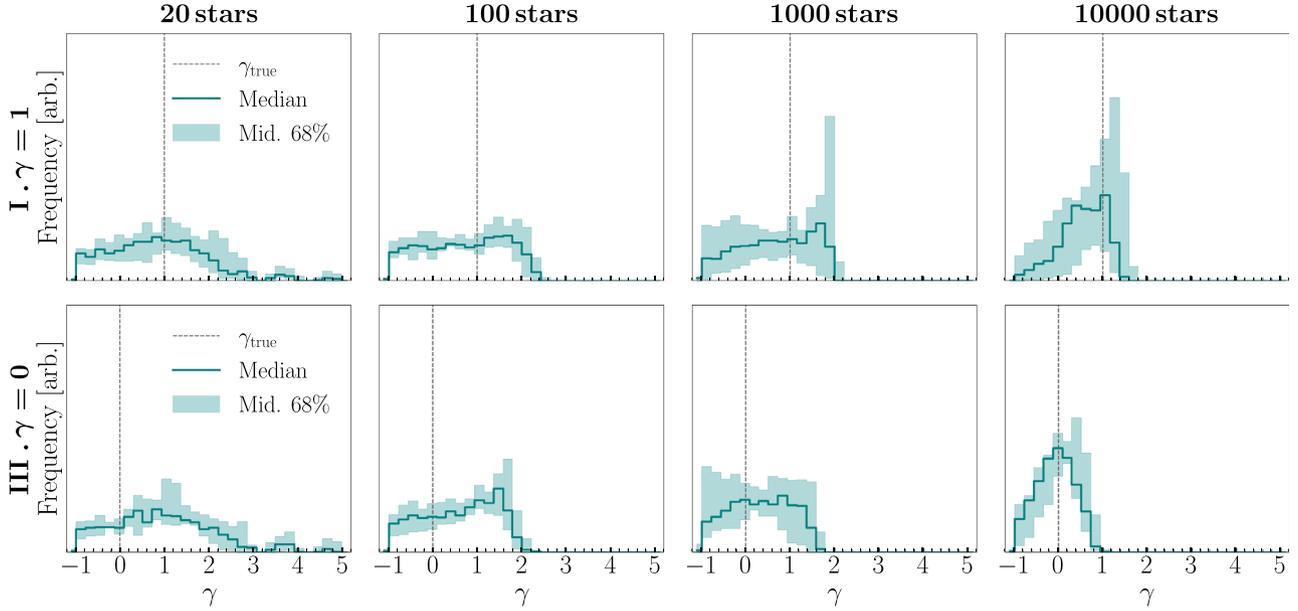
halo has an inner slope of  $\gamma = 0$  or  $\gamma = 1$  in a statistically significant manner, consistent with previous Jeans modelling-based results in the literature (e.g. Walker et al. 2009; Read, Walker & Steger 2018; Genina et al. 2020). We further note that, for the smaller sample sizes, the fact that the posterior  $\gamma$  distributions are unconstrained implies that the results are highly sensitive to the choice of priors on  $\gamma$ , and we therefore choose to present the full posterior distributions rather than to quote recovered median values or quantiles.

A separate method for quantifying the ability of this procedure to distinguish whether the underlying DM distribution has an inner cusp ( $\gamma = 1$ ) or core ( $\gamma = 0$ ) is to compare the statistical preference for a cuspy DM model over a cored DM model, or vice versa. In particular, we analyse the same data sets as before, this time fixing the value of  $\gamma$  in our model to either 1 or 0 in equation (17). The resulting models, respectively, correspond to the standard NFW distribution (equation 18) or the cored NFW distribution (equation 19). We then calculate the Bayes factor (BF) in preference for a model in which  $\gamma$  is fixed to the true value for the given data set, relative to a model in which  $\gamma$  is fixed to the alternative value, i.e.

$$\text{BF} = \frac{\text{Pr}(d|\gamma = \gamma_{\text{true}})}{\text{Pr}(d|\gamma = \gamma_{\text{alt.}})}. \quad (23)$$

On the Jeffreys scale, as amended by Kass & Raftery (1995),  $\text{BF} < 3.2$  is ‘not worth more than a bare mention,’  $\text{BF} \in [3.2, 10)$  provides substantial evidence,  $\text{BF} \in [10, 100)$  provides strong evidence, and  $\text{BF} \geq 100$  provides decisive evidence.

In Table 3, we list for parameter set I the median and  $\pm 1\sigma$  (second column) as well as the minimum (third column) and maximum (fourth column) BF values in preference for the true value of  $\gamma = 1$  across the 10 data sets. For the smaller samples, the BF values are generally indeterminate, which is consistent with the relatively unconstrained posterior distributions shown in the top row of Fig. 3. For a sample size of 10 000 stars, the median BF is also indeterminate, although we find that there is one realization for which there is decisive evidence, and two realizations for which there is strong



**Figure 3.** Posterior distributions for the inner slope  $\gamma$ . The top row corresponds to the scans shown in Fig. 1 (parameter set I), while the bottom row corresponds to the scans shown in Fig. 2 (parameter set III). The lines (bands) show the median (middle 68 per cent) in each  $\gamma$  bin across the 10 realizations. In both cases, the inner slope is generally poorly constrained for the smaller samples, with the median posterior distribution only peaking near the true value of  $\gamma$  (vertical dashed line in each panel) for the largest sample size of 10 000 stars. All panels in this figure share the same vertical scale.

**Table 3.** Values of the Bayes Factor (BF) from fitting parameter set I with a model assuming a cusp ( $\gamma = 1$ ) relative to a model assuming a core ( $\gamma = 0$ ). The second column lists the median and lower/upper  $1\sigma$ , while the third (fourth) column lists the minimum (maximum) BF value across the 10 data sets for each sample size. Of the 10 000-star samples, two realizations have  $10 \leq \text{BF}_{1,0} < 100$ , providing strong evidence, and one realization has  $\text{BF}_{1,0} \geq 100$ , providing decisive evidence in favour of a cusp over a core.

I . $\gamma = 1$ , $r_s = 1$ kpc			
$n_{\text{stars}}$	$\text{BF}_{1,0} = \frac{\text{Pr}(d \gamma=1)}{\text{Pr}(d \gamma=0)}$	$\min(\text{BF}_{1,0})$	$\max(\text{BF}_{1,0})$
20	$1.27^{+0.70}_{-0.52}$	0.44	2.60
100	$1.18^{+0.74}_{-0.40}$	0.56	3.05
1000	$1.77^{+1.57}_{-0.96}$	0.67	6.19
10 000	$2.02^{+54.45}_{-1.52}$	0.22	225.70

evidence, in favour of a model with a cusp. This is consistent with the rightmost panel in the top row of Fig. 3, in which the average posterior probability is non-negligible at  $\gamma = 0$  and there is significant variation in the height of the peak at  $\gamma \sim 1$  across realizations. Although there is significant scatter in the BF values between realizations, we emphasize that the BF in preference for the cored model over the cuspy one is always less than 10 – the minimum benchmark for claiming statistically significant preference for a cored DM profile – and therefore, even in cases where we are unable to robustly identify the presence of a cusp, we would not falsely claim the presence of a core.

We list the analogous results for parameter set III in Table 4. In this case, the median BF for a sample size of 10 000 stars is decisively in favour of a model with a core. This is also consistent with the posterior distribution shown in the bottom rightmost panel of Fig. 3, which is peaked at  $\gamma \sim 0$ , sharply drops near  $\gamma \sim 1$ , and has relatively little spread across realizations. Importantly, across all sample sizes and realizations for parameter set I(III), for which the true DM profile is cuspy (cored), the BF in preference for a cored (cuspy) profile over a cuspy (cored) one is always less than 10. This demonstrates that,

**Table 4.** Same as Table 3, but for parameter set III, in this case comparing a model assuming a core ( $\gamma = 0$ ) to a model assuming a cusp ( $\gamma = 1$ ). Of the 10 000-star samples, seven realizations have  $\text{BF}_{0,1} \geq 100$ , providing decisive evidence in favour of a cored distribution over a cuspy one.

III . $\gamma = 0$ , $r_s = 1$ kpc			
$n_{\text{stars}}$	$\text{BF}_{0,1} = \frac{\text{Pr}(d \gamma=0)}{\text{Pr}(d \gamma=1)}$	$\min(\text{BF}_{0,1})$	$\max(\text{BF}_{0,1})$
20	$0.58^{+0.25}_{-0.06}$	0.50	1.42
100	$0.58^{+0.19}_{-0.16}$	0.27	1.78
1000	$0.96^{+2.83}_{-0.60}$	0.17	6.78
10 000	$256.74^{+837.48}_{-235.60}$	2.98	46971.85

even when we are unable to recover statistical evidence for the true inner DM profile, we would not erroneously claim evidence for the wrong inner profile.

Thus far, we have demonstrated that, for data sets with  $\lesssim 1000$  measured stars – on par with the current measurements – we can robustly recover the total enclosed DM mass, but we cannot accurately reconstruct the inner profile or constrain the inner slope of the DM density distribution, even within our simplified framework. We have also tested samples with 5000 stars and found that the posterior  $\gamma$  distributions were typically flat as well, demonstrating that in order to constrain  $\gamma$  in our set-up, a sample size of  $\sim 10\,000$  stars is truly needed.

A natural question to ask is what drives the difficulty of accurately recovering  $\gamma$  for data sets with  $\lesssim 10\,000$  stars. In the Supplementary Material, we explore how degeneracies between the DM model parameters contribute to this challenge as well as the potential to better constrain  $\gamma$  by jointly analysing multiple data sets simultaneously. In Section 4, we investigate how the limitations on being able to accurately reconstruct the full density profile – which we emphasize is related to, but separate from, the issue of constraining the posterior distribution of  $\gamma$  – may affect the results of indirect detection analyses.

### 3.2 Velocity uncertainties

Looking towards future measurements, it is important to understand how increasingly precise measurements of line-of-sight velocities might affect our ability to reconstruct DM halo properties. To address this, we generate simulated data sets assuming different values of measurement error  $\Delta v$  (uniform across all generated stars), and repeat our analysis setting  $\Delta_{v_i} = \Delta v$  for all stars in equation (14). We compare our fiducial results, which assume a measurement error of  $\Delta v = 2 \text{ km s}^{-1}$ , to results assuming a more conservative value of  $\Delta v = 5 \text{ km s}^{-1}$ , as well as results in the limit of perfect measurements,  $\Delta v = 0 \text{ km s}^{-1}$ .<sup>6</sup> If the intrinsic velocity dispersion of a system is much larger than  $\sim 5 \text{ km s}^{-1}$ , we do not expect varying  $\Delta v$  in the range of 0–5  $\text{km s}^{-1}$  to have a significant effect on the analysis results. On the other hand, if the intrinsic velocity dispersion is  $\lesssim 5 \text{ km s}^{-1}$ , we expect the results to be dependent on the value of  $\Delta v$ , especially if the sample size is small. For parameter sets I and III discussed in Section 3.1, the intrinsic velocity dispersion is  $\sim 10$ – $15 \text{ km s}^{-1}$ . Parameter sets II and IV have the same DM inner slope and concentration as parameter sets I and III, respectively, but are approximately 100 times less massive and have an intrinsic velocity dispersion of  $\sim 2$ – $3 \text{ km s}^{-1}$ .

In Fig. 4, we show the inferred fractional DM density and enclosed mass profiles for parameter set IV.<sup>7</sup> From top to bottom, the rows correspond to  $\Delta v = 0, 2, 5 \text{ km s}^{-1}$ . The results are consistent with our intuition: because the typical intrinsic velocity dispersion for this set of systems is  $\sim 2 \text{ km s}^{-1}$ , a measurement error  $\Delta v \gtrsim 2 \text{ km s}^{-1}$  has a drastic effect on the inferred results, especially when combined with limited sample size. A value of  $\Delta v = 5 \text{ km s}^{-1}$  results in an inferred virial mass of  $M_{200} \sim 0$  for both the 20- and 100-star samples (detailed numbers tabulated in the Supplementary Material). For the 20-star samples, even our fiducial choice of  $\Delta v = 2 \text{ km s}^{-1}$  results in essentially no DM being recovered. This can be understood as the measurement error being large enough that the observed velocity dispersion can be statistically consistent with the complete absence of DM.

For the larger sample sizes, with 1000 and 10 000 stars, the Jeans analysis is able to recover the correct density profile even when the measurement errors are of the same order as the dispersion of the system. This can be attributed to the fact that with large enough statistics, the analysis can distinguish the radially dependent velocity dispersion  $\sigma_p(R)$  from the radially independent measurement error. These results indicate that in order to obtain accurate virial mass estimates for the dwarf galaxies with fewer than  $\sim 1000$  observed stars (which is the case for most of the observed Milky Way dwarf galaxies), it is crucial that the measurement error on the line-of-sight velocities be subdominant to the intrinsic velocity dispersion.

The analogous figures for parameter sets I, II, and III are presented in the Supplementary Material. For parameter set III, which has the same DM inner slope and concentration as parameter set IV but is 100 times more massive, varying the measurement error has negligible effect on the inferred DM halo properties as expected. The results for parameter sets I and II (which have  $\gamma = 1$ ) are qualitatively

similar to the results for parameters sets III and IV (which have  $\gamma = 0$ ), although quantitatively different due to slightly larger values of the intrinsic velocity dispersion.

Similarly,  $\Delta v$  affects the recovery of the inner slope more for the less massive haloes than for the more massive ones. Fig. 5 shows the posterior  $\gamma$  distributions corresponding to the scans shown in Fig. 4. As  $\Delta v$  is increased,  $\gamma$  becomes increasingly unconstrained for the smaller sample sizes, whereas for the 10 000-star samples, increasing  $\Delta v$  appears to lead to a bias in the best-fitting value of  $\gamma$ . The analogous figures for parameter sets I, II, and III are presented in the Supplementary Material. For the more massive halo with the same inner slope (parameter set III), the posterior  $\gamma$  distributions are mostly insensitive to these variations in the measurement error; the results for parameter sets I and II are qualitatively similar to the cases of parameter sets III and IV, respectively.

### 3.3 Location of stars

In this section, we explore how the DM inference depends not only on how many stars are measured, but also on where the observed stars are within the dwarf galaxy. To study this effect, we start with our data sets of initial size  $n_{\text{stars}}$  and apply the following selection functions, then repeat our analysis on the resulting data sets (where  $R_{1/2}$  is the projected half-light radius):

- (i) Inner stars analysis: keep only the stars in the inner region, with projected radius  $R < R_{1/2}$ .
- (ii) Outer stars analysis: keep only the stars in the outer region, with projected radius  $R > R_{1/2}$ .

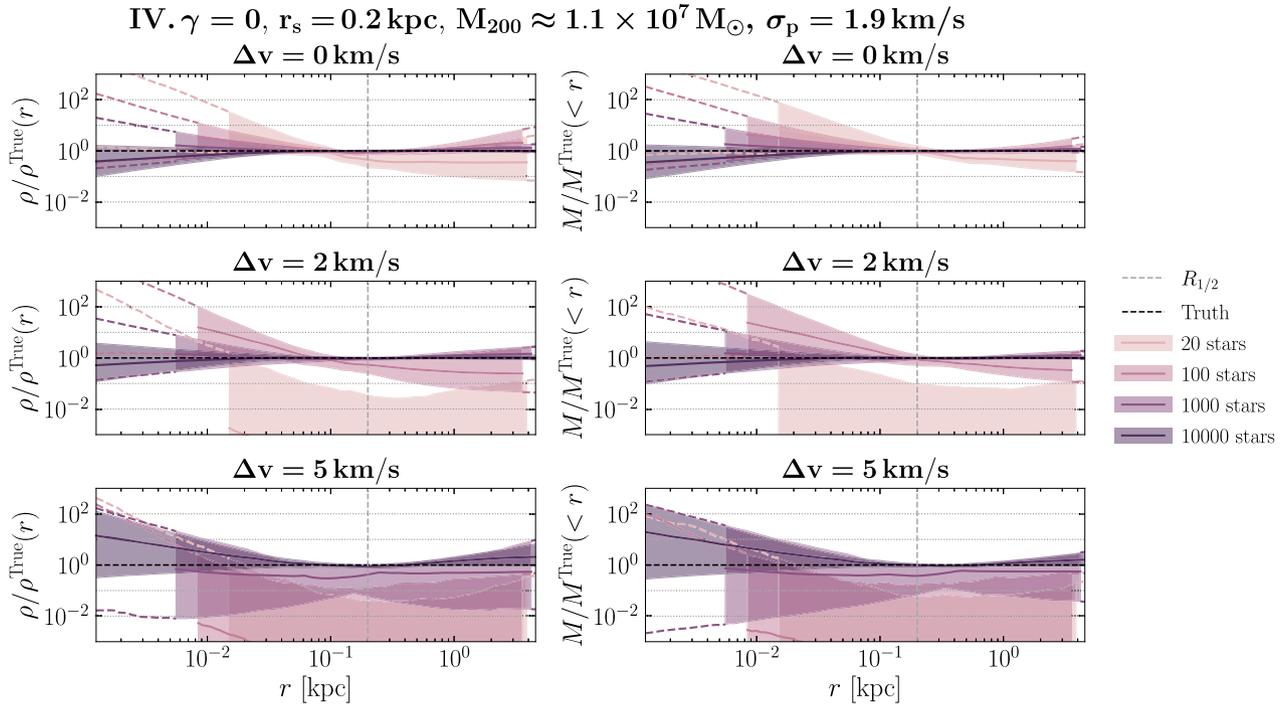
To account for the  $\sim 50$  per cent change in the number of stars from implementing these selection functions, we compare the results to ‘benchmark’ results on data sets with  $n_{\text{stars}}/2$  stars which are also generated from the original  $n_{\text{stars}}$ -star data sets, subsampled uniformly to preserve the radial probability distribution of the original data set. In doing so, we can compare the results for data sets that have approximately equal numbers ( $\sim n_{\text{stars}}/2$ ) but distinct spatial distributions of stars.

As before, we generate 10 independent data sets for each selection function. The distributions of the projected radius  $R$  as well as the 3d radius  $r$  for parameter set III with  $n_{\text{stars}} = 100$  (which is qualitatively representative of the distributions for all the parameter sets and sample sizes), for the three different selection functions, are presented in the Supplementary Material. We note that, because we implement the selection function on the projected radius, and  $r \geq R$  for all values of  $R$ , the  $R < R_{1/2}$  data sets extend slightly beyond a 3d radius of  $r = R_{1/2}$ . We test the effect of selection functions on data sets with initial sizes of  $n_{\text{stars}} = 100, 1000,$  and  $10,000$  stars for each of the four parameter sets. For the purpose of studying the effects of spatial distributions in the cleanest set-up, the studies presented in this section have been performed assuming  $\Delta v = 0 \text{ km s}^{-1}$ . As before, we perform a light profile fit to each data set prior to the Jeans fit.

In Fig. 6, we show the recovered DM density and enclosed mass profiles for the three different selection functions, for a particularly demonstrative example. This example is for parameter set IV, with an initial sample size of  $n_{\text{stars}} = 100$  stars; after applying each of the selection functions, we end up with a selected sample size of  $n_{\text{stars}}^{\text{sel}} \sim 50$  stars. From lightest to darkest colour, we show the results for the benchmark,  $R < R_{1/2}$ , and  $R > R_{1/2}$  data sets. Like before, the solid lines denote the median across the 10 realizations of the median recovered profiles, while the shaded bands depict the median of the 68 per cent containment ranges across the realizations. For ease of

<sup>6</sup>These values are chosen for reasonable comparison to current spectrographs such as Keck/DEIMOS (e.g. Martin et al. 2007; Simon & Geha 2007), Magellan/IMACS (e.g. Li et al. 2017; Simon et al. 2017, 2020), VLT/GIRAFFES+FLAMES (e.g. Pasquini et al. 2002; Li et al. 2018), and APOGEE (e.g. Majewski et al. 2017; Jönsson et al. 2020).

<sup>7</sup>We choose to present parameter set IV here because it has the smallest intrinsic velocity dispersion out of all of our parameter sets, and therefore is most drastically affected by increasing  $\Delta v$ .



**Figure 4.** Fractional recovered density profiles (left-hand panels) and enclosed mass profiles (right-hand panels) for parameter set IV, varying over the line-of-sight velocity measurement error  $\Delta v$  as well as the sample size. For each sample size, the solid line denotes the median (across our 10 independent realizations) of the median fractional recovered profiles, while the shaded band shows the median of the 68 per cent containment regions, plotted over the maximal radial range across all 10 data sets for that sample size. For the samples with fewer than 10000 stars, the median 68 per cent containment regions over the full radial range are shown bracketed by each pair of dashed lines in the colour corresponding to the sample size. Varying  $\Delta v$  has a particularly drastic effect on the smaller samples – for a sample size of 20 stars, a measurement error of  $\Delta v = 2$  km s $^{-1}$  is insufficient for recovering the DM density and enclosed mass profiles.

presentation, we choose in this case to show the solid line and band for each selection function from the median  $r_{\min}$  to the median  $r_{\max}$  across the realizations, where  $r_{\min}$  ( $r_{\max}$ ) is the galactocentric distance of the innermost (outermost) star in each individual realization. Beyond this radial range, the median 68 per cent containment ranges are shown by each pair of dashed lines in the colour corresponding to the selection function.

As expected, when the measured stars are all in the outer region of the dwarf, the DM profile is poorly constrained at small radii compared to the benchmark scenario. Conversely, when the measured stars are all in the inner region of the dwarf, the DM profile is poorly constrained at larger radii. Interestingly, for the  $R < R_{1/2}$  samples in this example, the DM profile is also typically less well-constrained at small radii; additionally, the density profile is biased high at small radii and low at large radii, to the extent that the total enclosed mass is also biased low (the recovered virial mass is  $M_{200} \sim 0.2^{+0.7}_{-0.1} \times 10^7 M_\odot$ , while the true value is  $M_{200} \sim 1.1 \times 10^7 M_\odot$ ). These biases, as well as the larger uncertainties on the DM profile in both the inner and outer regions, are present in spite of there being approximately twice as many stars within the half-light radius in the  $R < R_{1/2}$  data sets as in the benchmark data sets. In this particular example, the posterior  $\gamma$  distribution is unconstrained for all three selection functions due to the small size of the data set, so we do not recover a corresponding bias in  $\gamma$ .<sup>8</sup>

The specific behaviour of the results for the  $R < R_{1/2}$  samples noted in this example is not generic to all the variations we have tested –

in particular, for the data sets with larger selected sample size  $n_{\text{stars}}^{\text{sel}}$ , the bias in the DM density profile is less severe, and in some cases the median 68 per cent containment band on the inner density profile is slightly narrower than in the benchmark case (see Supplementary Material for results for an initial sample size of  $n_{\text{stars}} = 1000$ , i.e. for spatially selected data sets of size  $n_{\text{stars}}^{\text{sel}} \sim 500$ ).

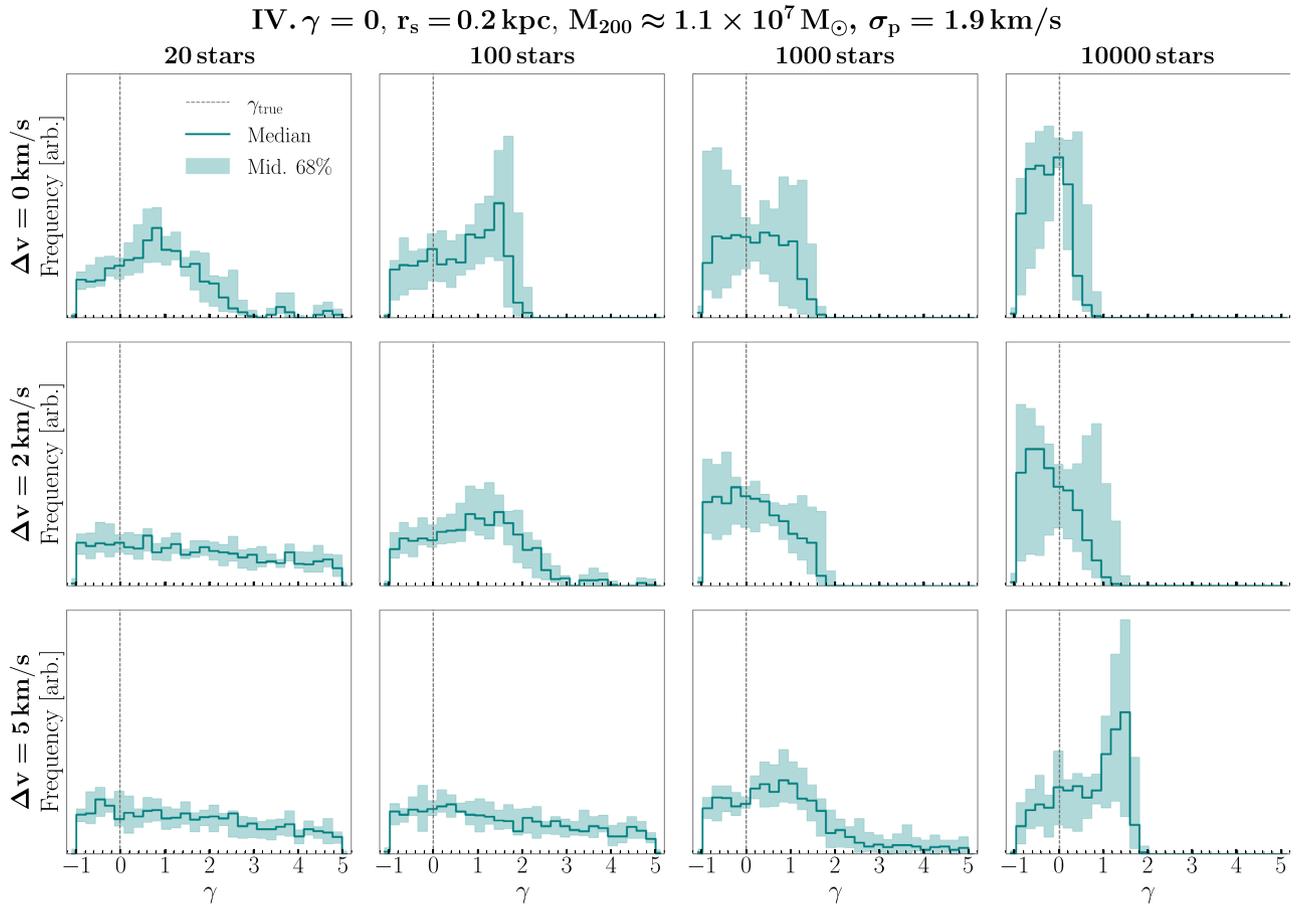
We can quantitatively compare the performance of the different selection functions, for different sample sizes  $n_{\text{stars}}^{\text{sel}}$ , by comparing the recovered virial mass estimates as well as the recovered  $J$ -factors (discussed more in Section 4), both tabulated in detail in the Supplementary Material. Across our four parameter sets, the results on spatial selection functions are the following:

(i) Inner stars analysis ( $R < R_{1/2}$ )

(a) For the smallest sample size  $n_{\text{stars}}^{\text{sel}} \sim 50$ , for all parameter sets, the inferred virial mass is systematically underestimated (inconsistent with the true value within  $1\sigma$  uncertainty for three of the four parameter sets). This becomes less severe as the sample size is increased, but across all four parameter sets for the larger sample sizes  $n_{\text{stars}}^{\text{sel}} \sim 500$  and  $n_{\text{stars}}^{\text{sel}} \sim 5000$ , the uncertainty on the estimated virial mass is consistently larger than for either the  $R > R_{1/2}$  data sets or the benchmark case, demonstrating that to achieve an accurate virial mass estimate, it is important to have measurements of outer stars.

(b) The behaviour of the posterior  $\gamma$  distribution varies across different sample sizes and different parameter sets – in some cases, the posterior  $\gamma$  distribution is biased high when the selection function is applied; in other cases, it is unchanged from the posterior distribution in the benchmark case. In all cases, the  $R < R_{1/2}$  selection function does not improve the

<sup>8</sup>For brevity, in this section we summarize in text the behaviour of the  $\gamma$  posteriors, when relevant, in place of showing additional figures.



**Figure 5.** Posterior distributions for the inner slope  $\gamma$ , varying over the line-of-sight velocity measurement error  $\Delta v$  as well as the sample size, shown for parameter set IV. These posteriors correspond to the results shown in Fig. 4. The lines (bands) show the median (middle 68 per cent) in each  $\gamma$  bin across the 10 realizations. Varying  $\Delta v$  has a drastic effect on the inference of  $\gamma$  in this case; in contrast, for the more massive halo of parameter set III, varying the measurement error has negligible effect (see Supplementary Material). All panels in this figure share the same vertical scale.

ability of the method to accurately constrain  $\gamma$ , relative to the benchmark case. Therefore, for the purpose of constraining  $\gamma$ , additional stars need to be measured across all radii.

(c) As we will discuss in Section 4, for the smallest sample size  $n_{\text{stars}}^{\text{sel}} \sim 50$ , for all parameter sets, the uncertainty on the  $J$ -factor estimate is larger than in the benchmark case. For the larger sample sizes, the uncertainty on the  $J$ -factor estimate is comparable to or slightly ( $\mathcal{O}(0.1 \text{ dex})$ ) smaller than in the benchmark case.

(ii) Outer stars analysis ( $R > R_{1/2}$ )

(a) For all parameter sets and all sample sizes  $n_{\text{stars}}^{\text{sel}}$ , the estimated virial mass is consistent with the true value, and the uncertainty on the virial mass estimate is comparable to or slightly smaller than in the benchmark case, demonstrating that having measurements of inner stars is not crucial to the recovery of the virial mass.

(b) For all parameter sets and all sample sizes  $n_{\text{stars}}^{\text{sel}}$ , the posterior  $\gamma$  distribution is comparable to (when the benchmark posterior distribution is already unconstrained) or less constrained than in the benchmark case.

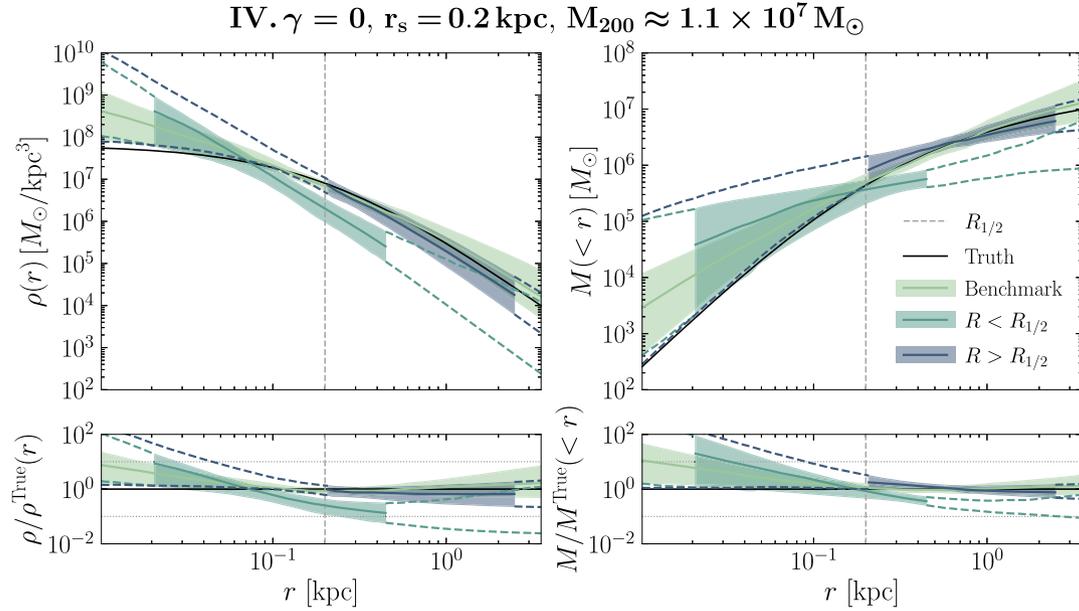
(c) As we will discuss in Section 4, for all parameter sets and all sample sizes  $n_{\text{stars}}^{\text{sel}}$ , the uncertainty on the  $J$ -factor estimate is comparable to or larger than in the benchmark case, indicating

that having measurements of inner stars is important for the purpose of constraining  $J$ -factors.

While the  $R > R_{1/2}$  data sets perform slightly better in terms of the uncertainty on the recovered virial mass relative to the two other selection functions, the improvement is marginal (see table in Supplementary Material for values). Therefore, based on the overall performance at inferring the full DM density profile and the inner slope  $\gamma$ , especially for the smallest samples, we find that it is crucial to have measurements of stars across the full radial distribution of the dwarf galaxy. Doing so allows the fit to anchor the DM distribution across the full radial range, and consistently results in comparable or better performance at accurately reconstructing both the inner and outer profile of the DM distribution, relative to the cases when the data consist purely of stars in either the inner or outer region of the system.

### 3.4 Embeddedness in DM halo

While our fiducial analyses have assumed for simplicity  $r_s = r_*$ , this may not be the case in realistic systems. To address this, we generate samples with our fiducial DM parameters but with varying degrees of embeddedness within the DM haloes, with  $r_* = \{0.2, 0.5, 1\} \times r_s$ , and repeat our analysis on these data sets. Intuitively, we expect that as the number of stars probing the innermost parts of the DM halo increases,



**Figure 6.** Inferred DM density profiles  $\rho(r)$  (left-hand panels) and corresponding enclosed mass profiles  $M(<r)$  (right-hand panels) for parameter set IV, starting with a sample size of  $n_{\text{stars}} = 100$  stars (resulting in selected samples of  $n_{\text{stars}}^{\text{sel}} \sim 50$  stars), with spatial selection functions applied. From lightest to darkest colour, the results are for the benchmark data sets, the data sets keeping only stars with  $R < R_{1/2}$ , and the data sets keeping only stars with  $R > R_{1/2}$ . For each selection function, the solid line denotes the median (across our 10 independent realizations) of the median recovered profiles, while the shaded band shows the median of the 68 per cent containment regions; these are plotted from the median  $r_{\text{min}}$  to the median  $r_{\text{max}}$  across the 10 realizations, where  $r_{\text{min}}(r_{\text{max}})$  is the galactocentric radius of the innermost(outermost) star. We also present the median 68 per cent containment regions over the full radial range of the benchmark samples, shown bracketed by each pair of dashed lines in the colour corresponding to the selection function. As expected, compared to the benchmark scenario, when the measured stars are all in the outer region of the dwarf, the DM profile is poorly constrained at small radii. Conversely, when the measured stars are all in the inner region of the dwarf, the DM profile is poorly constrained at larger radii. Moreover, the inner density profile is also less well-constrained for the  $R < R_{1/2}$  case than for the benchmark scenario, suggesting that in order to constrain the inner DM profile, it is important to have measured stars across the full radial distribution, and not only in the inner region.

the inference of the inner DM density profile and inner slope  $\gamma$  should become increasingly accurate. Indeed, as shown in Fig. 7, the  $\gamma$  inference is markedly improved as the ratio  $r_*/r_s$  is decreased, particularly for the larger sample sizes of  $\geq 1000$  stars.<sup>9</sup> Conversely, however, as the degree of embeddedness increases and there are consequently fewer stars at large  $r$ , the uncertainty on the inferred DM profile at large  $r$  increases. In the Supplementary Material, we show additional results demonstrating the increased uncertainty on the inferred  $M_{200}$  values, as well as the decreased uncertainty on the inferred  $J$ -factors, with higher degrees of embeddedness. Importantly, in spite of the decreased uncertainty on the inner DM profile, these results demonstrate that for the intermediate degree of embeddedness  $r_* = 0.5 \times r_s$ , we still need sample sizes of 10 000 stars to accurately and robustly constrain  $\gamma$  within our framework.

#### 4 IMPLICATIONS FOR INDIRECT DETECTION

In this section, we cast the results of our study into the context of indirect detection by calculating the inferred  $J$ -factors for the tests discussed in Section 3, using the public code CLUMPY (Charbonnier, Combet & Maurin 2012; Bonnivard et al. 2016a; Hütten, Combet & Maurin 2019) to the perform the  $J$ -factor computations (as defined

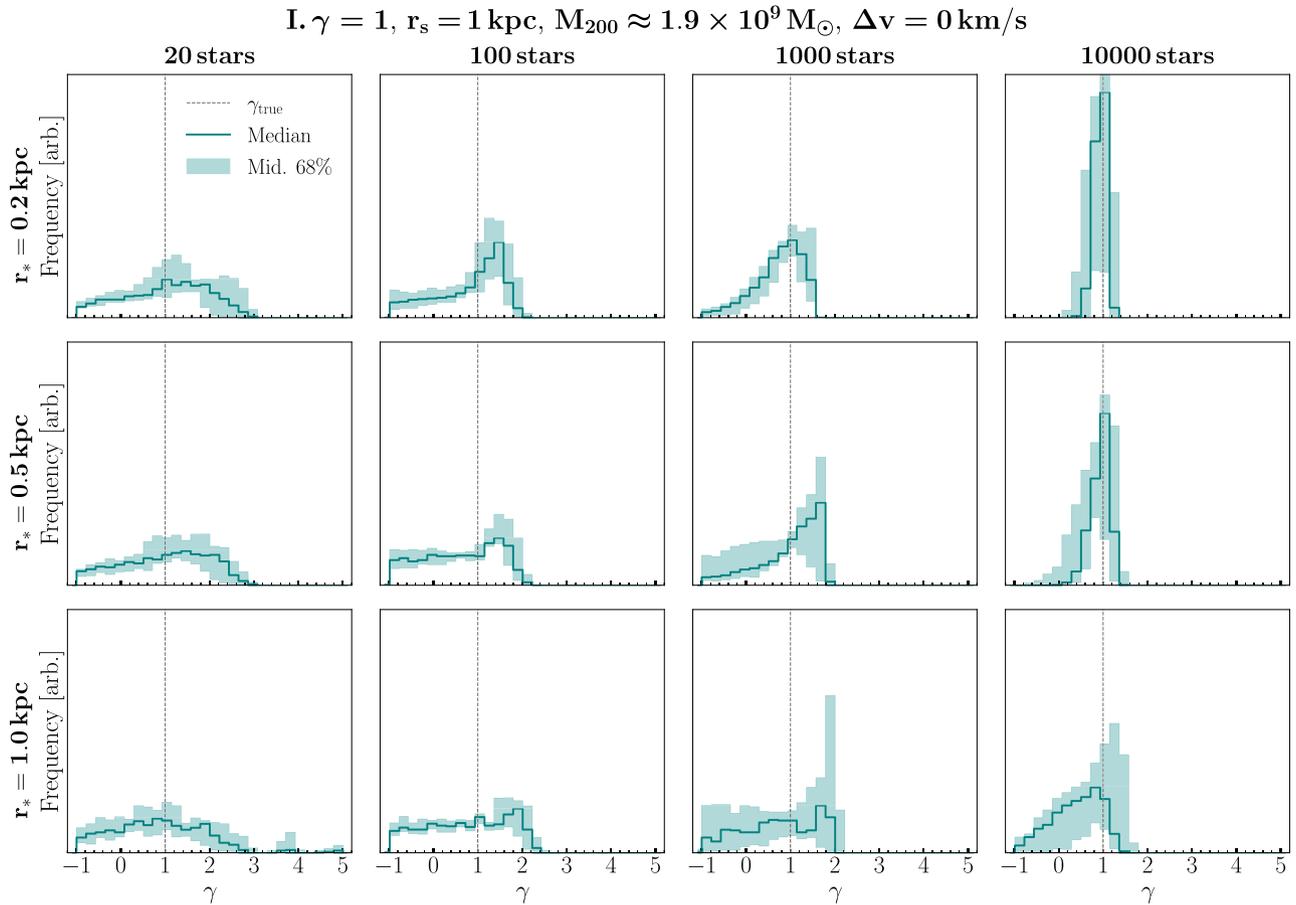
<sup>9</sup>For computational efficiency, the embeddedness tests are performed with a narrower prior range,  $\gamma \in [-1, 3]$ . As discussed in Section 2.5 as well as the Supplementary Material, this choice has negligible impact on the analysis results.

in equation 1). We examine the effects of sample size and line-of-sight velocity measurement error (Section 4.1), choices of priors in the Jeans analysis (Section 4.2), and spatial selection functions (Section 4.3) on the inferred  $J$ -factors. In Section 4.4, we discuss the implications of our findings on indirect detection results and make recommendations for future observations.

##### 4.1 Sample size and measurement error

First, we examine the effects of sample size and velocity measurement error,  $\Delta v$ , on the inferred  $J$ -factors. In the left column of Fig. 8, we show the inferred  $J$ -factors for parameter sets I (top) and III (bottom), for which  $M_{200} \sim 10^9 M_\odot$ , for the different sample sizes and values of  $\Delta v$ . For an individual scan, we evaluate the  $J$ -factor for each set of posterior parameters, assuming a distance of 50 kpc to the dwarf. Each data point in Fig. 8 shows the median across our 10 realizations of the median and middle 68 per cent containment range for the inferred values of  $\log_{10}[J(0.5^\circ)/(\text{GeV}^2 \text{ cm}^{-5})]$ . Within each cluster of three data points corresponding to a particular sample size, the blue circle, teal triangle, and green square show the results for  $\Delta v = 0, 2, 5 \text{ km s}^{-1}$ , respectively.

As expected, the uncertainties on the  $J$ -factor decrease as a function of increasing sample size. Additionally, the  $J$ -factors are nearly independent of  $\Delta v$ , which is expected for parameter sets I and III (see Section 3.2 for a discussion on the effects of  $\Delta v$ ). For parameter set I, our estimates of the  $J$ -factor are on average consistent with the true value for all sample sizes and values of  $\Delta v$ . For parameter set III, our estimates of the  $J$ -factor are systematically biased high,



**Figure 7.** Posterior distributions for the inner slope  $\gamma$ , varying over the degree of embeddedness of the stellar population within the DM halo, with  $r_* = \{0.2, 0.5, 1\} \times r_s$ , as well as the sample size, shown for parameter set I for the case of no measurement uncertainty ( $\Delta v = 0$  km s<sup>-1</sup>). The lines (bands) show the median (middle 68 per cent) in each  $\gamma$  bin across the 10 realizations. All panels in this figure share the same vertical scale. Increasing the degree of embeddedness improves the inference of  $\gamma$ , in particular for the 1000- and 10 000-star samples; we find qualitatively similar results for parameter sets II, III, and IV (see Supplementary Material).

although the median values are within a factor of 2 of the true values for the 1000- and 10 000-star samples – this is consistent with the inner density profiles being biased high for the smaller samples, as shown in Fig. 2. The typical values of the  $J$ -factor we recover for the different combinations of parameter set, sample size, and  $\Delta v$  are tabulated in the Supplementary Material.

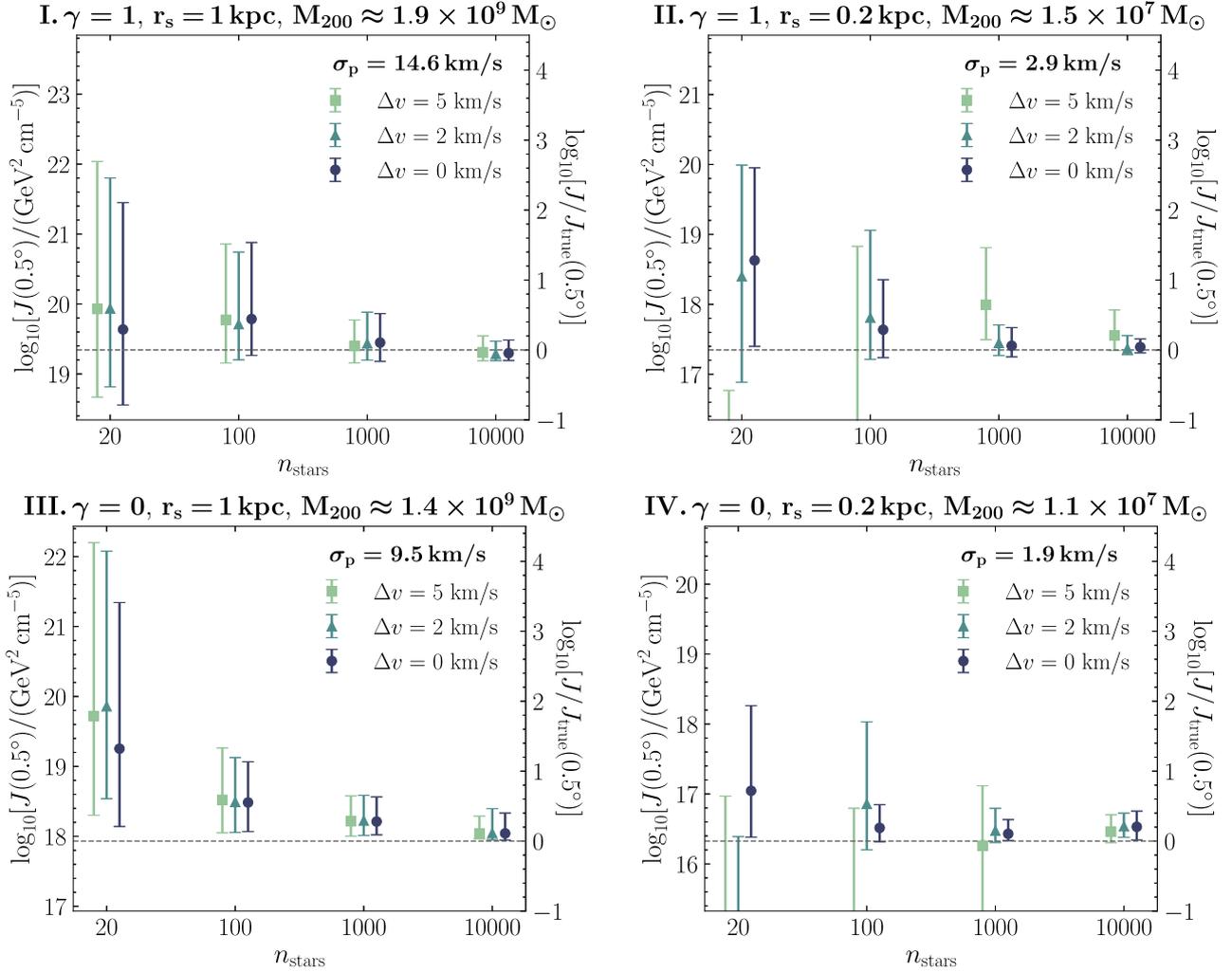
For parameter sets II and IV (shown in the top right and bottom right panels of Fig. 8, respectively), the  $J$ -factor estimates are highly sensitive to  $\Delta v$ , in a manner that is consistent with the results discussed in Section 3.2 (the corresponding fractional recovered density and enclosed mass profiles for parameter sets II and IV are shown in the Supplementary Material). In particular, the data points that extend below the range of the right-hand panels correspond to the cases of larger  $\Delta v$  where the recovered DM abundance is significantly underestimated.

#### 4.2 Dependence on priors

It is crucial to emphasize the dependence of the  $J$ -factor inference on the priors assumed for the DM halo parameters. The uncertainties on our inferred  $J$ -factors are notably larger than values commonly quoted in the literature for the data, such as the ones found in Geringer-Sameth et al. (2015b), hereafter GS15, which were used to derive the constraints on DM annihilation by the *Fermi*-

LAT collaboration in Albert et al. (2017), hereafter A17. The  $J$ -factors from GS15 are tabulated in the Supplementary Material for reference. For example, Ursa Major II, which has a sample size of 20 stars, is quoted to have a  $\sim \pm 0.5$  uncertainty on  $\log_{10}[J(0.5^\circ)/(\text{GeV}^2 \text{cm}^{-5})]$ , whereas on average, the uncertainties on  $\log_{10}[J(0.5^\circ)/(\text{GeV}^2 \text{cm}^{-5})]$  for our 20-star samples span  $\sim \pm 1$ –2 (when the values of  $\Delta v$  are sufficiently small for the DM to be recovered). This discrepancy is especially surprising because GS15 models the DM density distribution with the Hernquist/Zhao profile, which has two additional slope parameters compared to the gNFW model we use, and additionally models the velocity anisotropy – a model with more free parameters, combined with the added degeneracy between the anisotropy and mass profiles, should give rise to larger uncertainties on the inferred  $J$ -factors.

The primary source of this apparent discrepancy is that in this work, we have assumed wider prior ranges on the halo parameters than what was assumed in GS15 – in particular, the analysis in GS15 assumed a prior of  $0 \leq \gamma \leq 1.2$  on the inner DM slope. When we repeated our analysis assuming the same priors on  $\rho_0$ ,  $r_s$ , and  $\gamma$  as the ones used in GS15, the uncertainties on our  $J$ -factors decreased significantly. In Fig. 9, we show the median and middle 68 per cent range on  $\log_{10}[J(0.5^\circ)/(\text{GeV}^2 \text{cm}^{-5})]$  for each of our 10 different 20-star data sets from parameter set III (which is representative of the results for all parameter sets), assuming either our fiducial set-



**Figure 8.**  $J$ -factors as a function of the sample size and the measurement error on the line-of-sight velocities,  $\Delta v$ , for parameter sets I (top left), II (top right), III (bottom left), and IV (bottom right). We take the distance to the dwarf in each case to be 50 kpc and the angle of integration to be  $0.5^\circ$ . The results for  $\Delta v = 2 \text{ km s}^{-1}$  in the top (bottom) left-hand panel correspond to the recovered DM profiles shown in Fig. 1(2). For each realization of a given sample size and value of  $\Delta v$ , we build up a posterior  $J$ -factor distribution by calculating the  $J$ -factor for every set of posterior parameters, from which we can calculate the median and middle 68 per cent containment range of  $\log_{10}[J(0.5^\circ)/(\text{GeV}^2 \text{ cm}^{-5})]$  for that realization. Each data point shown here denotes the median across the 10 realizations of the median  $\log_{10}[J(0.5^\circ)/(\text{GeV}^2 \text{ cm}^{-5})]$ , and each set of error bars brackets the median across realizations of the 68 per cent containment. The data points that extend below the range of the right-hand panels correspond to the cases discussed in Section 3.2, where the recovered DM abundance is significantly underestimated. The values of all plotted  $J$ -factors are tabulated in the Supplementary Material.

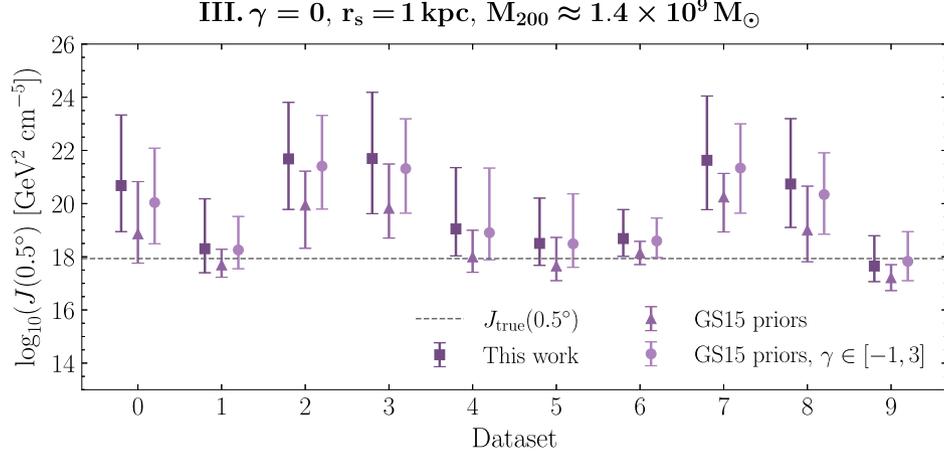
up (squares) or the priors from GS15 (triangles). Implementing the GS15 priors reduced the  $J$ -factor uncertainty in all 10 data sets, by as much as a factor of  $\sim 3$  in certain cases.

GS15 additionally takes the best-fitting Plummer radius from the literature and fixes it in their fit. Analogously, we have also repeated our analysis fixing our light-profile parameters to their best-fitting values while assuming the GS15 priors and found the results to be essentially unchanged from the case of GS15 priors without fixing light-profile parameters. Furthermore, GS15 truncates the  $J$ -factor integration at  $r_{\text{max}}$ , the galactocentric distance of the outermost star. We have tested this prescription as well, and found that it makes negligible difference to our values of the  $J$ -factor. This is expected, because the  $J$ -factor within the inner  $0.5^\circ$  is dominated by the most central regions of the DM halo, and is therefore insensitive to the outer truncation radius of the integration.

As an additional test, we set the priors on the normalization  $\rho_0$  and scale radius  $r_s$  for the DM profile to the GS15 priors, but rather

than using the GS15 prior of  $\gamma \in [0, 1.2]$  on the inner slope, we assume a wider prior range of  $\gamma \in [-1, 3]$ , which is equivalent to our fiducial prior range with the exclusion of the unphysical values of  $\gamma > 3$ . This directly tests how a wider prior range on  $\gamma$  affects the inferred  $J$ -factor. The results of this test are shown by the circles in Fig. 9, and are similar to our fiducial results (squares), indicating that the narrow prior range on  $\gamma$  is indeed what primarily drives our fit to reproduce the small  $J$ -factor uncertainties found in GS15. We have also verified that implementing the GS15 priors (with and without fixing the light profile parameters) on our 1000-star samples decreases the uncertainty on our estimated values of  $\log_{10}[J(0.5^\circ)/(\text{GeV}^2 \text{ cm}^{-5})]$  by a factor of  $\sim 2$ , making them broadly consistent with the uncertainties quoted in GS15 for the classical dwarfs.

An important takeaway from this exercise is that the  $J$ -factors inferred through the Jeans modelling procedure, for the currently accessible stellar sample sizes, depend sensitively on prior assump-



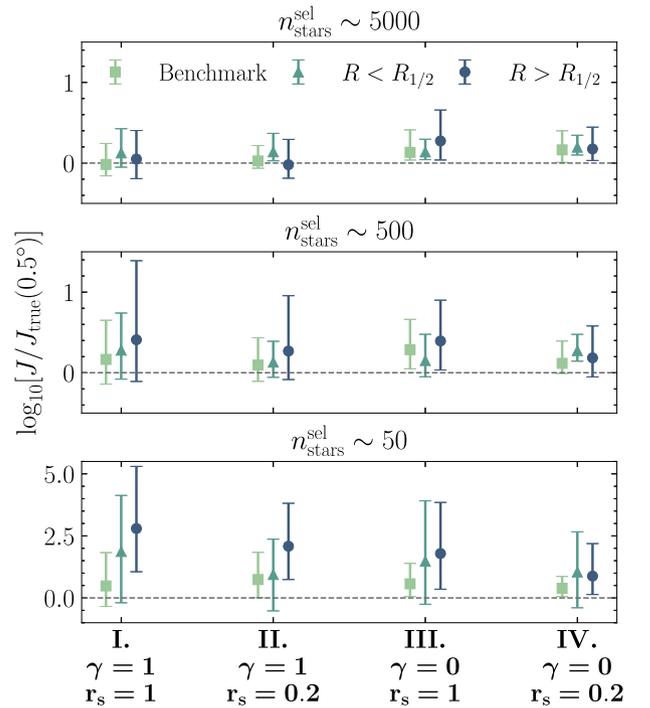
**Figure 9.** Comparing the  $J$ -factor results for each of the 10 20-star data sets in parameter set III: using our fiducial analysis set-up (dark purple squares), using the DM priors from Geringer-Sameth et al. (2015b) (medium purple triangles), and using the priors on the normalization  $\rho_0$  and scale radius  $r_s$  from Geringer-Sameth et al. (2015b) while setting the prior on the inner slope to be  $\gamma \in [-1, 3]$  (light purple circles). This demonstrates that for these small sample sizes, the results are highly prior-dependent, which is consistent with our findings that the DM profile and inner slope are poorly constrained for the data sets with limited statistics. This additionally demonstrates that the remarkably small uncertainties on the  $J$ -factors from Geringer-Sameth et al. (2015b) for the ultrafaint dwarfs may be driven by their narrow choice of prior on  $\gamma$ .

tions on  $\gamma$ , and therefore should be treated with caution. Motivated by the prior-dependence of  $J$ -factor estimates from Jeans analyses, a complementary method that has been proposed in the literature is a frequentist approach to deriving  $J$ -factors (Chiappo et al. 2017, 2019), which removes the prior-dependence but also loses the ability to construct full posterior probability distributions of the DM inner slope. A separate study (Ullio & Valli 2016) demonstrated that relaxing the prior assumptions on the DM density profile while keeping the dependence on the velocity anisotropy explicit can lead to a reduction of the minimal inferred  $J$ -factor.

### 4.3 Spatial selection

We can revisit the discussion of spatial selection functions detailed in Section 3.3, in the context of  $J$ -factors. In Section 3.3, we found that if we implemented a selection function of  $R > R_{1/2}$ , i.e. only included stars in the outer regions of the system, the resulting inferred DM density profile was more uncertain in the inner regions of the dwarf than in the benchmark scenario. We also found that if we implemented a selection function of  $R < R_{1/2}$ , i.e. only included stars in the inner regions of the system, the inferred DM density profile was more uncertain in the outer regions of the dwarf than in the benchmark scenario. Furthermore, for the  $R < R_{1/2}$  data sets, we found that the inner profile could be biased high, especially when the sample size was small. The degree of such biases and increased uncertainties on the DM density profile can be quantitatively captured by evaluating the  $J$ -factor. These results are shown in Fig. 10 and tabulated in detail in the Supplementary Material.

Overall, we find that for the data sets with  $\lesssim 50$  observed stars (comparable to the current sample sizes of ultrafaint dwarfs), observing more stars which are distributed across the full range of the radial distribution would have the most potential to decrease the uncertainty on estimates of the  $J$ -factors. This is demonstrated in the bottom panel of Fig. 10, in which the uncertainties on the  $J$ -factor are always smaller in the benchmark case (green squares) than for either of the other two cases (teal triangles and blue circles). For the systems with hundreds or thousands of observed stars, there is room for slight improvement on the accuracy of inferred  $J$ -factors



**Figure 10.**  $J$ -factors for a given selected sample size, for all parameter sets and spatial selection functions. From top to bottom, the panels show the results for  $n_{\text{stars}}^{\text{sel}} \sim 5000, 500, 50$ . From left to right, each cluster of three data points shows the results for parameter set I, II, III, and IV. Each data point denotes the median across the 10 realizations of the median  $\log_{10}[J(0.5^\circ)/(\text{GeV}^2 \text{cm}^{-5})]$ , and each set of error bars brackets the median of the 68 per cent containment across the realizations. For the smallest samples, the benchmark case consistently has smaller uncertainties than either the  $R > R_{1/2}$  or  $R < R_{1/2}$  cases. For some of the larger samples, the uncertainties are slightly reduced ( $\mathcal{O}(0.1)$  dex) smaller for  $R < R_{1/2}$  relative to the benchmark case. Note the different y-axis scale for the bottom panel.

by measuring more stars in the inner regions of the systems. This is demonstrated in the top and middle panels of Fig. 10, in which the uncertainties can be somewhat smaller ( $\mathcal{O}(0.1$  dex)) for the  $R < R_{1/2}$  data sets (teal triangles) than for the benchmark data sets (green squares) or  $R > R_{1/2}$  data sets (blue circles). As expected, the behaviour of the recovered  $J$ -factors is consistent with the ability of the Jeans modelling to accurately recover the inner density profile of the DM, as was discussed in Section 3.

#### 4.4 Dwarfs in need of more measurements

Within the literature, there are two approaches to dwarf galaxy indirect detection analyses – individual dwarfs may be analysed on their own (e.g. Essig et al. 2010; Baushev, Federici & Pohl 2012; Achterberg et al. 2015; Hooper & Linden 2015; Geringer-Sameth et al. 2015a; Zhao et al. 2018), or many systems may be stacked to obtain a more competitive limit on DM annihilation (e.g. Ackermann et al. 2011, 2015; Geringer-Sameth & Koushiappas 2011; Albert et al. 2017; Calore et al. 2018; Hoof et al. 2018). In both cases, achieving robust indirect detection results is dependent upon accurately estimating the  $J$ -factors for the dwarfs that dominate the limits. We emphasize that while the dwarfs that give rise to the strongest constraints on DM annihilation are among those with the largest  $J$ -factors, having a larger  $J$ -factor does not necessarily imply that the resulting limit from a given dwarf will be stronger, due to effects such as different levels of background contamination in different regions of the sky. In the following discussion, we will emphasize future observations which are important for obtaining more accurate estimates of the  $J$ -factors for the systems that dominate the A17 results.

As shown in Fig. 8 and detailed in the Supplementary Material, the typical uncertainty on  $\log_{10}[J(0.5^\circ)/(\text{GeV}^2 \text{cm}^{-5})]$  from our analysis is  $\sim \pm 1$ – $2$  for 20-star systems and  $\sim \pm 0.5$ – $1$  for 100-star systems (excluding the cases of small intrinsic dispersion and large  $\Delta v$  where the fit drastically underestimates the abundance of DM), as opposed to  $\sim \pm 0.5$  and  $\sim \pm 0.2$ , respectively, from GS15. We determined in Section 4.2 that this discrepancy may be due to different prior choices on  $\gamma$ . To test the effect of larger  $J$ -factor uncertainties on the resulting indirect detection constraints on DM annihilation, we can use the likelihood functions provided in A17<sup>10</sup> to derive limits assuming different values of the  $J$ -factor uncertainty. Similarly to A17, we use equation 3 of Ackermann et al. (2015) to profile over the  $J$ -factor uncertainty.

The three dwarfs from A17 that provide the strongest limits in the mass range relevant for the DM interpretation of the Galactic Center Excess (GCE) are Ursa Major II, Ursa Minor, and Draco. We first focus on Ursa Major II, which has a sample size of 20 stars. We find that increasing the uncertainty on  $\log_{10}[J(0.5^\circ)/(\text{GeV}^2 \text{cm}^{-5})]$  from 0.4 (which was assumed in the A17 analysis) to 1 weakens the limit by a factor of  $\sim 5$ – $8$  in the 10–100 GeV mass range for the  $b\bar{b}$  annihilation channel, resulting in a limit that no longer excludes or is in tension with the regions of parameter space consistent with the GCE from Gordon & Macias (2013), Daylan et al. (2016), Calore et al. (2015a), and Abazajian et al. (2014). Similarly, for Draco and Ursa Minor ( $\sim 300$  observed stars each), we find that increasing the uncertainty on  $\log_{10}[J(0.5^\circ)/(\text{GeV}^2 \text{cm}^{-5})]$  from their assumed values in A17 of 0.1 and 0.2, respectively, to 0.5(1) weakens the limit by a factor of  $\sim 2$ (10). We note that a factor of  $\sim 2$  weakening of the

strongest dwarf limits is sufficient to significantly reduce the tension with the DM interpretation of the GCE.

This demonstrates that for the current observed sample sizes, the dwarf galaxy indirect detection limits can be highly sensitive to the assumed priors for the inner DM slope  $\gamma$ . In order to derive robust indirect detection constraints from the dwarf galaxies, it is crucial to increase the number of observed stars in order to obtain more robust estimates of their  $J$ -factors. In particular, we emphasize the importance of increasing the sample sizes for Ursa Major II, Ursa Minor, and Draco, which dominate the indirect detection limits. Our results in Sections 3.3 and 4.3 suggest that measuring more stars spanning the entire spatial extent of the galaxies would be most effective at achieving more accurate estimates of their  $J$ -factors (see bottom panel of Fig. 10). If sample sizes are increased beyond  $\gtrsim 500$  stars, our findings suggest that focusing on measuring more stars in the inner regions of the dwarfs may provide additional constraining power on their  $J$ -factors (see top two panels of Fig. 10).

## 5 CONCLUSIONS

In this paper, we performed a systematic study of the spherical Jeans analysis method in the context of inferring the DM content in dwarf galaxies. We focused on simulated data for spherical, isotropic systems, and assessed the performance of the method at accurately recovering the overall dark matter density profile, the virial mass, and the inner slope of the dark matter density profile. Our primary conclusions are the following:

(i) For parameter sets I and III, which describe  $M_{200} \sim 10^9 M_\odot$  haloes (intrinsic velocity dispersion  $\sim 10$ – $15 \text{ km s}^{-1}$ ) with inner density slopes of  $\gamma = 1$  and  $\gamma = 0$ , respectively, we find that the virial mass we recover is always consistent with the true value, and is increasingly accurate as the sample size is increased. However, the inner profile of the DM density distribution is less well-constrained – for samples with  $\lesssim 1000$  stars, the posterior distributions on the inner slope  $\gamma$  are typically unconstrained, and there is no statistical preference for a cuspy or cored profile. We recover statistical evidence for the true (cuspy/cored) profile only for samples with 10 000 stars. For these parameter sets, the results are generally insensitive to varying the measurement error of the line-of-sight velocity,  $\Delta v$ , over the range  $\Delta v = 0$ – $5 \text{ km s}^{-1}$ .

(ii) For parameter sets II and IV, which describe  $M_{200} \sim 10^7 M_\odot$  haloes with inner density slopes of  $\gamma = 1$  and  $\gamma = 0$ , respectively, we find that the virial mass estimates depend sensitively on  $\Delta v$ , particularly for the samples with fewer stars. For parameter set II (intrinsic velocity dispersion  $\sim 3 \text{ km s}^{-1}$ ), the inferred virial mass for the 20- and 100-star samples is consistent with zero when  $\Delta v = 5 \text{ km s}^{-1}$ . Similarly, for parameter set IV (intrinsic velocity dispersion  $\sim 2 \text{ km s}^{-1}$ ), the recovered virial mass for the 20- and 100-star samples is consistent with zero for the cases of  $\Delta v = 2, 5 \text{ km s}^{-1}$ . This is rectified when the sample size is increased to 1000 or more stars.

(iii) From our study of spatial selection functions, we conclude based on the overall performance at inferring the DM density profile, the inner slope  $\gamma$ , and the virial mass that it is crucial to have measurements of stars across the full radial distribution of the dwarf galaxy, especially for the smallest samples. Doing so allows the fit to anchor the DM distribution across the full radial range, and consistently results in comparable or better performance at accurately reconstructing both the inner and outer profile of the DM distribution, relative to the cases where the data consist purely of stars in either the inner or outer region of the system. For systems with  $\lesssim 50$  observed

<sup>10</sup>[http://www-glast.stanford.edu/pub\\_data/1203/](http://www-glast.stanford.edu/pub_data/1203/)

stars, measuring more stars across the full radial extent can reduce the uncertainties on  $\log_{10}[J(0.5^\circ)/(\text{GeV}^2 \text{cm}^{-5})]$  by a factor of  $\sim 3$  compared to measuring the same number of stars only within the half-light radius.

(iv) Cast in the context of indirect detection, we find that for the 20-star samples across all parameter sets (in the cases of sufficiently small  $\Delta v$  for the DM to be recovered), the median  $1\sigma$  uncertainty on  $\log_{10}[J(0.5^\circ)/(\text{GeV}^2 \text{cm}^{-5})]$  across our 10 realizations is  $\sim \pm 1-2$ , in contrast with the uncertainties of  $\sim \pm 0.5$  quoted for some of the current ultrafaint dwarf measurements (with  $\sim 20$  stars) in [GS15](#), which were used to derive the dwarf galaxy constraints on DM annihilation in [A17](#). We find that this discrepancy may be driven by the more restrictive prior ranges for the DM profile parameters in [GS15](#) – in particular the prior range on the inner slope  $\gamma$  – and note that the resulting indirect detection results should be interpreted with this prior-dependence in mind.

(v) For the purpose of achieving more accurate, less prior-dependent  $J$ -factor estimates for the systems that dominate the indirect detection results presented in [A17](#), we identify Ursa Major II, Ursa Minor, and Draco as the dwarf galaxies that would most benefit from more stars being measured. Our preliminary analyses show that if we assume the typical  $J$ -factor uncertainties that we find in our work, the DM annihilation limits for these systems may be weakened to the degree of significantly affecting their implications on the DM interpretation of the GCE.

In our study, we have focused on the case of spherical, isotropic systems with the goal of understanding the limitations of Jeans analyses even in the absence of challenges that are known to complicate studies using this method, such as background contamination (e.g. [Bonnivard, Maurin & Walker 2016b](#); [Ichikawa et al. 2017, 2018](#); [Horigome et al. 2020](#)), the effect of assuming equilibrium for systems which are not in equilibrium ([El-Badry et al. 2017](#)), the effect of non-spherical models ([Bonnivard et al. 2015a](#); [Hayashi et al. 2016](#); [Klop et al. 2017](#)), the degeneracy between the enclosed mass and velocity anisotropy (e.g. [Merrifield & Kent 1990](#); [Wilkinson et al. 2002](#); [Lokas & Mamon 2003](#); [De Lorenzi et al. 2009](#); [Read & Steger 2017](#); [Genina et al. 2020](#)), and the presence of potentially large fractions of binary stars in the dwarf galaxies (e.g. [McConnachie & Côté 2010](#); [Minor et al. 2010](#); [Geha et al. 2013](#); [Spencer et al. 2017, 2018](#); [Minor et al. 2019](#)).

With regard to the core-cusp problem, we have found that even for the idealized systems we consider, and a relatively simple three-parameter halo model, the Jeans modelling method is severely limited in its ability to constrain the inner slope  $\gamma$  of the dark matter density profile. A crucial reason behind this is that there are degeneracies between the three parameters that describe our DM profiles. The fact that  $\gamma$  is difficult to pinpoint is consistent with previous Jeans modelling results in the literature (e.g. [Walker et al. 2009](#); [Read et al. 2018](#); [Genina et al. 2020](#)); we have additionally determined that, in order to constrain  $\gamma$  within this framework, it is necessary to measure  $\sim 10\,000$  stars within a single dwarf galaxy – even for the higher degree of embeddedness of  $r_* = 0.5 \times r_s$  – which is not practical within the near future. We therefore need to search for alternative approaches to addressing the core-cusp problem using Jeans analysis methods. Jointly fitting to multiple dwarf galaxies simultaneously is a promising method for leveraging moderately sized data sets to achieve better constraints on  $\gamma$ . Due to computational challenges, exploring this avenue systematically is beyond the scope of this paper.

The parameter degeneracy that limits our ability to reconstruct  $\gamma$  is distinct from the well-known mass-anisotropy degeneracy which

plagues Jeans analyses, for which a number of proposed solutions exist in the literature: using higher order moments of the velocity distribution ([Merrifield & Kent 1990](#); [Richardson & Fairbairn 2013, 2014](#); [Read & Steger 2017](#); [Kaplinghat, Valli & Yu 2019](#); [Genina et al. 2020](#)) and incorporating proper motion measurements of stars ([Strigari, Bullock & Kaplinghat 2007a](#); [Lazar & Bullock 2020](#)) are among the methods that have been demonstrated to ameliorate the mass-anisotropy degeneracy. It is worth exploring whether or not these methods would also lead to improved constraints on the inner slope of the DM density profile, the answer to which is not intuitively obvious. [Alvarez et al. \(2020\)](#) recently used the framework described in [Read & Steger \(2017\)](#), which parametrizes the DM density profile as a multiply-broken power law and employs higher order moments, to derive  $J$ -factors for the classical dwarfs. They obtained  $J$ -factor estimates which are consistent with the ones from [GS15](#), but with reduced uncertainties.

Given the swaths of stellar data that will become available from ongoing and future surveys (e.g. [Takada et al. 2014](#); [Ivezić, Ž., et al. 2019](#); [Li et al. 2019](#); [The MSE Science Team et al. 2019](#)), it is crucial to improve our ability to infer dark matter distributions from stellar data in dwarf galaxies. Doing so would allow us to more accurately interpret indirect detection results, more conclusively probe the core-cusp problem, and potentially uncover significant implications in our understanding of the particle nature of dark matter.

## ACKNOWLEDGEMENTS

We are particularly grateful to M. Geha and M. Lisanti for their helpful insight on the topic. We also thank P. Hopkins, A. Ji, E. Kirby, J. Read, J. Simon, and M. Walker for helpful discussions. LJC thanks S. Mishra-Sharma for fruitful discussions and moral support. LJC is supported by a Paul & Daisy Soros Fellowship and an NSF Graduate Research Fellowship under Grant Number DGE-1656466. LN is supported by the DOE under Award Number DESC0011632, the Sherman Fairchild fellowship, the University of California Presidential Fellowship, and a Carnegie Fellowship in Theoretical Astrophysics. The work presented in this paper was performed on computational resources managed and supported by Princeton Research Computing, a consortium of groups including the Princeton Institute for Computational Science and Engineering (PICSciE) and the Office of Information Technology’s High Performance Computing Center and Visualization Laboratory at Princeton University.

## DATA AVAILABILITY

The simulated data sets underlying this article were generated using the public code STARSAMPLER using the parameters described in the article, and will be shared on reasonable request to the authors.

## REFERENCES

- Abazajian K. N., Canac N., Horiuchi S., Kaplinghat M., 2014, *Phys. Rev. D*, 90, 023526
- Achterberg A., van Beekveld M., Beenakker W., Caron S., Hendriks L., 2015, *JCAP*, 12, 013
- Ackermann M. et al., 2011, *Phys. Rev. Lett.*, 107, 241302
- Ackermann M. et al., 2012, *ApJ*, 761, 91
- Ackermann M. et al., 2015, *Phys. Rev. Lett.*, 115, 231301
- Aghanim N. et al., 2020, *A&A*, 641, A6

- Ajello M. et al., 2016, *ApJ*, 819, 44
- Albert A. et al., 2017, *ApJ*, 834, 110
- Alvarez A., Calore F., Genina A., Read J., Serpico P. D., Zaldivar B., 2020, *J. Cosmol. Astropart. Phys.*, 9, 004
- Amorisco N. C., Evans N. W., 2012, *MNRAS*, 419, 184
- Atwood W. B. et al., 2009, *ApJ*, 697, 1071
- Battaglia G., Helmi A., Tolstoy E., Irwin M., Hill V., Jablonka P., 2008, *ApJ*, 681, L13
- Baushev A. N., Federici S., Pohl M., 2012, *Phys. Rev. D*, 86, 063521
- Behroozi P. S., Wechsler R. H., Conroy C., 2013, *ApJ*, 770, 57
- Benitez-Llambay A., Frenk C., 2020, *MNRAS*, 498, 4887
- Benitez-Llambay A., Frenk C. S., Ludlow A. D., Navarro J. F., 2019, *MNRAS*, 488, 2387
- Binney J., 1980, *MNRAS*, 190, 873
- Binney J., Mamon G. A., 1982, *MNRAS*, 200, 361
- Binney J., Tremaine S., 2008, *Galactic Dynamics*, 2nd edn. Princeton Series in Astrophysics, Princeton University Press
- Blumenthal G. R., Faber S. M., Flores R., Primack J. R., 1986, *ApJ*, 301, 27
- Bonnivard V. et al., 2015b, *MNRAS*, 453, 849
- Bonnivard V., Combet C., Maurin D., Walker M. G., 2015a, *MNRAS*, 446, 3002
- Bonnivard V., Hütten M., Nezri E., Charbonnier A., Combet C., Maurin D., 2016a, *Comput. Phys. Commun.*, 200, 336
- Bonnivard V., Maurin D., Walker M. G., 2016b, *MNRAS*, 462, 223
- Bose S. et al., 2019, *MNRAS*, 486, 4790
- Buchner J. et al., 2014, *A&A*, 564, A125
- Bullock J. S., Boylan-Kolchin M., 2017, *ARA&A*, 55, 343
- Calore F., Cholis I., McCabe C., Weniger C., 2015a, *Phys. Rev.*, D91, 063003
- Calore F., Cholis I., Weniger C., 2015b, *J. Cosmol. Astropart. Phys.*, 1503, 038
- Calore F., Serpico P. D., Zaldivar B., 2018, *J. Cosmol. Astropart. Phys.*, 1810, 029
- Chang L. J., Lisanti M., Mishra-Sharma S., 2018, *Phys. Rev.*, D98, 123004
- Charbonnier A. et al., 2011, *MNRAS*, 418, 1526
- Charbonnier A., Combet C., Maurin D., 2012, *Comput. Phys. Commun.*, 183, 656
- Chiappo A., Cohen-Tanugi J., Conrad J., Strigari L., Anderson B., Sanchez-Conde M., 2017, *MNRAS*, 466, 669
- Chiappo A., Cohen-Tanugi J., Conrad J., Strigari L., 2019, *MNRAS*, 488, 2616
- Daylan T., Finkbeiner D. P., Hooper D., Linden T., Portillo S. K. N., Rodd N. L., Slatyer T. R., 2016, *Phys. Dark Univ.*, 12, 1
- Dejonghe H., Merritt D., 1992, *ApJ*, 391, 531
- Del Popolo A., Le Delliou M., 2017, *Galaxies*, 5, 17
- Del Popolo A., Pace F., 2016, *Ap&SS*, 361, 162
- Despali G., Sparre M., Vegetti S., Vogelsberger M., Zavala J., Marinacci F., 2019, *MNRAS*, 484, 4563
- de Lorenzi F. et al., 2009, *MNRAS*, 395, 76
- Di Cintio A., Brook C. B., Macciò A. V., Stinson G. S., Knebe A., Dutton A. A., Wadsley J., 2014, *MNRAS*, 437, 415
- Di Mauro M., Hou X., Eckner C., Zaharijas G., Charles E., 2019, *Phys. Rev. D*, 99, 123027
- Dutton A. A., Macciò A. V., Buck T., Dixon K. L., Blank M., Obreja A., 2019, *MNRAS*, 486, 655
- El-Badry K., Wetzel A. R., Geha M., Quataert E., Hopkins P. F., Kereš D., Chan T. K., Faucher-Giguère C.-A., 2017, *ApJ*, 835, 193
- Essig R., Sehgal N., Strigari L. E., Geha M., Simon J. D., 2010, *Phys. Rev. D*, 82, 123503
- Feroz F., Hobson M. P., Bridges M., 2009, *MNRAS*, 398, 1601
- Fitts A. et al., 2019, *MNRAS*, 490, 962
- Flores R. A., Primack J. R., 1994, *ApJ Lett.*, 427, L1
- Gallagher J., Madsen G., Reynolds R., Grebel E. K., Smecker-Hane T., 2003, *ApJ*, 588, 326
- Garrison-Kimmel S., Bullock J. S., Boylan-Kolchin M., Bardwell E., 2017, *MNRAS*, 464, 3108
- Geha M. et al., 2013, *ApJ*, 771, 29
- Genina A. et al., 2020, *MNRAS*, 498, 144
- Geringer-Sameth A., Koushiappas S. M., 2011, *Phys. Rev. Lett.*, 107, 241303
- Geringer-Sameth A., Walker M. G., Koushiappas S. M., Kopusov S. E., Belokurov V., Torrealba G., Evans N., 2015a, *Phys. Rev. Lett.*, 115, 081101
- Geringer-Sameth A., Koushiappas S. M., Walker M., 2015b, *ApJ*, 801, 74
- González-Samaniego A., Bullock J. S., Boylan-Kolchin M., Fitts A., Elbert O. D., Hopkins P. F., Kereš D., Faucher-Giguère C.-A., 2017, *MNRAS*, 472, 4786
- Goodenough L., Hooper D., 2009, preprint ([arXiv: 0910.2998](https://arxiv.org/abs/0910.2998))
- Gordon C., Macias O., 2013, *Phys. Rev. D*, 88, 083521
- Grand R. J. J. et al., 2017, *MNRAS*, 467, 179
- Grcevich J., Putman M. E., 2009, *ApJ*, 696, 385
- Hayashi K., Ichikawa K., Matsumoto S., Ibe M., Ishigaki M. N., Sugai H., 2016, *MNRAS*, 461, 2914
- Hernquist L., 1990, *ApJ*, 356, 359
- Hlozek R. et al., 2012, *ApJ*, 749, 90
- Hoof S., Geringer-Sameth A., Trotta R., 2020, *JCAP*, 02, 012
- Hooper D., Linden T., 2015, *J. Cosmol. Astropart. Phys.*, 09, 016
- Horigome S.-i., Hayashi K., Ibe M., Ishigaki M. N., Matsumoto S., Sugai H., 2020, *MNRAS*, 499, 3320
- Hütten M., Combet C., Maurin D., 2019, *Comput. Phys. Commun.*, 235, 336
- Ichikawa K., Ishigaki M. N., Matsumoto S., Ibe M., Sugai H., Hayashi K., Horigome S.-i., 2017, *MNRAS*, 468, 2884
- Ichikawa K., Horigome S.-i., Ishigaki M. N., Matsumoto S., Ibe M., Sugai H., Hayashi K., 2018, *MNRAS*, 479, 64
- Ivezic Ž. et al., 2019, *ApJ*, 873, 111
- Jeans J. H., 1915, *MNRAS*, 76, 70
- Jönsson H. et al., 2020, *AJ*, 160, 120
- Kaplinghat M., Valli M., Yu H.-B., 2019, *MNRAS*, 490, 231
- Kass R. E., Raftery A. E., 1995, *J. Am. Stat. Assoc.*, 90, 773
- Klop N., Zandanel F., Hayashi K., Ando S., 2017, *Phys. Rev. D*, 95, 123012
- Lazar A. et al., 2020, *MNRAS*, 497, 2393
- Lazar A., Bullock J. S., 2020, *MNRAS*, 493, 5825
- Li T. et al., 2019, *BAAS*, 51, 252
- Lisanti M., Mishra-Sharma S., Rodd N. L., Safdi B. R., 2018a, *Phys. Rev. Lett.*, 120, 101101
- Lisanti M., Mishra-Sharma S., Rodd N. L., Safdi B. R., Wechsler R. H., 2018b, *Phys. Rev. D*, 97, 063005
- Liu M. S., 2019, *Inference of Dark Matter Density Profiles of Dwarf Spheroidal Galaxies via Distribution Functions*, Carnegie Mellon University
- Li T. S. et al., 2017, *ApJ*, 838, 8
- Li T. S. et al., 2018, *ApJ*, 857, 145
- Lokas E. L., Mamon G. A., 2003, *MNRAS*, 343, 401
- Majewski S. R. et al., 2017, *AJ*, 154, 94
- Mamon G. A., Boué G., 2010, *MNRAS*, 401, 2433
- Mamon G. A., Lokas E. L., 2005, *MNRAS*, 363, 705
- Mamon G. A., Biviano A., Boué G., 2013, *MNRAS*, 429, 3079
- Martin N., Ibata R., Chapman S., Irwin M., Lewis G., 2007, *MNRAS*, 380, 281
- Mashchenko S., 2015, preprint ([arXiv:1504.08273](https://arxiv.org/abs/1504.08273))
- Mashchenko S., Couchman H. M. P., Wadsley J., 2006, *Nature*, 442, 539
- McConnachie A. W., Côté P., 2010, *ApJ*, 722, L209
- Merrifield M. R., Kent S. M., 1990, *AJ*, 99, 1548
- Merritt D., 1985, *AJ*, 90, 1027
- Minor Q. E., Martinez G., Bullock J., Kaplinghat M., Trainor R., 2010, *ApJ*, 721, 1142
- Minor Q. E., Pace A. B., Marshall J. L., Strigari L. E., 2019, *MNRAS*, 487, 2961
- Moliné Á., Sánchez-Conde M. A., Palomares-Ruiz S., Prada F., 2017, *MNRAS*, 466, 4974
- Moore B., 1994, *Nature*, 370, 629
- Navarro J. F., Eke V. R., Frenk C. S., 1996, *MNRAS*, 283, L72
- Navarro J. F., Frenk C. S., White S. D. M., 1997, *ApJ*, 490, 493
- Oh S.-H. et al., 2015, *AJ*, 149, 180
- Oh S.-H., Brook C., Governato F., Brinks E., Mayer L., de Blok W. J. G., Brooks A., Walter F., 2011, *AJ*, 142, 24
- Pasquini L. et al., 2002, *The Messenger*, 110, 1
- Pieri L., Lavallo J., Bertone G., Branchini E., 2011, *Phys. Rev. D*, 83, 023518

- Plummer H. C., 1911, *MNRAS*, 71, 460  
 Pontzen A., Governato F., 2012, *MNRAS*, 421, 3464  
 Read J., Walker M., Steger P., 2018, *MNRAS*, 481, 860  
 Read J. I., Gilmore G., 2005, *MNRAS*, 356, 107  
 Read J. I., Steger P., 2017, *MNRAS*, 471, 4541  
 Read J. I., Agertz O., Collins M. L. M., 2016, *MNRAS*, 459, 2573  
 Read J. I., Iorio G., Agertz O., Fraternali F., 2017, *MNRAS*, 467, 2019  
 Richardson T., Fairbairn M., 2013, *MNRAS*, 432, 3361  
 Richardson T., Fairbairn M., 2014, *MNRAS*, 441, 1584  
 Robles V. H., Kelley T., Bullock J. S., Kaplinghat M., 2019, *MNRAS*, 490, 2117  
 Salucci P., Burkert A., 2000, *ApJ*, 537, L9  
 Sánchez-Conde M. A., Prada F., 2014, *MNRAS*, 442, 2271  
 Sawala T. et al., 2016, *MNRAS*, 457, 1931  
 Shen X., Hopkins P., Necib L., 2021, *MNRAS*, 506, 4421  
 Simon J. D. et al., 2017, *ApJ*, 838, 11  
 Simon J. D. et al., 2020, *ApJ*, 892, 137  
 Simon J. D., Geha M., 2007, *ApJ*, 670, 313  
 Spekkens K., Giovanelli R., Haynes M. P., 2005, *AJ*, 129, 2119  
 Spencer M. E., Mateo M., Walker M. G., Olszewski E. W., McConnachie A. W., Kirby E. N., Koch A., 2017, *AJ*, 153, 254  
 Spencer M. E., Mateo M., Olszewski E. W., Walker M. G., McConnachie A. W., Kirby E. N., 2018, *AJ*, 156, 257  
 Spergel D. N., Steinhardt P. J., 2000, *Phys. Rev. Lett.*, 84, 3760  
 Strigari L. E., Bullock J. S., Kaplinghat M., 2007a, *ApJ*, 657, L1  
 Strigari L. E., Koushiappas S. M., Bullock J. S., Kaplinghat M., 2007b, *Phys. Rev.*, D75, 083526  
 Strigari L. E., Koushiappas S. M., Bullock J. S., Kaplinghat M., Simon J. D., Geha M., Willman B., 2008, *ApJ*, 678, 614  
 Strigari L. E., Frenk C. S., White S. D. M., 2017, *ApJ*, 838, 123  
 Swaters R. A., Madore B. F., van den Bosch F. C., Balcells M., 2003, *ApJ*, 583, 732  
 Takada M. et al., 2014, *PASJ*, 66, R1  
 Tegmark M., Zaldarriaga M., 2002, *Phys. Rev. D*, 66, 103508  
 Tegmark M., Zaldarriaga M., 2009, *Phys. Rev. D*, 79, 083530  
 The MSE Science Team, 2019, preprint (arXiv:1904.04907)  
 Tollet E. et al., 2016, *MNRAS*, 456, 3542  
 Tulin S., Yu H.-B., 2018, *Phys. Rept.*, 730, 1  
 Ullio P., Valli M., 2016, *J. Cosmol. Astropart. Phys.*, 07, 025  
 Walker M. G., Peñarrubia J., 2011, *ApJ*, 742, 20  
 Walker M. G., Mateo M., Olszewski E. W., Peñarrubia J., Evans N. W., Gilmore G., 2009, *ApJ*, 704, 1274  
 Walter F., Brinks E., de Blok W. J. G., Bigiel F., Kennicutt R. C. J., Thornley M. D., Leroy A., 2008, *AJ*, 136, 2563  
 Wang J., Bose S., Frenk C. S., Gao L., Jenkins A., Springel V., White S. D. M., 2020, *Nature*, 585, 39  
 Wechsler R. H., Tinker J. L., 2018, *ARA&A*, 56, 435  
 Wilkinson M., Kleya J., Evans N., Gilmore G., 2002, *MNRAS*, 330, 778  
 Zechlin H.-S., Manconi S., Donato F., 2018, *Phys. Rev. D*, 98, 083022  
 Zhao H., 1996, *MNRAS*, 278, 488  
 Zhao Y., Bi X.-J., Yin P.-F., Zhang X.-M., 2018, *Chin. Phys. C*, 42, 025102

## SUPPORTING INFORMATION

Supplementary data are available at *MNRAS* online.

**Figure S1.** (Left) An example triangle plot of the posterior DM parameters from a scan of a 10 000-star sample from parameter set I, with  $\Delta v = 0 \text{ km s}^{-1}$ .

**Figure S2.** Posterior  $\gamma$  distributions for 100-star samples from parameter sets I (top row) and III (bottom row), with  $\Delta v = 2 \text{ km s}^{-1}$ .

**Figure S3.** Posterior  $\gamma$  distributions for the 20-star samples for parameter sets I (top row) and III (bottom row).

**Figure S4.** Example results from jointly analysing five 20-star samples from parameter set I, for which  $\gamma = 1$ .

**Figure S5.** An example light profile fit for a single 20-star sample.

**Figure S6.** Same as Fig. C1, but for a 1000-star sample.

**Figure S7.** Distributions of the projected radii (left-hand panel) and 3d galactocentric radii (right-hand panel) in our mock data sets for parameter set I.

**Figure S8.** Fractional recovered density profiles (left-hand panels) and enclosed mass profiles (right-hand panels) for parameter set III, varying over the line-of-sight velocity measurement error  $\Delta v$  as well as the sample size.

**Figure S9.** Same as Fig. D2, but for parameter set I.

**Figure S10.** Same as Fig. D2, but for parameter set II.

**Figure S11.** Posterior distributions for the inner slope  $\gamma$ , varying over the line-of-sight velocity measurement error  $\Delta v$  as well as the sample size, for parameter set III.

**Figure S12.** Same as Fig. D5, but for parameter set I, with vertical scale adjusted for presentation.

**Figure S13.** Same as Fig. D5, but for parameter set II, with vertical scale adjusted for presentation.

**Figure S14.** Inferred fractional virial mass,  $M_{200}/M_{200}^{\text{True}}$ , as a function of the sample size and the degree of embeddedness of the stellar population within the DM halo, with  $r_* = \{0.2, 0.5, 1\} \times r_s$ , for parameter sets I (top left), II (top right), III (bottom left), and IV (bottom right), for the case of no measurement uncertainty ( $\Delta v = 0 \text{ km s}^{-1}$ ).

**Figure S15.**  $J$ -factors as a function of the sample size and the degree of embeddedness of the stellar population within the DM halo, with  $r_* = \{0.2, 0.5, 1\} \times r_s$ , for parameter sets I (top left), II (top right), III (bottom left), and IV (bottom right), for the case of no measurement uncertainty ( $\Delta v = 0 \text{ km s}^{-1}$ ).

**Figure S16.** Posterior distributions for the inner slope  $\gamma$ , varying over the degree of embeddedness of the stellar population within the DM halo, with  $r_* = \{0.2, 0.5, 1\} \times r_s$ , as well as the sample size, shown for parameter set III for the case of no measurement uncertainty ( $\Delta v = 0 \text{ km s}^{-1}$ ).

**Figure S17.** Same as Fig. D10, but for parameter set II, with vertical scale adjusted for presentation.

**Figure S18.** Same as Fig. D10, but for parameter set IV, with vertical scale adjusted for presentation.

**Figure S19.** Histograms of the projected radius  $R$  (left-hand panel) and the 3d radius  $r$  (right-hand panel) for parameter set III starting with a sample size of  $n_{\text{stars}} = 100$  stars, resulting in selected samples of  $n_{\text{stars}}^{\text{sel}} \sim 50$  stars, for the three different spatial selection functions.

**Figure S20.** Inferred DM density profiles  $\rho(r)$  (left-hand panels) and corresponding enclosed mass profiles  $M(<r)$  (right-hand panels) for parameter set IV, starting with a sample size of  $n_{\text{stars}} = 1000$  stars (resulting in selected samples of  $n_{\text{stars}}^{\text{sel}} \sim 500$  stars), with spatial selection functions applied.

**Table S1.** List of dwarf galaxies used in Albert et al. (2017) ordered by decreasing  $J$ -factor.

**Table S2.** Inferred values of the virial mass  $M_{200}$  and  $J$ -factor for the different parameter sets, sample sizes, and values of  $\Delta v$ .

**Table S3.** Inferred values of the virial mass  $M_{200}$  and  $J$ -factor for the different parameter sets, selected sample sizes, and spatial selection functions, with  $\Delta v = 0$  in all cases.

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

This paper has been typeset from a  $\text{\TeX}/\text{\LaTeX}$  file prepared by the author.