This article has been accepted for publication in Monthly Notices of the Royal Astronomical Society. ©: 2023 The Authors. Published by Oxford University Press on behalf of the Royal Astronomical Society. All rights reserved.

Link to article on OUP website:

https://academic.oup.com/mnras/article/522/2/3058/7129015

Advance Access publication 2023 April 18

Downloaded from https://academic.oup.com/mnras/article/522/2/3058/7129015 by guest on 18 June 2024



Salvador Cardona-Barrero <sup>®</sup>, <sup>1,2</sup>★ Giuseppina Battaglia <sup>®</sup>, <sup>1,2</sup> Carlo Nipoti <sup>®3</sup> and Arianna Di Cintio <sup>®1,2</sup>

Accepted 2023 April 10. Received 2023 April 10; in original form 2022 December 20

#### ABSTRACT

An anticorrelation between the central density of the dark matter (DM) halo ( $\rho_{150, DM}$ ) and the pericentric distances ( $r_p$ ) of the Milky Way's (MW's) dwarf spheroidal galaxies (dSphs) has been reported in the literature. The existence and origin of such anticorrelation are, however, controversial, one possibility being that only the densest dSphs can survive the tidal field towards the centre of our Galaxy. In this work, we place particular emphasis on quantifying the statistical significance of such anticorrelation, by using available literature data in order to explore its robustness under different assumptions on the MW gravitational potential, and for various derivations of  $\rho_{150}$  and  $r_p$ . We consider models in which the MW is isolated and has low ( $8.8 \times 10^{11} \, \mathrm{M}_{\odot}$ ) and high ( $1.6 \times 10^{12} \, \mathrm{M}_{\odot}$ ) halo masses, respectively, as well as configurations in which the MW's potential is perturbed by a Large Magellanic Cloud (LMC) infall. We find that, while data generally support models in which the dSphs' central DM density decreases as a function of their pericentric radius, this anticorrelation is statistically significant at  $3\sigma$  level only in  $\sim$ 12 per cent of the combinations of  $\rho_{150}$  and  $r_p$  explored. Moreover, including the impact of the LMC's infall on to the MW weakens or even washes away this anticorrelation, with respect to models in which the MW is isolated. Our results suggest that the strength and existence of such anticorrelation are still debatable: exploring it with high-resolution simulations including baryonic physics and different DM flavours will help us to understand its emergence.

**Key words:** galaxies: dwarf – galaxies: kinematics and dynamics – Local Group.

## 1 INTRODUCTION

Thanks to several observational campaigns and theoretical work on dynamical modelling, the dark matter (DM) content and orbital parameters of the Milky Way's (MW) dwarf spheroidal galaxies (dSphs)<sup>1</sup> are now relatively well known (e.g. Battaglia & Nipoti 2022; Battaglia et al. 2022, and references therein). The inner DM density of dSphs is often quantified by measuring  $\rho_{150}$ , defined as the DM density at a distance of 150 pc from the centre of the dwarf (e.g. Read, Walker & Steger 2019, hereafter R19). The orbit of each dSph, which depends on the gravitational potential assumed for the MW, can be described by different parameters, among which the pericentric radius  $r_p$  (minimum distance of the dwarf centre of mass from the Galactic Centre), which gives an indication of the importance of tidal effects.

Using pericentres inferred from the second data release (DR2) from the *Gaia* mission (Prusti et al. 2016; Gaia Collaboration 2018), Kaplinghat, Valli & Yu (2019, hereafter K19) claimed that a sample of nine MW dSphs exhibits an anticorrelation between  $\rho_{150}$  and  $r_p$ . This anticorrelation, which is also found for DM sub-haloes in some cosmological simulations (Robles & Bullock 2021; Genina et al.

2022), might be a consequence of survivor bias, i.e. the fact that lower density satellites on small pericentre orbits have not survived the tidal field of the MW (Hayashi, Chiba & Ishiyama 2020; Genina et al. 2022). Alternatively, it may be a signature of self-interacting dark matter, as gravothermal core collapse (Balberg, Shapiro & Inagaki 2002) is accelerated in dwarfs that undergo tidal stripping (Nishikawa, Boddy & Kaplinghat 2020), leading to larger central densities.

The existence and the strength of such anticorrelation for MW satellite galaxies are a matter of debate (Hayashi et al. 2020; Genina et al. 2022; Hayashi et al. 2022). There is no evidence of anticorrelation when samples of ultra-faint dwarfs (UFDs) are analysed (K19). It is also to be considered that while the mass (and average density) of these pressure-supported galaxies is determined with the highest precision within the half-light radius (e.g. Wolf et al. 2010) or 1.8 times the half-light radius (Errani, Peñarrubia & Walker 2018), where the mass-anisotropy degeneracy is minimized, mass and density estimates at other locations carry larger uncertainties.

In this paper, we revisit the question of the possible anticorrelation between  $\rho_{150}$  and  $r_{\rm p}$  of MW dSphs with a quantitative approach, by performing a systematic statistical analysis. In particular, we address the question of whether the result is sensitive to the set of literature estimates of  $\rho_{150}$  and  $r_{\rm p}$  considered. For the latter quantity, several new determinations have recently been obtained using the more accurate and precise data from the *Gaia* early third data release (eDR3; Gaia Collaboration 2021) and also taking into account self-

<sup>&</sup>lt;sup>1</sup>Instituto de Astrofísica de Canarias, Calle Vía Láctea s/n, E-38206 La Laguna, Tenerife, Spain

<sup>&</sup>lt;sup>2</sup>Departamento de Astrofísica, Universidad de La Laguna, Avda. Astrofísico Fco. Sánchez, E-38205 La Laguna, Tenerife, Spain

<sup>&</sup>lt;sup>3</sup>Dipartimento di Fisica e Astronomia 'Augusto Righi', Università di Bologna, Via Gobetti 93/2, I-40129 Bologna, Italy

<sup>\*</sup> E-mail: scardona@iac.es

<sup>&</sup>lt;sup>1</sup>Following Simon (2019), we adopt the nomenclature 'dSph'/'UFD' for the galaxies brighter/fainter than absolute *V*-band magnitude  $M_V = -7.7$ .

**Table 1.** Values of the pericentric radius and central DM density of each dSph (columns) compiled from the literature (rows); we refer the reader to Section 2 for the labelling of each model.  $\rho_{150}$  and  $r_p$  are given as the 50th percentiles of the distribution of values, and the lower and upper error bars bracket the 16th and 84th percentiles, respectively.

	Draco	UMi	Carina	Sextans	LeoI	LeoII	Sculptor	Fornax
Pericentres: $r_{\rm p}$ (kpc)								
F18	$42.0^{+16.0}_{+11.0}$	$44.0^{+12.0}_{+10.0}$	$103.0^{+8.0}_{+23.0}$	$79.0^{+9.0}_{+8.0}$	$63.0^{+221.0}_{+47.0}$	$67.0^{+154.0}_{+52.0}$	$69.0^{+10.0}_{+9.0}$	$100.0^{+28.0}_{+33.0}$
B22_L	$51.7^{+4.0}_{+6.0}$	$48.9^{+3.0}_{+3.0}$	$106.7^{+6.0}_{+5.0}$	$74.5^{+4.0}_{+6.0}$	$46.6^{+30.0}_{+26.0}$	$115.5^{+88.0}_{+59.0}$	$63.6^{+4.0}_{+3.0}$	$89.4^{+31.0}_{+26.0}$
B22_H	$37.6^{+4.0}_{+4.0}$	$34.9^{+3.0}_{+3.0}$	$102.8^{+10.0}_{+32.0}$	$64.0^{+5.0}_{+6.5}$	$35.0^{+24.0}_{+20.0}$	$69.0^{+64.0}_{+29.0}$	$48.7^{+4.0}_{+4.0}$	$56.2^{+22.0}_{+15.0}$
B22_LMC	$100.0^{+22.0}_{+19.0}$	$70.4^{+8.5}_{+5.0}$	$98.0^{+14.0}_{+24.0}$	$71.6^{+4.5}_{+6.0}$	$40.1^{+29.0}_{+24.0}$	$105.4^{+115.0}_{+50.0}$	$47.7^{+3.4}_{+3.4}$	$91.7^{+32.0}_{+25.0}$
P22	$40.4^{+6.5}_{+5.4}$	$41.8^{+5.3}_{+4.5}$	$114.4_{+11.8}^{+49.7}$	$82.8_{+4.0}^{+3.7}$	$42.9_{+23.2}^{+28.9}$	$54.3_{+31.6}^{+55.7}$	$55.0^{+5.5}_{+5.2}$	$85.2^{+38.6}_{+29.3}$
P22_LMC	$58.0^{+11.4}_{+9.5}$	$55.7^{+8.4}_{+7.0}$	$77.9_{+17.9}^{+24.1}$	$82.2_{+4.3}^{+3.8}$	$47.5^{+30.9}_{+24.0}$	$61.4_{+34.7}^{+62.3}$	$44.9_{+3.9}^{+4.3}$	$76.7^{+43.1}_{+27.0}$
Central densities: $\rho_{150}$ (10 <sup>7</sup> M $_{\odot}$ kpc <sup>-3</sup> )								
R19	$23.6_{+2.9}^{+2.9}$	$15.3^{+3.5}_{+3.2}$	$11.6^{+2.0}_{+2.2}$	$12.8^{+3.4}_{+2.9}$	$17.7^{+3.3}_{+3.4}$	$18.4^{+1.7}_{+1.6}$	$14.9^{+2.8}_{+2.3}$	$7.9_{+1.9}^{+2.7}$
H20	$23.5^{+12.8}_{+6.3}$	$23.8^{+38.6}_{+7.2}$	$10.9_{+3.2}^{+8.2}$	$5.2^{+3.6}_{+2.3}$	$26.4^{+22.3}_{+9.1}$	$20.2^{+12.7}_{+6.1}$	$21.4^{+12.6}_{+6.3}$	$12.2^{+3.2}_{+2.3}$
K19_NFW	$21.7_{+2.2}^{+2.7}$	$25.2^{+2.9}_{+4.5}$	$10.3^{+1.1}_{+0.9}$	$11.0^{+2.9}_{+1.8}$	$15.1^{+3.4}_{+2.4}$	$17.1_{+3.8}^{+2.4}$	$17.1_{+2.2}^{+2.1}$	$7.5^{+2.0}_{+1.4}$
K19_ISO	$21.3_{+4.7}^{+5.4}$	$25.4^{+6.1}_{+5.7}$	$5.7^{+3.2}_{+1.7}$	$8.5^{+5.0}_{+3.5}$	$14.1_{+4.5}^{+5.5}$	$13.5^{+4.2}_{+1.7}$	$16.1_{+3.3}^{+2.9}$	$3.4_{+1.3}^{+1.7}$

consistently the impact of the infall of a massive Large Magellanic Cloud (LMC) on to the MW, which can strongly affect the orbital history of MW dSphs (e.g. Patel et al. 2020; Battaglia et al. 2022; Pace, Erkal & Li 2022). This article is structured in the following way: In Section 2, we introduce the set of MW dSphs considered, and the sets of literature estimates for  $\rho_{150}$  and  $r_p$ ; in Section 3, we present the statistical approaches undertaken, including a method of general validity for data sets with asymmetric error bars, and test them on mock data sets; and in Section 4, we discuss our results and present our conclusions in Section 5.

## 2 SAMPLE AND DATA SETS

We focus on eight of the MW dSphs, specifically the objects in common between the studies of K19, R19, and Hayashi et al. (2020, hereafter H20). K19 also included Canes Venatici I, but, as we will see later, the exclusion of this system does not change the conclusions on the existence of the anticorrelation between  $\rho_{150}$  and  $r_p$ .

The values of  $\rho_{150}$  are taken from K19, R19, and H20.<sup>2</sup> In R19, they were determined with GRAVSPHERE, which solves the spherical Jeans equation for the projected line-of-sight (l.o.s.) velocity dispersion profile of the stellar component and also fits two higher order 'virial shape parameters' (Merrifield & Kent 1990; Richardson & Fairbairn 2014; Read & Steger 2017); it uses a non-parametric form for the enclosed mass as a function of radius, M(< r). Also, K19 solved the spherical Jeans equation for the projected l.o.s. velocity dispersion profile but considered one 'virial shape parameter'; for the DM halo density profile, they considered separately a NFW (Navarro, Frenk & White 1996) model and a cored isothermal model; therefore, we have two sets of  $\rho_{150}$  for the K19 study (which we label K19\_NFW and K19\_ISO, respectively). In H20, the mass modelling was performed by solving the axisymmetric Jeans equations for the second moment of the l.o.s. velocity distribution at a given projected 2D position; the DM halo was modelled with a generalized Hernquist (Hernquist 1990) profile, therefore probing cuspy and cored models and considering non-spherical DM haloes;  $\rho_{150}$  is calculated along

the major axis of the DM halo, which is assumed to have the same orientation as the stellar component.

We also consider several determinations of the pericentric radii for the MW dSphs: the *Gaia* DR2-based determinations for an MW of mass  $0.8 \times 10^{12} \, \mathrm{M}_\odot$  by Fritz et al. (2018, hereafter F18), as used in K19³; the *Gaia* eDR3-based pericentric distances by Battaglia et al. (2022, hereafter B22) in three gravitational potentials, two of them with isolated MW of mass  $8.8 \times 10^{11} \, \mathrm{M}_\odot$  ('Light', hereafter B22\_L) and  $1.6 \times 10^{12} \, \mathrm{M}_\odot$  ('Heavy', hereafter B22\_H), respectively, and one with an  $8.8 \times 10^{11} \, \mathrm{M}_\odot$  MW perturbed by a  $1.5 \times 10^{11} \, \mathrm{M}_\odot$  LMC (hereafter B22\_LMC); and the *Gaia*-eDR3 based values by Pace et al. (2022, hereafter P22) in an isolated  $1.3 \times 10^{12} \, \mathrm{M}_\odot$  MW (hereafter, labelled as P22) and in an MW + LMC potential (P22\_LMC), having the LMC mass of  $1.38 \times 10^{11} \, \mathrm{M}_\odot$ .

Table 1 and Fig. 1 (black points with error bars) present all the sets of  $\rho_{150}$  and  $r_p$  used in this work. As can be gathered from Fig. 1 in a qualitative way, the relationship between  $r_p$  and  $\rho_{150}$  varies, depending on the pericentric radii adopted as well as on the determinations of the central DM densities. In the next sections, we quantify this visual impression with a quantitative statistical analysis.

## 3 STATISTICAL ANALYSIS

In this section, we outline the methodology we use to quantify a possible anticorrelation between  $\rho_{150}$  and  $r_p$ .

As in K19, we model the relationship between  $\rho_{150}$  and  $r_p$  as a power-law (PL) of the form

$$\log_{10} \left( \frac{\rho_{150}}{10^7 \,\mathrm{M}_{\odot} \,\mathrm{kpc}^{-3}} \right) = q + m \log_{10} \left( \frac{r_{\mathrm{p}}}{\mathrm{kpc}} \right), \tag{1}$$

with q and m being the logarithmic zero-point and slope, respectively. One of the aspects that we wish to take into account is that usually the error bars in the sets of  $\rho_{150}$  and  $r_{\rm p}$  are highly asymmetric. However, this asymmetry has not been considered when fitting equation (1) in the literature. For example, when quantifying the relation between  $\rho_{150}$  and  $r_{\rm p}$ , K19 symmetrize the errors by averaging the upper and lower errors, and shift the median to the mid-point (see

<sup>&</sup>lt;sup>2</sup>While for R19 and H20 the values are tabulated, in K19 they are not and we have digitized their fig. 2.

<sup>&</sup>lt;sup>3</sup>K19 also demonstrate that their results are essentially unchanged when using the values for a twice as massive MW.

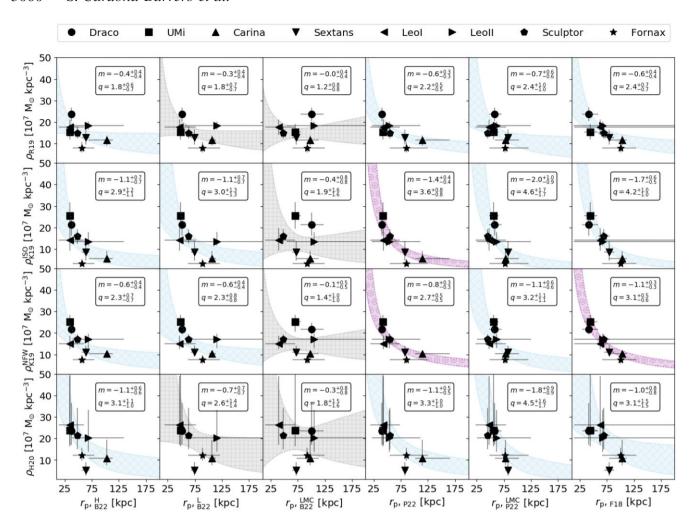


Figure 1. Central DM density versus pericentric distance for eight classical dSphs. From top to bottom, we show the densities as derived in R19, K19 (assuming isothermal and NFW DM density profiles, respectively), and H20. The pericentres are (from left to right): B22 assuming a heavy MW, a light MW and a light MW with the inclusion of the LMC; P22 without and with LMC; and F18 with a light MW. The details of each data set can be found in Section 2. The shaded region depicts the 68 per cent confidence-level region of the PL fit between the central DM density and the pericentric distance, obtained with the percentile fitting method (see Appendix B). In each panel, we show the recovered median logarithmic slope (m) and logarithmic zero-point (q). The uncertainties of the parameters are indicated with the 16th and 84th percentiles. The colours of the reported relations have been chosen to represent the statistical significance of the logarithmic slope: whether it is compatible with 0 within  $1\sigma$  (grey), between  $1\sigma$  and  $3\sigma$  (light blue), or within more than  $3\sigma$  (magenta).

their appendix C). In this work, we explore two different methods that retain the information of the asymmetries of the errors.

The gathered data (Table 1 and Fig. 1) are provided in the literature as the median and the 16th and 84th percentiles ( $x_{50th}$ ,  $x_{16th}$ , and  $x_{84th}$ , respectively) of an underlying distribution. Our first approach ('simulating errors' method), described in Appendix A, consists in resimulating the error distribution via the reported percentiles: assuming a probability distribution, we fit its corresponding cumulative distribution function (CDF) to the percentiles of each pair of pericentre and central density values. Then, via random sampling from the fitted probability distribution we can obtain different random realizations of the original data. The second approach ('percentile fitting' method), described in Appendix B, consists in treating each percentile as a random variable and modelling its probability distribution; thus, a fully Bayesian approach is possible via order statistics. Finally, for comparison, we have also applied the fitting method used in K19, which we will refer to as 'symmetrized errors' method.

In order to test the performance of these three methods, we applied them to mock data sets, with a radial coverage and noise level mimicking those of the data in the literature. We generate a set of eight pericentric distances  $r_{\rm p}^*$  with similar CDF as the one obtained from the pericentres of P22; the corresponding DM central densities  $\rho_{150}^*$  are obtained from equation (1) assuming q=3 and m=-1. We introduce noise by simulating N=20 samples extracted from a log-normal distribution centred on  $r_{\rm p}^*$  and  $\rho_{150}^*$  with logarithmic variance  $\sigma_{\rm LN}^2$ . For each data point, we derive the 16th, 50th, and 84th percentiles. In this way, we are able to obtain samples of mock data with asymmetric errors similar to those in the literature.

In Table 2, we show the average uncertainty ( $<\sigma_{\rm data}>$ ) for each of the sets of  $r_{\rm p}$  and  $\rho_{150}$  analysed, obtained from fitting the CDF of a log-normal distribution to the reported percentiles. To mimic these typical uncertainties, we then choose three different values of  $\sigma_{\rm LN}=\{0.1,\ 0.25,\ 0.5\}$  (which hereafter we also refer to as 'noise levels' 1,

<sup>4</sup>Note that the choice of N = 20 does not meet the conditions described in Appendix B; i.e. 20 is not a large number of samples and the fraction  $P_i(N - 1)/100$  is not an integer. Thus, we may not expect an excellent performance when using the percentile method.

**Table 2.** Typical uncertainties in the quantities used in this work. Col. 1 lists the set of  $r_{\rm p}$  and  $\rho_{150}$  being considered; col. 2 ( $<\sigma_{\rm data}>$ ) gives the corresponding average uncertainty for that given quantity.

Data set	$<\sigma_{ m data}>$
ρ <sub>R19</sub>	$0.190 \pm 0.059$
ρ <sub>H20</sub>	$0.458 \pm 0.129$
$ ho_{ m K19}^{ m NFW}$	$0.164 \pm 0.045$
$ ho_{ m K19}^{ m ISO}$	$0.324 \pm 0.110$
$r_{\rm p},_{\rm F18}$	$0.514 \pm 0.517$
$r_{ m p}, _{ m B22}^{ m LMC}$	$0.300 \pm 0.252$
$r_{\mathrm{p}},^{\mathrm{L}}_{\mathrm{B22}}$	$0.246 \pm 0.250$
$r_{\mathrm{p}},_{\mathrm{B22}}^{\mathrm{H}}$	$0.281 \pm 0.229$
$r_{\rm p}$ , LMC $_{\rm P22}$	$0.318 \pm 0.245$
<i>r</i> <sub>p</sub> , <sub>P22</sub>	$0.313 \pm 0.260$

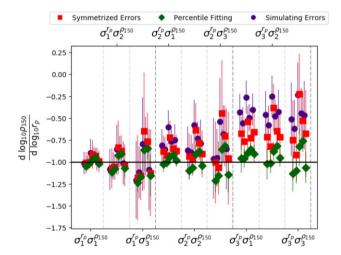


Figure 2. Recovered logarithmic slope for different mock data sets. The horizontal black line shows the input logarithmic slope (-1). The different symbols and colours indicate the methods tested (as indicated in the legend). The groups of points refer to the noise levels explored, as indicated by the subscripts 1, 2, and 3 on  $\sigma$ .

2, and 3, respectively) with which to produce the mock data sets. For each combination of these three noise levels, we explore five random realizations, getting a total of  $3 \times 3 \times 5 = 45$  pairs of mock data sets  $\xi = \{\rho_{150}, r_p\}$ .

Fig. 2 shows the recovered logarithmic slope *m* using the three different fitting procedures applied to the mocks. This analysis suggests that both the method of K19, i.e. the 'symmetrized errors' method, and the 'simulating errors' method suffer from a bias towards flatter relations. Thus, the slopes derived with these two methods will be considered as upper limits on the actual slope. The bias worsens for larger noise levels and is mainly driven by the error level on the pericentric radii. The K19 method yields slopes closer to the true values than the 'simulating errors' method; therefore, in the remainder of the article, we will not consider the 'simulating errors' method further. On the other hand, the percentile method appears to be the one with the best performance overall, yielding unbiased values of the slope. This will be our reference method.

#### 4 RESULTS

Fig. 3 shows the logarithmic slope recovered with the 'percentile fitting' and 'symmetrized errors' methods for all the data set pairs considered. In general, the logarithmic slopes cover the range between -0.5 and -1.5, hinting at an underlying anticorrelation, in agreement with previous studies. We also note that we recover the K19 results, when using the same set of densities and pericentric radii as in their work (we recover them exactly when applying their same methodology, and well within  $1\sigma$  when using the percentile fitting method; see the last panel of Fig. 3). The slope obtained with the H20 densities is the most uncertain, due to the larger errors on  $\rho_{150}$ , most likely due to the more complex modelling performed by the authors, which allows deviations from spherical symmetry in both the stellar and DM components of the dwarf galaxy.

Since the percentile fitting method appears to be the most robust, in Fig. 1 we show the best fits obtained using this method as a reference. In each panel, we show the 68 per cent confidence region of the PL fit as a shaded band. The colours of the shaded bands indicate whether the logarithmic slope m is compatible with 0 within  $<1\sigma$  (grey), between  $1\sigma$  and  $3\sigma$  (blue), or more than  $>3\sigma$  (magenta). As previously discussed, the value of m is negative in most cases, but with a varying statistical significance depending on the combination of data sets considered. In general, the relation obtained using the B22  $r_p$  in the potential including the LMC is flatter than that in the other cases with a logarithmic slope consistent with zero within  $1\sigma$ .

When using the same pair of data sets as K19, we recover their result of a statistically significant (at  $\gtrsim 3\sigma$ ) anticorrelation (see panels in the second and third rows and rightmost column of Fig. 1). The only other combination of data sets for which we find that m differs from zero in a statistically significant way is when the K19  $\rho_{150}$  are paired with the P22  $r_{\rm p}$ . Apart from variations in  $r_{\rm p}$ , a contributing factor is that the K19 densities for the galaxies with the smallest and largest pericentres tend to be, respectively, higher and lower than the corresponding estimates obtained in other works and also have smaller error bars.

In order to test the effect of the small sample size in the derivation of the significance of the logarithmic slope, we repeated the analysis removing one of the galaxies at a time. On average, the effect of removing one galaxy (from any of the data sets) is to diminish the significance of the recovered logarithmic slope by a factor of 0.66. Furthermore, we do not find any galaxy whose removal systematically increases the significance of the relation or that would make it become statistically significant if the original significance is below  $3\sigma$ .

As a further investigation of what relation might be underlying the inferred  $r_{\rm p}$  and  $\rho_{150}$ , we wish to compare models  $(f_{\theta})$  with a different dependence of  $\rho_{150}$  on  $r_{\rm p}$ . Specifically, we test two models in which  $\rho_{150}$  depends on  $r_{\rm p}$  and one in which it is independent:

- (i) PL: As in K19, the model is defined as  $f_{\theta} \equiv f(r \mid m, q) = 10^q r^m$ , and the scatter  $\sigma_i^2 = \delta \rho_i^2 + \left(m \cdot 10^q r_i^{m-1} \delta r_i\right)^2$ .
- (ii) Exponential (EX): The model is defined as  $f_{\theta} \equiv f(r \mid m, q) = \exp[q + mr]$ , and the scatter  $\sigma_i^2 = \delta \rho_i^2 + (m \exp[q + mr_i] \delta r_i)^2$ .
- (iii) Intrinsic scatter (IS): The central density is independent of the pericentric distances. The model is defined as  $f_{\theta} \equiv f(r \mid m, \sigma_0) = m$ , and the scatter  $\sigma_i^2 = \delta \rho_i^2 + \sigma_0^2$ , with  $\sigma_0$  the IS.

In the formulae described earlier, we use the same normalization as in equation (1), with the DM density  $\rho$  (and its uncertainty  $\delta\rho$ ) normalized to  $10^7~\rm M_{\odot}\,kpc^{-3}$ , while the pericentric distance r (and its uncertainty  $\delta\rho$ ) is normalized to 1 kpc.

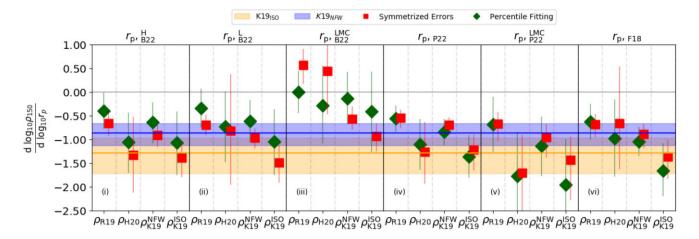


Figure 3. Measured logarithmic slope of the  $\rho_{150}$ - $r_p$  relation for the different pairs of data sets. The panels refer to the different sources of the measured pericentres: B22 (i, ii, iii); P22 (iv, v); and F18 (vi). The source of the central density measurement is indicated in the horizontal axis of each panel. Blue and orange lines indicate the reported values for the logarithmic slope in K19 using F18 pericentres, for an isothermal sphere (yellow) and a NFW profile (blue). Different symbols indicate different fitting methods as indicated in the legend.

We compare the quality of each of the previous models by making use of the Akaike information criterion (AIC; Akaike 1974) including a correction due to small sample size (Burnham 1998):

$$AIC_c = 2k - 2\log \mathcal{L}^{\max} + \frac{2k(k+1)}{n-k-1},$$
(2)

with k the number of free parameters of the model tested. The model comparison is done using the metric  $\Delta_i = \text{AIC}_c^i - \min_j \text{AIC}_c^j$ . Those models with  $\Delta_i < 2$  have substantial empirical support and cannot be rejected (Burnham 1998).

Since calculating a likelihood for the percentile fitting method is a rather complex business, we take advantage of the fact that the percentile fitting and the symmetrized error methods give comparable results to perform the model comparison via the latter method, which allows for a simpler Gaussian likelihood; then, we define the likelihood as

$$\mathcal{L}(\lbrace r_i \rbrace, \lbrace \rho_i \rbrace | \theta) = \prod_i \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left[-\frac{(\rho_i - f_\theta(r_i))^2}{2\sigma_i^2}\right],\tag{3}$$

with  $\sigma_i^2$  the variance,  $f_\theta$  the model, and  $\theta$  a vector gathering the model parameters. For sampling the likelihood, we have performed a Markov Chain Monte Carlo (MCMC) analysis as implemented in the public PYTHON package EMCEE (Foreman-Mackey et al. 2013).

A summary of the model comparison can be found in Table 3 (see Table C1 for the parameters of the best-fitting models to the different pairs of data sets). We find that, for an isolated MW, the two models with a decreasing central DM density of the dSphs as a function of  $r_{\rm p}$  (PL and EX) are preferred and perform similarly well. This preference is approximately independent of the source of the central density estimates. The only relevant exception is the densities from H20 for which models with or without a dependence on  $r_{\rm p}$  have similar empirical support. Also, with this analysis the K19 densities are those that provide the largest support to the models with decreasing  $\rho_{150}$  as a function of  $r_{\rm p}$ .

The effect of the inclusion of the LMC infall on the orbital integration of the dSphs appears to yield discrepant results between the two works that consider it. The pericentric distances derived by B22 seem to prefer the model where  $\rho_{150}$  and  $r_{\rm p}$  are uncorrelated, while three out of the four data pairs using P22 pericentres prefer the models in which  $\rho_{150}$  depends on  $r_{\rm p}$ . This is due to differences between the two sets of determinations that, while within  $1\sigma$  or at

**Table 3.** Results of model comparison for the different data set pairs. The model comparison is done using the metric  $\Delta_i = \mathrm{AIC}_c^i - \min_j \mathrm{AIC}_c^j$ . Those models with  $\Delta_i < 2$  have substantial empirical support and cannot be rejected (Burnham 1998). We explicitly identify them in the column 'Preferred models' and their  $\Delta_i$  is highlighted in boldface in the following columns.

$r_{\rm p}$	ρ <sub>150</sub>	Preferred models	$\Delta_{\mathrm{EX}}$	$\Delta_{ m PL}$	$\Delta_{IS}$
B22_H	R19	EX\PL	0.0	0.5	2.1
B22_H	H20	EX\PL\IS	0.5	0.0	1.1
B22_H	K19_NFW	$EX\PL$	0.9	0.0	8.4
B22_H	K19_ISO	$EX\PL$	0.0	0.3	6.8
B22_L	R19	$EX\PL$	0.0	1.3	3.4
B22_L	H20	$EX\PL\IS$	1.4	1.4	0.0
B22_L	K19_NFW	$EX\PL$	0.6	0.0	9.1
B22_L	K19_ISO	$EX\PL$	0.0	0.8	7.3
B22_LMC	R19	IS	4.1	5.3	0.0
B22_LMC	H20	$EX\PL\IS$	0.8	1.0	0.0
B22_LMC	K19_NFW	IS	8.7	11.9	0.0
B22_LMC	K19_ISO	IS	3.0	5.6	0.0
P22	R19	$EX\PL$	0.6	0.0	6.1
P22	H20	$EX\PL$	0.9	0.0	2.6
P22	K19_NFW	$EX\PL$	1.2	0.0	10.5
P22	K19_ISO	EX\PL	0.0	0.1	9.0
P22_LMC	R19	$EX\PL\IS$	0.0	0.7	1.9
P22_LMC	H20	$EX\PL$	0.1	0.0	2.6
P22_LMC	K19_NFW	$EX\PL$	0.0	0.6	3.3
P22_LMC	K19_ISO	$EX\PL$	0.0	0.8	4.2
F18	R19	$EX\PL$	0.4	0.0	3.5
F18	H20	$EX\PL\IS$	1.2	1.0	0.0
F18	K19_NFW	$EX\PL$	0.5	0.0	7.3
F18	K19_ISO	$EX\PL$	0.0	0.6	5.9

most  $2\sigma$ , do nevertheless move the points on the  $\rho_{150}$  versus  $r_p$  plane enough to change the significance of the relation. None the less, the empirical support for the models with a dependence on pericentric distance becomes much milder when including the effect of the LMC with respect to when considering an isolated potential (see Table 3), and also when using the P22 determinations of  $r_p$ . The inclusion of the LMC infall on to the MW seems to be an important ingredient for exploring the existence of this possible anticorrelation, and further

efforts to understand the effect of the LMC on the orbital properties of the MW satellites are needed.

#### 5 DISCUSSION AND SUMMARY

In this contribution, we reassess the question of whether an anticorrelation exists between the central DM density ( $\rho_{150}$ ) and orbital pericentre ( $r_p$ ) of classical dSph satellites of the MW. We explore in a quantitative way how the existence and strength of the relation depend on the adopted sets of inferred central DM densities and pericentric distances. Specifically, we consider the  $\rho_{150}$  determinations by R19, K19, and H20 and the pericentric radii by B22 and P22 in different gravitational potentials for the MW, both isolated or including the infall of a massive LMC. We also test a new method designed to handle the asymmetric error bars that naturally appear during the measurement of non-negative quantities such as radial distances or densities; this method is of general applicability to all situations in which asymmetric errors are present.

In general, we find that only the adoption of the K19 DM densities with  $r_p$  from the earlier Gaia DR2 determinations or the isolated MW potential by P22 results in strong empirical support for models in which  $\rho_{150}$  and  $r_p$  are anticorrelated and a logarithmic slope differing from zero at  $>3\sigma$  when exploring a PL dependence. In contrast, the H20 DM densities do not result in a preference for models in which  $\rho_{150}$  and  $r_p$  are anticorrelated with respect to models where they are independent. Clearly, uncertainties in the determinations of  $\rho_{150}$  play a role in establishing whether the existence of such anticorrelation is robust. The same can be said for the determinations of the pericentric radii. Even though the 3D bulk motions of the MW classical dSphs are now known in exquisite detail, transforming these observables into pericentric radii requires the adoption of a gravitational potential in which to integrate the orbits, which results in additional uncertainties. Besides the fact that the mass of the MW is still unknown within a factor of 2 (see e.g. Fritz et al. 2020; Wang, Hammer & Yang 2022), it is clear that the inclusion of the infall of a massive LMC in the determinations of the orbits of MW dSphs impacts the resulting  $r_p$ , and that there is some variation in the values determined for these  $r_p$ across different studies.

The issue is exacerbated by the small number of galaxies involved in the analysis, which implies that any conclusion should be taken with care. Works by K19 and Hayashi et al. (2022) suggest that increasing the sample through the inclusion of MW UFD galaxies completely washes away the anticorrelation present in the MW dSphs. It is not clear whether the reason is to be ascribed to the larger errors in the inferences of the DM densities and pericentric radii of UFDs or to some more fundamental property of these galaxies. For example, a fraction of the UFDs may be on their first infall on to the MW as can be gathered by results of orbital integration (e.g. B22; P22) and arguments on conservation of energy and angular momentum (e.g. Hammer et al. 2021). Thus, those galaxies may not have suffered from the tidal field of the MW, and consequently can present small pericentre distances and central densities unaffected by their orbits, thus obscuring the possible relation. We plan to explore the dependence of this anticorrelation on the inclusion of UFDs in future work.

The aforementioned possible effect of first infall is not necessarily limited to UFDs. The orbital integration of LeoI indicates that this classical dSph is on its first infall, having just passed its first pericentre. Furthermore, close interaction with the LMC may also introduce outliers in the relation. In particular, Carina has a small but non-negligible probability of being related to the LMC (B22; P22). However, as discussed in Section 4, the significance of the obtained

relations is not driven by any specific subset of satellites, indicating that our results are robust.

Not only is the existence of the anticorrelation controversial, but also its origin. A caveat against the survival bias hypothesis is the lack of MW satellites with high inner DM densities on external orbits, i.e. with large pericentres. This problem is reduced when adopting the pericentres of B22 (with the LMC): in this case, Draco and Leo I populate the high- $\rho_{150}$  and high- $r_p$  region of the parameter space.

The results from simulations are also contradictory. On the one hand, satellites in DM-only simulations seem to show an anticorrelation between the central DM density and the distance of the most recent pericentric passage (see H20). On the other hand, the inclusion of baryons in the simulations may have a strong impact on the proposed relation. As shown in Robles et al. (2019), the inclusion of the potential of the MW disc in the simulations can strongly alter the density of the satellites, diminishing the DM densities of the satellites with smaller pericentres and inverting the relation.

In summary, out of the 24 combinations of  $r_p$  and  $\rho_{150}$  explored, we found that only three strongly support (at more than  $3\sigma$  level) the presence of an anticorrelation between those two quantities: these represent the 12.5 per cent of the models explored. When making use of the AIC for model comparison, we find the data to be better described by models in which the central density  $\rho_{150}$  decreases as a function of  $r_p$  (PL and EX), and these perform similarly well. Only in one of the MW potentials explored are the data better described by a model with no dependence of  $\rho_{150}$  on  $r_p$ . Our results suggest that the strength and the existence of the  $\rho_{150}$ - $r_p$  anticorrelation on the MW's dSphs are still debatable. Exploring the existence and characteristics of this relationship with cosmological simulations at very high resolution, including properly modelled baryonic effects, as well as different DM flavours, will be a step forward towards understanding its emergence and its likelihood in a Lambda cold dark matter universe. We plan to do this in a future contribution.

# **ACKNOWLEDGEMENTS**

SCB acknowledges support from the Spanish Ministry of Economy and Competitiveness (MINECO) under the grant SEV-2015-0548-18-3 and the Spanish Ministry of Science and Innovation (MICIU/FEDER) through research grant PGC2018-094975-C22. GB and SCB acknowledge support from the Agencia Estatal de Investigación del Ministerio de Ciencia en Innovación (AEI-MICIN) and the European Regional Development Fund (ERDF) under grant number PID2020-118778GB-I00/10.13039/501100011033. GB acknowledges the AEI under grant number CEX2019-000920-S. ADC was supported by a Junior Leader fellowship from 'La Caixa' Foundation (ID 100010434), fellowship code LCF/BQ/PR20/11770010.

Data analysis was performed using the PYTHON<sup>5</sup> programming language. The following PYTHON modules were used for the analysis: PANDAS (The Pandas Development Team 2020); NUMPY (Harris et al. 2020); SCIPY (Virtanen et al. 2020); MATPLOTLIB (Hunter 2007); CORNER (Foreman-Mackey 2016); NUMBA (Lam, Pitrou & Seibert 2015); H5PY (Collette et al. 2021); and EMCEE (Foreman-Mackey et al. 2013).

## DATA AVAILABILITY

The data underlying this article will be shared on reasonable request to the corresponding author.

<sup>&</sup>lt;sup>5</sup>https://www.python.org

#### REFERENCES

3064

Akaike H., 1974, IEEE Trans. Autom. Control, 19, 716

Arnold B. C., Balakrishnan N., Nagaraja H. N., 2008, A First Course in Order Statistics, SIAM, Philadelphia, USA.

Balberg S., Shapiro S. L., Inagaki S., 2002, ApJ, 568, 475

Battaglia G., Nipoti C., 2022, Nat. Astron., 6, 659

Battaglia G., Taibi S., Thomas G. F., Fritz T. K., 2022, A&A, 657, A54

Burnham K. P., 1998, Model Selection and Multimodel Inference, A Practical Information – Theoretic Approach. Springer, Berlin

Collette A. et al., 2021, h5py/h5py: 3.5.0, available at https://doi.org/10.528 1/zenodo.5585380

Errani R., Peñarrubia J., Walker M. G., 2018, MNRAS, 481, 5073

Foreman-Mackey D., 2016, J. Open Source Softw., 1, 24

Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, PASP, 125, 306

Fritz T. K., Battaglia G., Pawlowski M. S., Kallivayalil N., van der Marel R., Sohn S. T., Brook C., Besla G., 2018, A&A, 619, A103

Fritz T. K., Di Cintio A., Battaglia G., Brook C., Taibi S., 2020, MNRAS, 494, 5178

Gaia Collaboration, 2018, A&A, 616, A1

Gaia Collaboration, 2021, A&A, 649, A1

Genina A., Read J. I., Fattahi A., Frenk C. S., 2022, MNRAS, 510, 2186

Hammer F., Wang J., Pawlowski M. S., Yang Y., Bonifacio P., Li H., Babusiaux C., Arenou F., 2021, ApJ, 922, 93

Harris C. R. et al., 2020, Nature, 585, 357

Hayashi K., Chiba M., Ishiyama T., 2020, ApJ, 904, 45

Hayashi K., Hirai Y., Chiba M., Ishiyama T., 2022, preprint (arXiv:2206.02821)

Hernquist L., 1990, ApJ, 356, 359

Hunter J. D., 2007, Comput. Sci. Eng., 9, 90

Kaplinghat M., Valli M., Yu H.-B., 2019, MNRAS, 490, 231

Lam S. K., Pitrou A., Seibert S., 2015, Proc. Second Workshop LLVM Compiler Infrastruct. HPC, LLVM'15, Numba: A LLVM-Based Python JIT Compiler, Assoc. Comput. Mach., New York, NY, p. 1

Merrifield M. R., Kent S. M., 1990, AJ, 99, 1548

Navarro J. F., Frenk C. S., White S. D. M., 1996, ApJ, 462, 563

Nishikawa H., Boddy K. K., Kaplinghat M., 2020, Phys. Rev. D, 101, 063009

Pace A. B., Erkal D., Li T. S., 2022, ApJ, 940, 136

Patel E. et al., 2020, ApJ, 893, 121

Prusti T. et al., 2016, A&A, 595, A1

Read J. I., Steger P., 2017, MNRAS, 471, 4541

Read J. I., Walker M. G., Steger P., 2019, MNRAS, 484, 1401

Richardson T., Fairbairn M., 2014, MNRAS, 441, 1584

Robles V. H., Bullock J. S., 2021, MNRAS, 503, 5232

Robles V. H., Kelley T., Bullock J. S., Kaplinghat M., 2019, MNRAS, 490, 2117

Simon J. D., 2019, ARA&A, 57, 375

The Pandas Development Team, 2020, Pandas-dev/Pandas: Pandas, available at https://doi.org/10.5281/zenodo.3509134

Virtanen P. et al., 2020, Nat. Methods, 17, 261

Wang J., Hammer F., Yang Y., 2022, MNRAS, 510, 2242

Wolf J., Martinez G. D., Bullock J. S., Kaplinghat M., Geha M., Muñoz R. R., Simon J. D., Avedo F. F., 2010, MNRAS, 406, 1220

# APPENDIX A: SIMULATED ERRORS

Our first approach for handling asymmetric errors is to describe the error distribution of each data point with some probability distribution. Once this distribution has been chosen, we sample new data from it and fit them.

The procedure can be summarized as follows:

- (i) Choose a probability distribution to be fitted to the percentiles  $x_{50\text{th}}$ ,  $x_{16\text{th}}$ , and  $x_{84\text{th}}$ .
- (ii) For each galaxy (g) and variable  $\xi^g = \{r_p^g, \rho_{150}^g\}$ , obtain the best-fitting parameters of the CDF corresponding to item (i).

- (iii) Randomly sample for each galaxy one value of the density  $\rho_{150}^{g,i}$  and pericentric distance  $r_p^{g,i}$  following the distribution obtained in (ii). Let us denote these new data as  $\xi^{g,i}$ .
- (iv) Fit the desired relation between resampled densities and pericentres, obtaining this way the set of parameters  $\theta^i$ .
- (v) Repeat (iii) and (iv) N times:  $\boldsymbol{\theta} = \{\theta^1, \theta^2, ..., \theta^N\}$ , with  $N = 10^4$ 
  - (vi) Store the results as the median of  $\theta$ .

For point (i), we tested three different distributions: a log-normal (LN) distribution, a log-logistic (LL) distribution, and a Gaussian distribution truncated at zero (TG). Since we find that these give similar results, in the main text we only show those for the LN distribution, and give the corresponding equation for the PDF,  $f_{\rm LN}$ , and CDF,  $F_{\rm LN}$ , below:

$$f_{\rm LN}(x \mid \mu_{\rm LN}, \sigma_{\rm LN}) = \frac{1}{x\sigma_{\rm LN}\sqrt{2\pi}} \exp\left[-\frac{\log x - \log \mu_{\rm LN}}{2\sigma_{\rm LN}^2}\right], \quad (A1)$$

$$F_{\rm LN}(x \mid \mu_{\rm LN}, \sigma_{\rm LN}) = \frac{1}{2} \left[ 1 + \operatorname{erf} \left( \frac{\log x - \log \mu_{\rm LN}}{\sqrt{2}\sigma_{\rm LN}} \right) \right]. \tag{A2}$$

# APPENDIX B: PERCENTILES AS RANDOM VARIABLES

Our second approach for handling asymmetric errors is to treat each percentile as a random variable and model its probability distribution via order statistics.

Let us denote the available sets of  $\rho_{150}$  and  $r_{\rm p}$  for each galaxy as  ${\bf d}=\{\rho_{150},r_{\rm p}\}$ . In general, these data will be different from the 'real' values, which we will denote as  ${\bf d}^*=\{\rho_{150}^*,r_{\rm p}^*\}$ . We indicate with  ${\boldsymbol \eta}$  all the extra parameters of the model. Our goal is to obtain the probability distribution of the parameters  $\{{\bf d}^*,\eta\}$  given our data  ${\bf d}$ , i.e.  $P({\bf d}^*,\eta|{\bf d})$ .

We can decompose the probability described earlier via the Bayes' theorem as

$$P(\mathbf{d}^*, \eta \mid \mathbf{d}) = \frac{1}{P(\mathbf{d})} P(\eta) P(\mathbf{d}, \mathbf{d}^* \mid \eta).$$
(B1)

The term in the denominator is a normalization constant that we can obviate. The second term on the r.h.s. is the a priori probability distribution of the parameters  $\eta$ . The last term is the joint probability distribution of the measured data  $\mathbf{d}$  and the 'real' values  $\mathbf{d}^*$ . We can further decompose this last term as

$$P(\mathbf{d}, \mathbf{d}^* \mid \eta) = P(\mathbf{d} \mid \mathbf{d}^*, \eta) P(\mathbf{d}^* \mid \eta)$$
(B2)

The probability of the data given the 'real' values  $P(\mathbf{d}|\mathbf{d}^*, \eta)$  and the probability distribution of the 'real' values given the model we are testing  $P(\mathbf{d}^*|\eta)$  remain to be solved for.

Let us start with the model. The second term on the r.h.s. of equation (B2) can be decomposed as:  $P(\mathbf{d}^*|\eta) = P(\rho_{150}^*|r_p^*, \eta)P(r_p^*|\eta)$ . The first term is the conditional probability of the central DM density of the dSphs given the pericentric distance. The second term is the a priori probability distribution of the pericentres, which we will assume uniform between  $r_{\min}$  and  $r_{\max}$ :  $P(r_p^*|\eta) = U_{[r_{\min}, r_{\max}]}(r_p^*)$ . We model the relation between the central densities and the pericentres as equation (1); the conditional probability of  $\rho_{150}$  given  $r_p$  can be written as

$$\begin{split} &P(\rho_{150}{}^* \mid r_p{}^*, \eta) = \mathcal{N}\left(\log_{10}(\rho_{150}{}^*) \mid q + m \log_{10}(r_p{}^*) \; ; \; \sigma^2\right) = \\ &\times \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2} \left[\log_{10}\left(\frac{\rho_{150}^*}{10^7 \; \mathrm{M}_{\odot}/\mathrm{kpc}^3}\right) - \left(q + m \log_{10} \; \frac{r_p^*}{\mathrm{kpc}}\right)\right)\right]^2\right\}. \end{split} \tag{B3}$$

The only term that remains to be solved for in equation (B2) is the conditional probability of the measured data **d** given the 'real'

Table B1. Fixed parameters of the model B1.

Parameter	Value
$N_{\min}, N_{\max}$	10, 10 <sup>4</sup>
$r_{\min}, r_{\max}$	$10^{-3}$ , $5 \times 10^2$
$m_{\min}, m_{\max}$	-15, 15
$q_{\min}, q_{\max}$	-15, 15
$\nu$ , $\tau^2$	1, 0.1

ones  $\mathbf{d}^*$ , i.e.  $P(\mathbf{d}|\mathbf{d}^*, \eta)$ . To characterize this last term, we need to give more details on the data gathering process. First, we make the reasonable assumption that the measurements of the central density and the pericentric distance are independent. Thus, we can split the probability distribution as

$$P(\mathbf{d} \mid \mathbf{d}^*, \eta) = P(\rho_{150} \mid {\rho_{150}}^*, \eta) P(r_p \mid r_p^*, \eta) . \tag{B4}$$

The available estimates of  $\rho_{150}$  and  $r_{\rm p}$  are in the form of percentiles, namely  $x_{16\text{th}}$ ,  $x_{50\text{th}}$ , and  $x_{84\text{th}}$ , where x is either  $\rho_{150}$  or  $r_{\rm p}$ . We can handle these data without losing information via order statistics.

Let us assume a continuous PDF for X,  $f(x; \beta)$  defined by the set of parameters  $\beta$ . If we get N samples from the distribution f and we sort them in such a way that:  $x_1 < x_2 < \ldots < x_{N-1} < x_N$ , the joint probability distribution of all the order statistics described previously (Arnold, Balakrishnan & Nagaraja 2008) is

$$f_{X_1,...,X_N}(x_1,...,x_N) = N! \prod_{i=1}^N f(x_i;\beta).$$
 (B5)

As we have only three of the N order statistics  $(x_i < x_k < x_j)$ , we shall marginalize over the unwanted N-3 parameters. The result of this procedure is the joint probability distribution of three order statistics:

$$P(x_{i}, x_{k}, x_{j} \mid \beta, N) = \frac{N!}{(i-1)!(N-j)!(k-i-1)!(j-k-1)!} \times f(x_{i} \mid \beta) f(x_{k} \mid \beta) f(x_{j} \mid \beta) F(x_{i} \mid \beta)^{i-1} (1 - F(x_{j} \mid \beta))^{N-j} \times (F(x_{k} \mid \beta) - F(x_{i} \mid \beta))^{k-i-1} (F(x_{j} \mid \beta) - F(x_{k} \mid \beta))^{j-k-1}.$$
(B6)

The indices  $\{i, k, j\}$  of the order statistics can be easily related to the  $p_{\{i, k, j\}}$ -th percentile as  $\{i, k, j\} = \inf[1 + p_{\{i, k, j\}}(N-1)/100]$ . Note that this probability depends on the family chosen for the PDF (CDF) f(F), the set of parameters  $\beta$ , and on the number of samples N. In particular, in this work  $\beta = \{x^*, s^2\}$ , where  $x^*$  is a location

<sup>6</sup>This is only true if the fraction is an integer. If not, the percentile is usually computed as the weighted mean between the closest order statistics. Taking into account this fact is possible but increases the complexity of the model. Moreover, the difference is only significant when the number of samples *N* is small. As this is not generally the case, we decided to stick with the simpler model

parameter while  $s^2$  is a parameter characterizing the width of the distribution. In particular, in this work we chose a log-normal PDF parametrized by the median  $x^*$  and the variance in log space  $s^2$ :

$$f(x \mid \beta = \{x^*, s^2\}) = \frac{1}{x\sqrt{2\pi s^2}} \exp\left[-\frac{\ln^2 x/x^*}{2s^2}\right].$$
 (B7)

If we introduce a prior over N and  $s^2$ , we can remove these two parameters via marginalization. We have assumed a uniform prior for N, between  $N_{\min} = 20$  and  $N_{\max} = 10^4$  such as  $P(N) = U_{[N_{\min}, N_{\max})}(N)$ . For  $s^2$ , we impose as prior a scaled inverse chi-square distribution:

$$P_{\text{Scale-Inv}-\chi^{2}}\left(s^{2} \mid \nu, \tau^{2}\right) = \frac{\left(\tau^{2}\nu/2\right)^{\nu/2}}{\Gamma(\nu/2)} \frac{\exp\left[-\frac{\tau^{2}\nu}{2s^{2}}\right]}{(s^{2})^{1+\nu/2}},$$
 (B8)

with 1 degree of freedom ( $\nu = 1$ ) and fixed scale parameter ( $\tau^2 = 0.1$ ), with  $\Gamma$  being the gamma function.

Finally, the probability distribution of the measured data given the 'real' values  $P(\mathbf{d}|d^*, \eta)$  can be written as

$$P(\mathbf{d}|d^*, \eta) = \prod_{\xi = \{\rho_{150}, r_p\}} \int ds^2 \sum_{N} \left[ U_{[N_{\min}, N_{\max})}(N) \right] \\ \times P_{\text{Scale-Inv-}\chi^2}(s^2 \mid \nu = 1, \tau^2 = 0.1) \\ \times P\left(\xi_i, \xi_k, \xi_j \mid \beta = \{\xi^*, s^2\}, N\right) \right].$$
 (B9)

The parameters we are finally left to fit are  $\{r_p^*\}$ ,  $\{\rho_{150}^*\}$ , and  $\eta=\{m,q,\sigma^2\}$ , i.e. 2n+3 parameters with n=8 the number of galaxies. We also need to choose the prior probability over  $\eta$ . For m and q, we choose uniform probability distributions  $P(m)=U_{(m_{\min},m_{\max})}(m)$  and  $P(q)=U_{(q_{\min},q_{\max})}(q)$ . For  $\sigma^2$ , we chose a prior of the form  $P(\sigma^2) \propto 1/\sigma^2$ . The values used in this work can be found in Table B1.

## APPENDIX C: MODEL COMPARISON

We report here Table C1, containing the results of the MCMC sampling of the likelihood described in Section 4.

**Table C1.** Results of the MCMC sampling of the likelihood described in Section 4. The first two columns indicate the data set pair used in the regression. The following columns indicate the median of the parameters of the three different models tested in this work. The lower and upper error bars indicate the 16th and 84th percentiles, respectively.

		I	S	PL		EX	
$r_{\rm p}$	$\rho_{150}$	m	$\sigma_0^2$	q	m	q	m
B22_H	R19	15.3+2.2	30+41	$2.3^{+0.4}_{-0.3}$	$-0.9^{+0.2}_{-0.2}$	$3.4^{+0.3}_{-0.2}$	$-0.012^{+0.004}_{-0.005}$
B22_H	H20	$17.3^{+6.0}_{-4.7}$	$111^{+189}_{-73}$	$3.4^{+1.3}_{-1.4}$	$-1.3^{+0.8}_{-0.8}$	$4.1^{+0.8}_{-1.0}$	$-0.027^{+0.017}_{-0.016}$
B22_H	K19_NFW	$15.3^{+2.7}_{-2.5}$	$45^{+57}_{-23}$	$2.8^{+0.4}_{-0.3}$	$-0.9^{+0.2}_{-0.2}$	$3.7^{+0.3}_{-0.2}$	$-0.018^{+0.005}_{-0.005}$
B22_H	K19_ISO	$13.1^{+3.6}_{-3.2}$	$71.5^{+94.4}_{-38}$	$3.5^{+0.7}_{-0.6}$	$-1.4^{+0.3}_{-0.4}$	$4.0_{-0.4}^{+0.4}$	$-0.027^{+0.007}_{-0.009}$
B22_L	R19	$15.3^{+2.2}_{-2.2}$	$29.9^{+41.3}_{-16.4}$	$2.5^{+0.4}_{-0.4}$	$-0.7^{+0.2}_{-0.2}$	$3.4^{+0.2}_{-0.2}$	$-0.009^{+0.003}_{-0.003}$
B22_L	H20	$17.2^{+5.9}_{-4.8}$	$110^{+183}_{-72.7}$	$2.6^{+2.1}_{-2.2}$	$-0.8^{+1.2}_{-1.1}$	$3.2^{+1.4}_{-1.2}$	$-0.010^{+0.015}_{-0.020}$
B22_L	K19_NFW	$15.3^{+2.6}_{-2.5}$	$45^{+59}_{-23}$	$3.0^{+0.4}_{-0.3}$	$-1.0^{+0.2}_{-0.2}$	$3.6^{+0.2}_{-0.3}$	$-0.012^{+0.003}_{-0.003}$
B22_L	K19_ISO	$13.1_{-3.2}^{+3.5}$	$71^{+96}_{-38}$	$3.9_{-0.6}^{+0.7}$	$-1.5^{+0.4}_{-0.4}$	$4.0^{+0.4}_{-0.4}$	$-0.020^{+0.005}_{-0.006}$
B22_LMC	R19	$15.3^{+2.3}_{-2.2}$	$30^{+42}_{-17}$	$0.1^{+0.7}_{-0.6}$	$0.6^{+0.3}_{-0.4}$	$2.2^{+0.8}_{-0.3}$	$+0.006^{+0.004}_{-0.011}$
B22_LMC	H20	$17.2^{+6.0}_{-4.7}$	$111^{+187}_{-73}$	$0.2^{+1.8}_{-1.6}$	$0.4^{+0.8}_{-0.9}$	$2.0^{+0.9}_{-0.8}$	$+0.005^{+0.008}_{-0.010}$
B22_LMC	K19_NFW	$15.3^{+2.7}_{-2.6}$	$45^{+57}_{-23}$	$2.2^{+0.4}_{-0.5}$	$-0.6^{+0.2}_{-0.2}$	$3.3^{+0.2}_{-0.3}$	$-0.008^{+0.003}_{-0.003}$
B22_LMC	K19_ISO	$13.1_{-3.3}^{+3.6}$	$72^{+96}_{-38}$	$2.8^{+0.6}_{-0.5}$	$-0.9^{+0.3}_{-0.3}$	$3.4^{+0.3}_{-0.3}$	$-0.013^{+0.003}_{-0.004}$
P22	R19	$15.3^{+2.2}_{-2.2}$	$30^{+41}_{-16}$	$2.2^{+0.4}_{-0.3}$	$-0.6^{+0.2}_{-0.2}$	$3.2^{+0.2}_{-0.2}$	$-0.008^{+0.002}_{-0.003}$
P22	H20	$17.2^{+5.9}_{-4.7}$	$111^{+191}_{-73}$	$3.5^{+1.2}_{-1.1}$	$-1.3^{+0.6}_{-0.7}$	$4.1^{+0.8}_{-0.8}$	$-0.022^{+0.012}_{-0.012}$
P22	K19_NFW	$15.3^{+2.6}_{-2.5}$	$44^{+56}_{-23}$	$2.5^{+0.3}_{-0.3}$	$-0.7^{+0.2}_{-0.2}$	$3.5^{+0.2}_{-0.2}$	$-0.011^{+0.003}_{-0.003}$
P22	K19_ISO	$13.1^{+3.6}_{-3.3}$	$71^{+94}_{-38}$	$3.3^{+0.7}_{-0.6}$	$-1.2^{+0.3}_{-0.4}$	$3.7^{+0.4}_{-0.3}$	$-0.017^{+0.005}_{-0.007}$
P22_LMC	R19	$15.3^{+2.7}_{-2.2}$	$30^{+41}_{-17}$	$2.4^{+0.6}_{-0.4}$	$-0.7^{+0.2}_{-0.4}$	$3.3^{+0.3}_{-0.2}$	$-0.010^{+0.004}_{-0.005}$
P22_LMC	H20	$17.4^{+5.8}_{-4.8}$	$110^{+186}_{-73}$	$4.2^{+1.6}_{-1.5}$	$-1.7^{+0.8}_{-0.9}$	$4.5^{+0.9}_{-0.9}$	$-0.029^{+0.013}_{-0.015}$
P22_LMC	K19_NFW	$15.3^{+2.6}_{-2.6}$	$45^{+56}_{-23}$	$2.9^{+0.7}_{-0.5}$	$-1.0^{+0.3}_{-0.4}$	$3.6^{+0.3}_{-0.2}$	$-0.014^{+0.004}_{-0.006}$
P22_LMC	K19_ISO	$13.1^{+3.6}_{-3.3}$	$73^{+94}_{-40}$	$3.6^{+1.4}_{-0.8}$	$-1.4^{+0.5}_{-0.8}$	$3.8^{+0.7}_{-0.4}$	$-0.020^{+0.007}_{-0.013}$
F18	R19	$15.3^{+2.2}_{-2.2}$	$29^{+40}_{-16}$	$2.4^{+0.5}_{-0.4}$	$-0.7^{+0.2}_{-0.3}$	$3.4^{+0.3}_{-0.2}$	$-0.010^{+0.003}_{-0.004}$
F18	H20	$17.4^{+5.9}_{-4.8}$	$111^{+190}_{-73}$	$2.4_{-2.3}^{+1.7}$	$-0.7^{+1.2}_{-0.9}$	$3.4_{-1.4}^{+1.2}$	$-0.012^{+0.016}_{-0.015}$
F18	K19_NFW	$15.2^{+2.7}_{-2.5}$	$45^{+58}_{-23}$	$2.8^{+0.5}_{-0.4}$	$-0.9^{+0.2}_{-0.3}$	$3.7^{+0.3}_{-0.2}$	$-0.014^{+0.003}_{-0.004}$
F18	K19_ISO	$13.1_{-3.3}^{+3.6}$	$72^{+94}_{-39}$	$3.6^{+0.9}_{-0.6}$	$-1.4^{+0.3}_{-0.5}$	$3.9^{+0.4}_{-0.4}$	$-0.020^{+0.005}_{-0.006}$

This paper has been typeset from a TEX/IATEX file prepared by the author.