The Astrophysical Journal, 970:1 (26pp), 2024 July 20 https://doi.org/10.3847/1538-4357/ad571c

HSTPROMO Internal Proper-motion Kinematics of Dwarf Spheroidal Galaxies. I. Velocity Anisotropy and Dark Matter Cusp Slope of Draco

Eduardo Vitral[1,2], Roeland P. van der Marel[1,3], Sangmo Tony Sohn[1], Mattia Libralato[4,5], Andrés del Pino[6], Laura L. Watkins[4], Andrea Bellini[1], Matthew G. Walker[7], Gurtina Besla[8], Marcel S. Pawlowski[9], and Gary A. Mamon[2]

[1] Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA; evitral@stsci.edu
[2] Institut d'Astrophysique de Paris, CNRS, Sorbonne Université, 98 bis Boulevard Arago, F-75014, Paris, France
[3] Center for Astrophysical Sciences, The William H. miller III Department of Physics & Astronomy, Johns Hopkins University, Baltimore, MD 21218, USA
[4] AURA for the European Space Agency (ESA), Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA
[5] INAF - Osservatorio Astronomico di Padova, Vicolo dell'Osservatorio 5, Padova I-35122, Italy
[6] Centro de Estudios de Física del Cosmos de Aragón (CEFCA), Unidad Asociada al CSIC, Plaza San Juan 1, E-44001, Teruel, Spain
[7] McWilliams Center for Cosmology, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213, USA
[8] Steward Observatory, University of Arizona, 933 North Cherry Avenue, Tucson, AZ 85721, USA
[9] Leibniz-Institute for Astrophysics Potsdam (AIP), An der Sternwarte 16, 14482 Potsdam, Germany
Received 2024 March 21; revised 2024 May 22; accepted 2024 June 7; published 2024 July 11

Abstract

We analyze four epochs of Hubble Space Telescope imaging over 18 yr for the Draco dwarf spheroidal galaxy. We measure precise proper motions for hundreds of stars and combine these with existing line-of-sight (LOS) velocities. This provides the first radially resolved 3D velocity dispersion profiles for any dwarf galaxy. These constrain the intrinsic velocity anisotropy and resolve the mass–anisotropy degeneracy. We solve the Jeans equations in oblate axisymmetric geometry to infer the mass profile. We find the velocity dispersion to be radially anisotropic along the symmetry axis and tangentially anisotropic in the equatorial plane, with a globally averaged value $\beta_{\rm B} = -0.20^{+0.28}_{-0.53}$, (where $1 - \beta_{\rm B} \equiv \langle v^2_{\rm tan}\rangle / \langle v^2_{\rm rad}\rangle$ in 3D). The logarithmic dark matter (DM) density slope over the observed radial range, $\Gamma_{\rm dark}$, is $-0.83^{+0.32}_{-0.37}$, consistent with the inner cusp predicted in ΛCDM cosmology. As expected given Draco's low mass and ancient star formation history, it does not appear to have been dissolved by baryonic processes. We rule out cores larger than 487, 717, and 942 pc at 1σ, 2σ, and 3σ confidence, respectively, thus imposing important constraints on the self-interacting DM cross section. Spherical models yield biased estimates for both the velocity anisotropy and the inferred slope. The circular velocity at our outermost data point (900 pc) is $24.19^{+6.31}_{-2.97}$ km s^{-1}. We infer a dynamical distance of $75.37^{+4.73}_{-4.00}$ kpc and show that Draco has a modest LOS rotation, with $\langle v/\sigma\rangle = 0.22 \pm 0.09$. Our results provide a new stringent test of the so-called "cusp–core" problem that can be readily extended to other dwarfs.

Unified Astronomy Thesaurus concepts: Dark matter (353); Dwarf spheroidal galaxies (420); Astronomy data analysis (1858); Proper motions (1295); Stellar kinematics (1608); Stellar dynamics (1596); Galaxy dynamics (591); Galaxy structure (622)

1. Introduction

Decades of astrophysical evidence support the notion that most of the matter in the Universe is dark. However, the nature of this dark matter (DM) remains a mystery. The most likely candidate is some form of cold DM (CDM), consisting of collisionless particles that cannot (yet) be detected directly but that interact through gravity.

Some of the best systems to study DM are the "classical" dwarf spheroidal galaxies (dSphs) in the Milky Way (MW). They are strongly DM dominated (Pryor & Kormendy 1990) and have a large number of bright stars that can be resolved, due to their proximity. The stars' motions contain information about the gravitational potential in which they move, and thus a large observational effort has been invested in obtaining their line-of-sight (LOS) velocities ($v_{\rm LOS}$, e.g., Tolstoy et al. 2004; Walker et al. 2007; Gilmore et al. 2022). Results from analyzing these data have been inconclusive about some CDM predictions. A conspicuous example of this is the so-called "cusp–core problem": the tension around the predicted and observed DM mass density profiles of galaxies. CDM halos in collisionless cosmological *N*-body simulations follow a nearly universal mass density profile that increases and diverges toward the center, forming a "cusp" (Navarro et al. 1997). In contrast, observations of some dSphs favor shallower density profile slopes, consistent with a constant-density "core" at the center (e.g., Battaglia et al. 2008; Walker & Peñarrubia 2011; Amorisco & Evans 2012; Brownsberger & Randall 2021).

Various solutions have been proposed to explain this and other discrepancies. Some propose fundamental changes in the nature of DM, such as warm DM (WDM), e.g., sterile neutrinos and gravitinos, that predict lower central DM densities and cored profiles (Dalcanton & Hogan 2001), or self-interacting DM (SIDM), for which DM particles in the central region thermalize via collisions and thereby form a cored profile (e.g., Sameie et al. 2020). Others include the impact of baryons, which may transform cusps into cores by transferring energy and mass to the outer parts of the halos, e.g., via supernova feedback (Read & Gilmore 2005; Pontzen & Governato 2012; Brooks & Zolotov 2014), or star formation events (Read et al. 2018).

Recent studies have also found that the orientation of a galaxy with respect to the viewer has a large impact on the derived velocity dispersion, resulting in a range of density slopes fitting the data (Genina et al. 2018).

Significant uncertainties are introduced by the fact that most observational studies are based solely on $v_{\rm LOS}$ measurements, which constrain only one component of motion. Consequently, interpretations rely on substantial assumptions, in particular that $v_{\rm LOS}$ is representative of the three-dimensional (3D) velocities.[10] These assumptions have been challenged by alternatives implying that the inferred, excessive dynamical mass-to-light ratios could be due to, e.g., modified gravity (McGaugh & Wolf 2010) or out-of-equilibrium dynamics caused by tidal interaction with the MW (Klessen & Kroupa 1998; Hammer et al. 2018), although the latter is hard to explain for satellites on orbits having higher pericenter values reported by Li et al. (2021), Battaglia et al. (2022), and Pace et al. (2022) from Gaia-based systemic proper motions (PMs).

Multiple techniques have been used to model $v_{\rm LOS}$ dispersion ($\sigma_{\rm LOS}$) profiles and, thus, constrain mass density profiles of dSphs. Examples include Jeans models (Walker et al. 2009; Zhu et al. 2016; Read & Steger 2017), distribution function (DF) fitting (Wilkinson et al. 2002; Vasiliev 2019), and Schwarzschild orbit superposition modeling (Breddels et al. 2013; Kowalczyk et al. 2019), each with their own strengths and weaknesses. However, all modeling techniques face the same problem: when only $v_{\rm LOS}$ are used, there is a strong degeneracy between the mass density profile $\rho(r)$ and the velocity anisotropy profile $\beta(r)$, which quantifies differences in velocity dispersions in orthogonal directions (Binney & Mamon 1982; Binney & Tremaine 1987). Some models mitigate this degeneracy by restricting parameter space or using higher-order moments (Vasiliev 2019; Genina et al. 2020; Read et al. 2021), but having only the LOS component of motion fundamentally limits what can be achieved.

The key to progress is to measure the internal PM kinematics of stars. The radial and tangential PM components directly measure the projected velocity dispersion anisotropy, which, under assumptions of inclination and intrinsic shape, uniquely determines $\beta(r)$ without requiring any dynamical modeling (e.g., van der Marel & Anderson 2010). This makes PMs crucial for dynamical modeling of dSphs, with models making use of PMs performing consistently better than those based solely on $v_{\rm LOS}$ (Read et al. 2021). Different techniques can be used to measure internal PM kinematics, but all of them require combining two or more epochs of observations to determine PMs of individual stars. At typical distances of MW dSphs, the only feasible instruments currently available for measuring individual PMs are Gaia, the Hubble Space Telescope (HST), and JWST.

Gaia has been tremendously successful in revolutionizing our view of the MW and its satellites, but the relatively shallow limiting magnitude ($G \sim 21$ mag) and its large PM uncertainties for typical dSphs stars in the MW halo (e.g., Vitral 2021; Pace et al. 2022) hinder its use for a direct measurement of internal PM dispersions (Martínez-García et al. 2021). An alternative is to combine Gaia astrometry with that from another instrument (e.g., HST) to achieve longer time baselines and thus lower PM uncertainties. This procedure has been applied in Massari et al. (2017, 2020) and del Pino et al. (2022). However, even with the Gaia end-of-mission PM uncertainties reduced by a factor of 3, the number of stars available to measure PM dispersion profiles will always be confined to those near the tip of the red giant branch, due to the limiting magnitude of Gaia. This is insufficient to discriminate between cusp and core models (Strigari et al. 2018; Guerra et al. 2023).

While the JWST time baseline is still too short (due to its recent launch) for a robust JWST versus JWST PM computation, comparing positions of stars in images obtained with the same detectors on board HST over time is the best means to obtain precise PMs of thousands of individual stars. HST is exquisitely well suited for astrometric and PM science, due to its stability, high spatial resolution, and well-determined point-spread functions (PSFs) and geometric distortions. By combining two or more epochs of space-based imaging, it is possible to measure precise internal PMs in nearby stellar systems (e.g., Libralato et al. 2022).

In the current work, we combine 18 yr of HST data, mostly obtained in the context of our High-Resolution Space Telescope PROper MOtion (HSTPROMO) Collaboration,[11] to measure PMs of hundreds of stars in the Draco dSph. With this, we measure its internal PM dispersion profile for the first time and thus provide unprecedented constraints on its DM density slope. We describe the data sets we used in Section 2, we explain the methods used to analyze the data in Section 3, and we present our results in Section 4. We comment on the robustness of our findings in Section 5, and we discuss and conclude our work in Sections 6 and 7, respectively. Throughout the paper, we use lowercase r to denote 3D distances and uppercase R to denote 2D projected distances.

2. Draco Data and General Characteristics

The Draco dSph is an excellent candidate to test the predictions of CDM scenarios. Its star formation shut down long ago ($\sim$10 Gyr; Aparicio et al. 2001), making it a prime candidate for hosting a "pristine" DM cusp, unaffected by baryonic processes (Read et al. 2018). Interestingly, the stellar mass of Draco is clearly below the limit where stellar feedback, as implemented in current cosmological simulations, should still produce a core (Fitts et al. 2017).[12] Furthermore, Draco is one of the most DM-dominated satellites of the MW (Kleyna et al. 2002) and is seemingly unaffected by Galactic tides that might heat its velocity dispersion profile (Odenkirchen et al. 2001; Ségall et al. 2007).

Recent efforts to infer the DM density of Draco from Jeans modeling of LOS velocities have yielded similar results: Read et al. (2018) fit rotationless spherical Jeans models combined with higher-order LOS moments and report a DM density slope at 150 pc of $-0.95^{+0.50}_{-0.46}$ (95% intervals); Hayashi et al. (2020) applied rotationless axisymmetric models to LOS data and report a cusp with "high probability" and a formal measurement of the asymptotic DM slope of $-1.03^{+0.14}_{-0.15}$ (68% intervals). Meanwhile, when formulating dynamical mass estimators based on PM dispersions, Lazar & Bullock (2020) found the then-available data to be insufficient for the purpose of constraining the asymptotic DM slope. Below, we describe the main characteristics of the new data sets we employ and how those are able to grasp the dynamical status of Draco in more detail and with better accuracy.

[10] By this statement, we mean that $v_{\rm LOS}$ is sometimes used to infer mass and/or anisotropy properties that are not uniquely constrained solely by the second-order moments of this single dimension.

[11] https://www.stsci.edu/~marel/hstpromo.html

[12] Fitts et al. (2017) place this limit at $M_\star = 2 \times 10^6\, M_\odot$.

Table 1
Overview of Draco Parameters

Reference	Data	α_0 (hh mm ss)	δ_0 (dd mm ss)	θ (deg)	ϵ	r_h (arcmin)	$\mu_{\rm LOS}$ (km s^{-1})	$\langle v_{\rm LOS}/\sigma_{\rm LOS}\rangle$
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
This work	Gaia EDR3	$17\ 20\ 16.42^{+0.02}_{-0.01}$	$+57\ 55\ 06.60^{+0.07}_{-0.32}$	89^{+10}_{-6}	$0.25^{+0.03}_{-0.03}$	$10.4^{+0.3}_{-0.4}$	$-291.73^{+0.48}_{-0.48}$	$0.22^{+0.09}_{-0.09}$
Martin et al. (2008)	SDSS	$17\ 20\ 14.4^{+0.6}_{-0.6}$	$+57\ 57\ 54^{+8}_{-8}$	89^{+2}_{-2}	$0.31^{+0.02}_{-0.02}$	$10.0^{+0.3}_{-0.2}$	...	...
Odenkirchen et al. (2001)	SDSS	$17\ 20\ 13.2^{+1.4}_{-1.4}$	$+57\ 54\ 60^{+0.07}_{-0.07}$	89^{+3}_{-3}	$0.28^{+0.01}_{-0.01}$	$12.4^{+1.6}_{-1.2}$	...	...
Wilson (1955)	Palomar	17 20 12.4	+57 54 55	...	...	...	...	...

Note. Column (1): reference where the values are reported. Column (2): data source of respective estimates (columns (3)–(7)). Column (3): R.A. of Draco center. Column (4): decl. of Draco center. Column (5): projected angle in the sky, from north to east. Column (6): projected ellipticity in the sky, defined as $1 - b/a$, with a and b the major and minor axes of the projected ellipse, respectively. Column (7): 3D half-number radius of a Plummer model fit for this work, of an exponential model fit for Martin et al. (2008), and of a Sérsic model fit for Odenkirchen et al. (2001) (using the 3D deprojection method from Vitral & Mamon 2021). Column (8): bulk LOS velocity. Column (9): mean rotation fraction in the LOS.

2.1. Projected Density

2.1.1. Center

The quoted center for Draco in the McConnachie (2012) catalog is the one from Wilson (1955), when the dSph was discovered. After that, more detailed sky surveys have allowed further refinement of this measurement. In particular, Odenkirchen et al. (2001) and Martin et al. (2008) used Sloan Digital Sky Survey (SDSS) data to estimate the values quoted in Table 1, and Vitral (2021) computed its center from Gaia EDR3 data assuming a Plummer (Plummer 1911) spherical model. In this paper, we compute it again from Gaia EDR3, but using a more refined version of the Vitral (2021) algorithm, which allows for an elliptical Plummer distribution (see Appendix A for analytical expressions of density profiles). The overall parameterizations are thus similar to the ones reported in Vitral (2021), with the exception of the elliptical Plummer, which adds two extra free parameters to the fit: (i) a projected angle θ in the sky, and (ii) the ellipticity implied by the minor-axis scale length of the projected ellipse.

Our fit was performed by a Markov Chain Monte Carlo (MCMC) routine that uses the software EMCEE (Foreman-mackey et al. 2013). We selected the most probable values from the joint MCMC posterior chain as the parameters and assigned uncertainties based on its difference to the 16th–84th percentiles of the respective posterior distribution. Our best-fit center, projected angle, and projected ellipticity are listed in Table 1. Overall, the fits agree very well with the estimates from Wilson (1955), Odenkirchen et al. (2001), and Martin et al. (2008).

2.1.2. Surface Density Profile

Later on, we will set up not only axisymmetric Jeans models but also spherical models to fit our data set (see Section 3). For that purpose, it is of interest to know the best scale radius of the observed data, assuming a spherical density profile, so that we can set reasonable priors to our models. Following Hayashi et al. (2020) and Massari et al. (2020), we assume a Plummer model. We derive the Plummer scale radius of Draco using Gaia EDR3 data, using the same formalism as in Vitral (2021), and with the (α_0, δ_0) centers calculated in Section 2.1.1.

Figure 1 displays the goodness of fit of our spherical Plummer profile to the Gaia EDR3 data. This satisfactory agreement yields a 3D half-number radius of $10\farcm4^{+0\farcm3}_{-0\farcm4}$, which lies between the values of $10\farcm0^{+0\farcm3}_{-0\farcm2}$ and $12\farcm4^{+1\farcm6}_{-1\farcm2}$, estimated by Martin et al. (2008) and Odenkirchen et al. (2001) with SDSS data, respectively, for an exponential model and a Sérsic model.

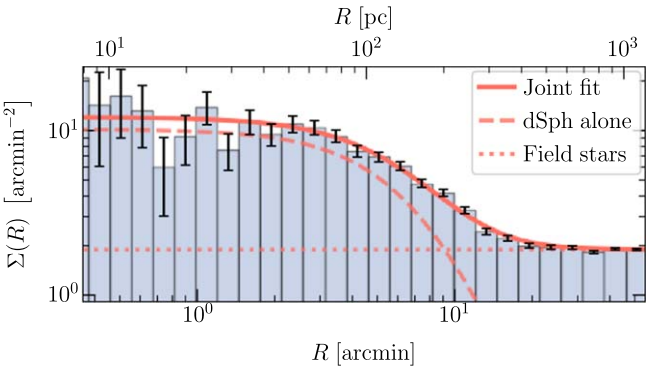

Figure 1. Surface density: goodness of fit of a Plummer (Plummer 1911) spherical model to Gaia EDR3 data of Draco, with significant field-star interlopers. The surface density fits follow the formalism detailed in Vitral (2021), which assumes a constant contribution of field stars.

2.1.3. Inclination

Due to the elliptical projected shape of Draco, we choose to model it as an oblate spheroid with a flattening parameter (i.e., intrinsic axial ratio) q. This relates to the projected axial ratio of Draco, q_p, through the equation (Binney & Tremaine 1987)

$$q_p^2 = \cos^2 i + q^2 \sin^2 i, \quad (1)$$

where i is the inclination of the spheroid (see Appendix B.2 below, where an edge-on model is defined to have $i = 90°$). We derive q_p from the ellipticity value in Table 1, which yields $q_p = 0.745 \pm 0.051$. The intrinsic axial ratio q is not known, but its probability distribution can be assumed to follow the general flattening probability distribution of oblate elliptical galaxies in the nearby Universe. This is given by Equation (4) from Lambas et al. (1992),

$$\psi_{\rm obl}(q) = \frac{2}{\pi}\sqrt{1-q^2}\,(0.344 + 21.272\,q^2 - 29.24\,q^4), \quad (2)$$

with $q \in [0, 0.9]$. This can be used to obtain a probability density function (pdf) for the inclination of Draco, as follows:

1. We draw N values[13] of flattening from the pdf in Lambas et al. (1992), which we label $q_{\rm all}$.
2. We draw N inclination values according to $i_{\rm all} = \arccos(U)$, where U is the uniform distribution within the [0, 1] interval. This ensures that the inclinations are sampled

[13] We use $N = 10^7$.

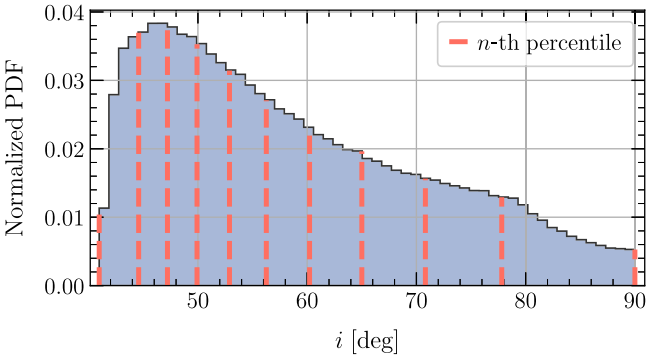

Figure 2. Inclinations: pdf of inclinations for Draco, computed according to the methodology described in Section 2.1.3. The dashed red lines indicate the nth percentiles of the distribution, with n spaced from 0 to 100, on intervals of 10.

uniformly on the surface of a unit sphere from face-on ($i = 0°$) to edge-on ($i = 90°$) cases.

3. For each of those inclinations, we compute the respective projected flattening from Equation (1), using the $q_{\rm all}$ values previously drawn, and we label these as $q_{p,\rm all}$.
4. From those ($q_{p,\rm all}$, $i_{\rm all}$) pairs, we keep the ones that satisfy $|q_{p,\rm all} - q_{p,\rm Draco}| < 0.01$.

The remaining pairs from the last step above yield the projection of the inclination pdf onto the observed projected axial ratio of Draco, which is depicted in Figure 2. The resulting [16th, 50th, 84th] percentiles of the inclination and flattening final distributions are [46.°1, 56.°3, 73.°4] and [0.38, 0.60, 0.72], respectively.

2.2. Line-of-sight Velocities

Draco has been the subject of many observational campaigns to obtain LOS velocities of samples of individual stars (e.g., Armandroff et al. 1995; Wilkinson et al. 2004; Walker et al. 2015). Recently, Walker et al. (2023) provided the most complete catalog of dwarf galaxy LOS kinematics, including also metallicities and stellar parameters. Here we make use of this catalog to complement our PM data set. In this section, we study some of the main aspects of this LOS data set, including its interloper contribution, the implied galaxy rotation, and the influence of binaries on the inferred kinematics.

2.2.1. Interloper Cleaning

To best interpret our results based on LOS data, we need to remove interlopers (essentially, stars in the foreground and background). Hence, we perform a multidimensional mixture model to assign membership probabilities to each star in our subset and then select it (or not) based on a threshold probability.

We first narrow our study to catalog stars that satisfy `good_obs == 1`, as suggested in Walker et al. (2023, Section 5). This essentially removes stars having high $v_{\rm LOS}$ uncertainties. Next, we select the parameterizations that we use to model the joint pdf of Draco stars (tracers) and interlopers, in each dimension of the data (i.e., $v_{\rm LOS}$, $T_{\rm eff}$, $\log g$,[14] [Fe/H], [Mg/Fe]). The tracer pdf's of $v_{\rm LOS}$, $\log g$, [Fe/H], and [Mg/Fe] were modeled as a Gaussian, while the tracer pdf of $T_{\rm eff}$ was modeled as a double Gaussian. The interloper pdf's of $\log g$, [Fe/H], and [Mg/Fe] were modeled as a triple Gaussian, while the interloper pdf's of $v_{\rm LOS}$ and $T_{\rm eff}$ were modeled as log-Gauss and double-Gaussian distributions, respectively. These choices of pdf's were made so as to maximize the goodness of fit.

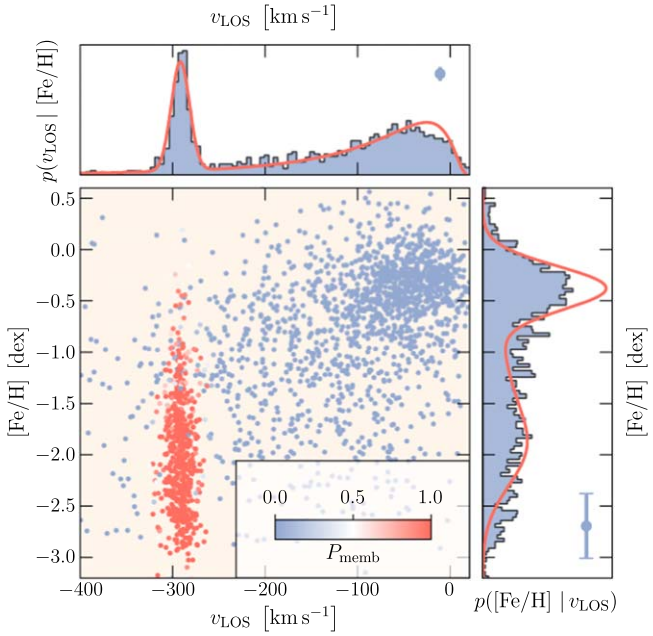

Figure 3. Interloper cleaning: goodness of fit of our multidimensional mixture model, projected on the $v_{\rm LOS}$ and [Fe/H] dimensions. The central panel shows the stars from the Walker et al. (2023) catalog on the sky region of Draco, in the $v_{\rm LOS}$−[Fe/H] plane, color coded by their Draco membership probability. The upper and right panels display the histogram of this plane projected on the $v_{\rm LOS}$ and [Fe/H] dimensions, respectively. In the corner of each side panel, we represent the median uncertainty on the respective dimension by an error bar. Stars with $P_{\rm memb} > 0.99$ were selected for inclusion in our analysis.

We fitted this multidimensional distribution through an MCMC routine on discrete data and considered the region between the 16th and 84th percentiles of each posterior distribution as the uncertainty on our fits and the most probable values from the joint MCMC posterior chain as the best parameters. Figure 3 showcases the goodness of our fit, projected on the $v_{\rm LOS}$−[Fe/H] dimensions. From our fits, we assigned as Draco members the stars having a membership probability higher than 99%, which removes most of the stars beyond a little more than $3\sigma_{\rm LOS}$ from the bulk $v_{\rm LOS}$. This final subset was composed of 435 stars with $v_{\rm LOS}$ data and uncertainties smaller than the value of $\sigma_{\rm LOS}$. Our measured value for the bulk LOS motion, $\mu_{\rm LOS}$,[15] is presented in Table 1. For this subset, we chose not to correct for perspective effects caused by Draco's bulk motion, since those have negligible effects. Indeed, the rms correction for the sample stars implied by Equation (13) of van der Marel et al. (2002) is only 0.2 km s^{-1}, while the rotation and velocity dispersion profiles inferred further in this work change by at most 0.1 km s^{-1}, which is well below their respective measurement uncertainties.

2.2.2. Rotation

Like most galaxies (e.g., Martínez-García et al. 2023), it is possible that Draco possesses detectable mean rotation. We

[14] Throughout this work, we denote the base-10 logarithm as log and the logarithm in the natural base as ln.

[15] We label the bulk LOS motion of Draco as $\mu_{\rm LOS}$ and the first-order moment over the major axis, which relates to rotation, as $\langle v_{\rm LOS} \rangle$.

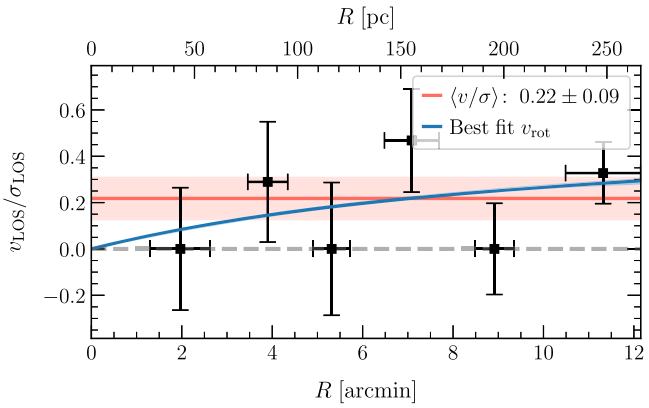

Figure 4. Rotation: rotation profile of Draco, as a function of its projected radius, computed with LOS data. The black circles and error bars show the v/σ ratio per bin, while the red and blue curves (and respective uncertainty regions) show, respectively, the overall mean through the Draco field and the best fit of Equation (3). We add a dashed gray line to represent the case of no rotation. From this plot, we conclude that Draco has a modest amount of rotation.

estimate here the rotation fraction of Draco, $\langle v/\sigma \rangle$, from its $v_{\rm LOS}$.

We partition the LOS data into six concentric annuli on the sky, all of which have the nearly the same number ($N \sim 50$) of stars. We then perform a sinusoidal fit with free amplitude for each partition (i.e., to the respective $v_{\rm LOS}$ vs. projected angle quantities) and free mean velocity and phase. The mean and phase[16] are forced to be the same for all annuli (thus, in total, eight free parameters). The measured amplitude and $\sigma_{\rm LOS}$ per annulus allow us to construct the rotation profile displayed in Figure 4, which has a mean $\langle v_{\rm LOS}/\sigma_{\rm LOS} \rangle = 0.22 \pm 0.09$ averaged over all radii (listed in Table 1). For visualization, the blue line in the figure displays the best fit using a parameterization of the form

$$\left(\frac{v}{\sigma}\right)_{\rm rot} = \left(\frac{v}{\sigma}\right)_0 \frac{(R/R_0)}{[1 + (R/R_0)]^{1+\zeta}}, \tag{3}$$

which increases linearly at small projected radii and falls as a power law at higher projected radii.

Our results lie between those found by Hargreaves et al. (1996) and Kleyna et al. (2002), who used less complete data sets than ours. They found a rotation amplitude of 0.7 km s^{-1} around Draco's minor axis and of 6 km s^{-1} at 30′. Meanwhile, our $(v/\sigma)_{\rm rot}$ fit predicts $v/\sigma = 0.4$ at 30′, which translates to a rotation amplitude, $v_{\rm rot}$, of less than 4 km s^{-1} at this projected radius. Our fit also agrees well with Martínez-García et al. (2021), where internal rotation was confirmed using Gaia EDR3 PMs. From those results, we conclude that although Draco has some mean rotation, it is small when compared to the overall velocity dispersion, especially at inner radii, where most of the data are concentrated (see Figure 1 for comparison).

[16] The phase is taken with respect to the position angle of the minor axis (see column (5) of Table 1). We found that while the rotation curve was robustly constrained by the data, the exact angle of the rotation axis was not. The fits showed angle variations between annuli, and the best-fit angle also depended on the exact choice of annuli, both in excess of the formal uncertainties. While a kinematic axis intermediate between the major and minor photometric axes appeared formally preferred by the fits, we concluded after experimentation that an oblate model with rotation around the photometric minor axis was acceptable. In any case, the data do show a preference for one spin sign (i.e., receding relative velocities on the western longitudes) rather than another.

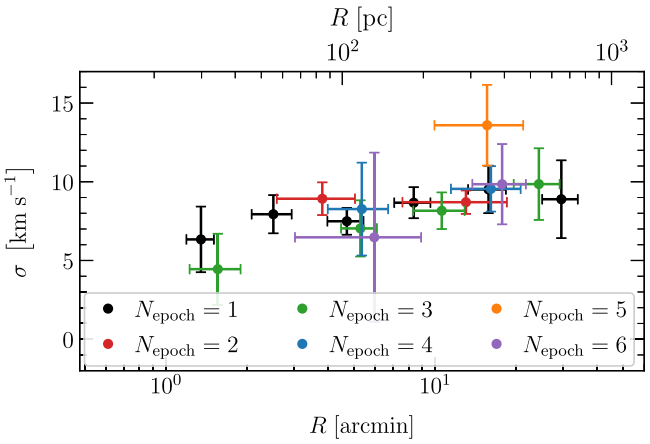

Figure 5. Impact of unresolved binaries: multiepoch velocity dispersion profiles (as a function of projected radii) for groups of LOS data constructed from a different number of epochs. The number of stars having multiepoch observations is scarcer, and thus we have fewer radial bins for those. The excellent agreement within 1σ between all $N_{\rm epoch}$ and the fact that the $\sigma_{\rm LOS}$ profiles associated with higher $N_{\rm epoch}$ are not considerably smaller than the $N_{\rm epoch} = 1$ subset give us good confidence that unresolved binaries do not affect our mass estimates beyond the statistical uncertainties.

2.2.3. Binaries

Per construction, measurements of $v_{\rm LOS}$ from spectral lines are subject to Doppler shifts from unresolved binary motion. As a consequence, single-epoch LOS measurements can carry an overestimated $\sigma_{\rm LOS}$ and thus an overestimation of the system's total mass due to binary motion. Meanwhile, given the multiepoch requirement of PM measurements, those end up averaging the motion of unresolved binaries to zero, such that the mentioned overestimation becomes negligible. For example, while Bianchini et al. (2016) showed that globular star clusters with unresolved binary fractions up to 50% should introduce changes <6% on the PM velocity dispersion, Pianta et al. (2022) recently performed simulations of dSphs to argue that one could reach much higher changes when using only LOS data.

Such an undesirable effect can be almost completely corrected by obtaining multiepoch LOS observations, as recently argued by Wang et al. (2023). Given Draco's high binary fraction of 50% (Spencer et al. 2018), we perform here a multiepoch test to gauge the influence of unresolved binaries on our cleaned Draco LOS data set. To do so, we plot the radial profile of $\sigma_{\rm LOS}$ for groups of LOS data constructed from a different number of epochs. If unresolved binaries are to affect our LOS data as proposed by Pianta et al. (2022), then one should expect multiepoch velocity dispersions to be considerably smaller than the ones computed from single-epoch exposures.

Figure 5 displays our multiepoch comparison, where velocity dispersion profiles are computed according to van der Marel & Anderson (2010, Appendix A) and Vitral et al. (2023a, Section 3.2.1). The number of stars having multiepoch observations is scarcer, and thus we have fewer radial bins for those. In any case, all our multiepoch radial dispersion profiles agree within 1σ with the single-epoch measurement. Hence, Figure 5 reassures us that for the cleaned LOS subset of Walker et al. (2023) the effects of unresolved binaries in the velocity dispersion of Draco are within the statistical uncertainties. Finally, we revisit this conclusion in Section 5.1, where we compare our results for subsets with and without LOS data.

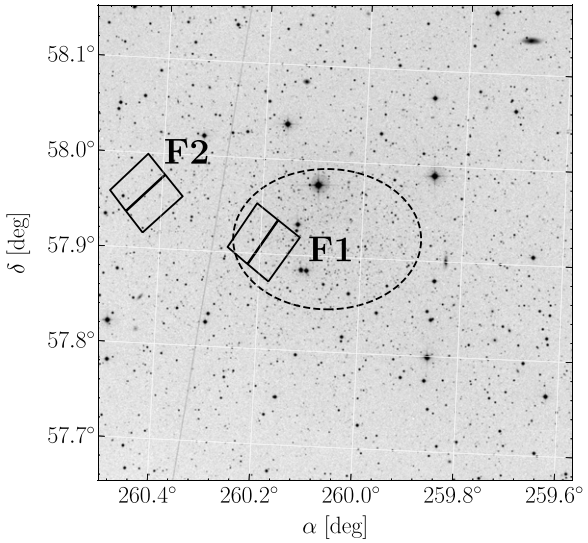

Figure 6. Observed fields: HST target fields (background image is from the STScI Digitized Sky Survey; see acknowledgments), with a black ellipse showing Draco's half-number radius.

2.3. Proper Motions

2.3.1. Observations and Astrometric Catalogs

For our new PM measurements of Draco stars, we used multiepoch HST ACS/WFC imaging data. Descriptions about field locations and observations during earlier epochs for our target fields F1 and F2 are provided in Sohn et al. (2017). The field locations are also shown in Figure 6. In summary, the F1 field had three epochs of imaging data obtained in 2004, 2006, and 2013, while F2 had two epochs of imaging data obtained in 2004 and 2012. All fields were observed once again on 2022 October–November through our HST program GO-16737 (Sohn et al. 2021) using the same filter (F606W), telescope pointing, and orientation as in the previous epochs.[17] In this latest epoch, we obtained 15 individual exposures, with each exposure lasting 430 s for each field.

The data analysis largely followed the procedures described in Bellini et al. (2018) and Libralato et al. (2018). Here we provide only a high-level outline of the PM derivation process and refer the reader to those papers for more details about the methodology. We downloaded the flat-fielded `_flt.fits` images of all target fields for each epoch from the Mikulski Archive for Space Telescopes and processed them using the `hst1pass` program (Anderson 2022) to derive a position and a flux for each star in each exposure. Instead of working on the `_flc.fits` images that are corrected for charge transfer efficiency (CTE) losses, we utilized the table-based CTE correction option in `hst1pass`, which is an improved version of the ones used in previous works (Anderson 2022; J. Anderson, in preparation). We applied corrections to the positions using the ACS/WFC geometric distortions based on Kozhurina-Platais et al. (2015); these were further extended to include time-dependent distortion variations beyond 2020 (V. Kozhurina-Platais, private communication).

For each field, we constructed a "master frame" using the average positions of stars from the repeated first-epoch exposures. The (X, Y) axes of these master frames were aligned with (α, δ) by registering the stellar positions to the Gaia DR3 astrometric system. We aligned the positions of stars from the other epochs to these master frames using a six-parameter linear transformation and determined average positions for each epoch. By construction, this procedure aligns the star fields between different epochs, leading to zero PM on average for the Draco dSph stars themselves. This does not affect our results since we are mostly interested in measuring the internal velocity dispersion on the plane of the sky (POS; see Section 2.3.3 below for more discussion of this topic). Uncertainties on the average positions were determined from the repeated measurements as the rms divided by the square root of the number of exposures. In the end, for each field per epoch, we prepared a catalog that includes positions of stars measured as described above and average instrumental[18] F606W and F814W magnitudes (from the data in the 2012–2013 epoch) output by `hst1pass`.

2.3.2. Photometric Cleaning

Once our observations are reduced and we have the master frame (X, Y) positions of sources at each epoch, for each field, we first perform a photometric cleaning of the data. The goal of this step is mainly (i) to remove interlopers, (ii) to remove background galaxies, and (iii) to remove stars associated with poor photometry that might bias our PM analysis. Points (i) and (ii)–(iii) are performed independently, and we further select stars that simultaneously survived both sets of cleaning cuts.

Point (i) is accomplished by performing a cleaning on the color–magnitude diagram (CMD) of each field, at each epoch. We use a friends-of-friends procedure, where we assign as an interloper a star whose distances to other stars in the CMD[19] are greater than typical distances in the subset. To do so, we define, after inspection of the CMD–distance distribution at each epoch, fiducial distance thresholds to use in this cleaning. Since this step is likely to remove bright stars on the tip of the red giant branch and the horizontal branch (this is because they do not have a high number of neighbors), we reintroduce them to the cleaned subset. They are likely dSph members and would be filtered in further steps if they are not. As an example, Figure 7 (top panels) displays the results of this CMD cleaning for the three epochs of Field 2.

Next, to address point (ii), we remove sources likely to be background galaxies, which lie on the upper side of the `QFIT`—F606W diagram,[20] departing from the bulk set of stars. This step is performed with a similar friends-of-friends analysis to that for point (i), with different distance thresholds per field and per magnitude range. Finally, we proceed to point (iii) by removing stars that satisfy `QFIT` $\geqslant 0.2$, as they are associated with poor PSF fits. The final `QFIT` cleaning of our subset is displayed in the bottom panels of Figure 7.

2.3.3. Local Corrections

After having a photometrically cleaned subset that is also devoid, at least to a large extent, of interlopers, we proceed to compute the PM of each star in our subset. Essentially, the raw

[17] We also observed another field, F3, but found the resulting PMs of insufficient accuracy for the present purpose, due to the use of different HST filters per epoch.

[18] The instrumental magnitude in a given filter is defined here as mag $= -2.5 \log c$, where c is the number of photon counts per exposure for a source.

[19] The distance is defined in F606W vs. (F606W–F814W) space using percentiles to normalize each dimension, similarly to what is explained in Section 3.3 of Vitral et al. (2022, Equations (1)–(2)).

[20] The `QFIT` parameter is a combined measure of goodness of fit and signal-to-noise ratio (see Anderson et al. 2006; Libralato et al. 2014 for details).

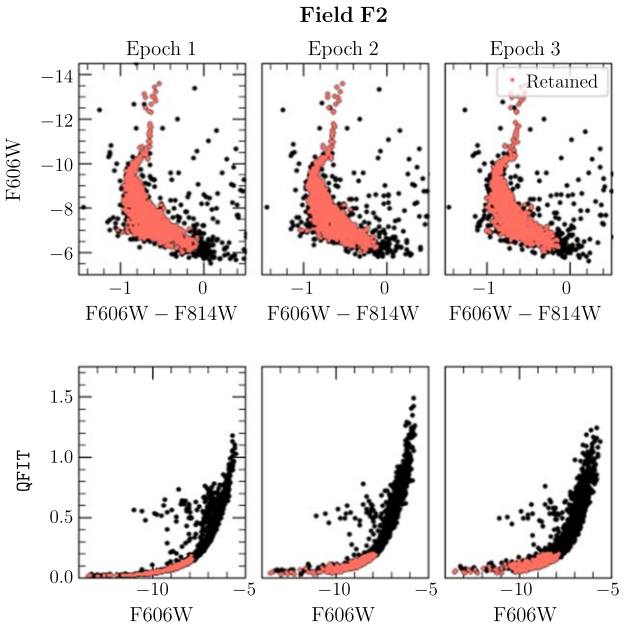

Figure 7. Photometric cleaning. The top panels show the CMDs of our initial interloper cleaning of Field 2 (using instrumental magnitudes), with the original stars in the subset in black and the retained ones in orange. The bottom panels show the QFIT-based cleaning of Field 2, with the original stars in the subset in black and the retained ones in orange. The plots attest to the effectiveness of our friends-of-friends photometric and interloper cleaning.

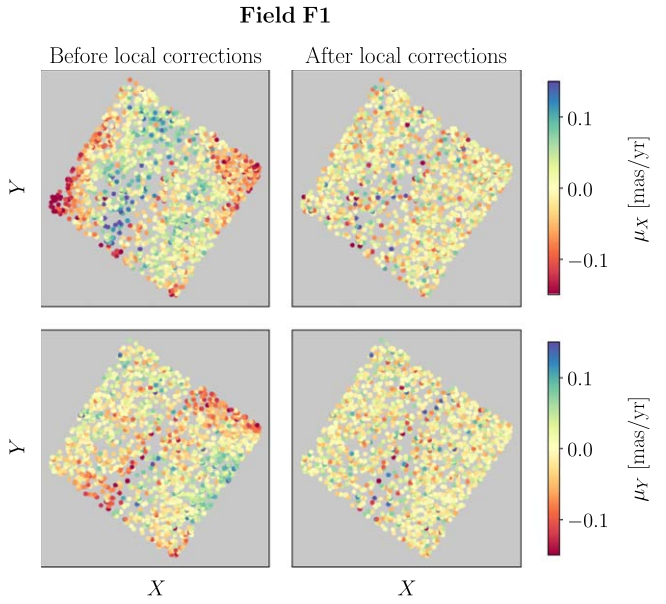

Figure 8. Local corrections: effect of local PM corrections on our data set (shown here for Field 1). The panels show the master frame (X, Y) positions of the sources, color coded according to the PM in each direction (for reference, Draco's typical PM dispersion peaks around 0.03 mas yr^{-1}). The left panels show clearly artificial PM shifts related to uncorrected residual CTE and geometric distortion effects, especially around the detector edges and boundaries. The right panels show our corrected sample, which has a generally homogeneous PM distribution throughout the whole field, centered around null PMs.

PMs are computed by a least-squares line fit of the master frame (X, Y) positions as a function of the epoch time. We use the NUMPY.POLYFIT routine from PYTHON, assuming the (X, Y) uncertainties calculated in Section 2.3.1 and no χ^2 rescaling.[21] We store the χ^2 of the fit for later data cleaning.

The raw PMs may contain low-level systematic effects related to the CTE issues of HST's degrading charge-coupled devices, as well as from subtle variations in geometrical distortion between epochs. As a result, some regions of the observed fields may present systematically higher/lower PMs. This problem has been previously reported, for instance, in Bellini et al. (2014) and Libralato et al. (2022). We display this for our Field 1 in the left panels of Figure 8. The best procedure to correct for these effects is to perform a local PM correction that shifts those regions back to the bulk PM of the field. In practice, we follow the procedures laid out in previous works (e.g., Bellini et al. 2014; Libralato et al. 2022) that have constructed HST PMs by looping over each star and removing the median PM of a local net of the 10 closest[22] stars. This process adds an extra layer of uncertainties (basically the uncertainty on the median,[23] which we add quadratically to the original PM uncertainty) but successfully renders the PM data set more homogeneous. The right panels of Figure 8 show that this successfully removes most of the systematic effects.

This local correction step removes not only streaming artifacts from the data but also any variations in mean streaming intrinsic to the galaxy. However, it preserves the local velocity dispersion, which is most critical to perform mass modeling. We verified with axisymmetric rotating mock data sets that this step does not significantly change the second-order velocity moment of the data (changes remain smaller than ∼5% for all possible inclinations). Moreover, our axisymmetric model fitting in Section 4.2.2 below explicitly accounts for the fact that any mean streaming in the PM directions is not observationally constrained.

2.3.4. Sky Coordinates

As explained in Section 2.3.1, our master frame (X, Y) positions are already aligned, per construction, with sky coordinates, by using Gaia reference frames. This means that our PMs computed in (X, Y) directions are straightforwardly converted as $\mu_{\alpha,*} = -\mu_X$ and $\mu_\delta = \mu_Y$, where we denote $\mu_{\alpha,*} \equiv \mu_\alpha \cos\delta$.

To convert (X, Y) positions to (α, δ), we first perform a naive translation/rotation such that the coordinates approximately match the true ones. Next, we select brighter sources and associate them with Gaia EDR3 catalog sources. We perform a final translation/rotation to minimize the logarithm of the sum of distances between those matches, and we apply the respective conversion parameters to our data set. We verified that our matches are performed correctly by comparison to the Draco subset from del Pino et al. (2022).

2.3.5. Outliers and Underestimated Errors

After the conversion to sky coordinates, our PM subset is nearly ready for use. However, there might still be hidden interlopers in the data with unusually high PMs, or stars with underestimated errors that might bias our results.

[21] Essentially, the χ^2 quantity is defined as $\chi^2 = \sum_i^N \frac{(y_i - f(x_i))^2}{(N - N_{\rm free})}$, with $N_{\rm free} = 2$ (i.e., a line) and N being the number of epochs used in the fit.

[22] Here "closest" refers to (X, Y) spatial positions. We verified that the systematics observed in Figure 8 pertained mostly to geometrical distortions (rather than CTE), where distances in magnitude space are not relevant and could instead bring farther-away stars into the local net.

[23] The uncertainty on the median for a Gaussian distribution is given by $\epsilon_{\rm median} = \sqrt{\frac{\pi}{2}\frac{n}{(n-1)}}\ \epsilon_{\rm mean} = \sqrt{\frac{\pi}{2}\frac{n}{(n-1)}}\ \frac{\sigma}{\sqrt{n}}$ (Kenney & Keeping 1963), where σ and n are, respectively, the distribution standard deviation and number of samples, which we fix to 10.

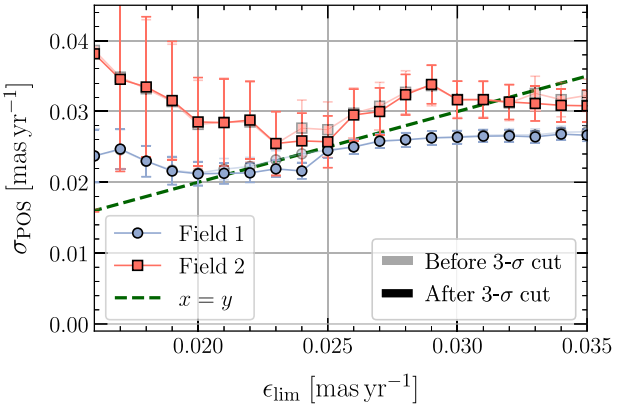

Figure 9. Underestimated errors: measured PM dispersion, $\sigma_{\rm POS}$, as a function of the maximum PM uncertainty in the data set, $\epsilon_{\rm lim}$. Blue curves relate to Field 1, while the red ones relate to Field 2. Opaque curves relate to the PM subsets without the 3σ cleaning explained in Section 2.3.5, while the solid curves show results for the PM subsets with such a cut. For reference, we display an $x = y$ line in dashed green. This plot shows that it is important to impose a cut in PM errors at $\epsilon_{\rm lim} \sim 0.024$ mas yr^{-1}. Inclusion of stars with PM uncertainties in excess of the intrinsic galaxy dispersion yields an overestimated value for the galaxy dispersion.

A rapid test to probe the number of such stars is to fit the PM distribution with a Gaussian (in both radial and tangential directions) and to compare the fraction of stars beyond 3σ to the fraction ∼0.27% predicted in this Gaussian. When performing this exercise, we observe that the fraction of stars in the wings of the distribution increases as we consider stars with higher PM uncertainties. This not only shows that we could be encompassing interlopers but also points to the possibility of underestimated errors toward fainter stars.

When measuring the velocity dispersion of a subset (as explained in Section 2.3.6), we are actually fitting the quadratic sum of the intrinsic dispersion and the errors associated with the tracers. If the errors are underestimated, the intrinsic dispersion will be overestimated. Figure 9 shows the PM velocity dispersion (i.e., $\sigma_{\rm POS}$) of stars with maximum PM uncertainty $\epsilon_{\rm lim}$. The curves show that the fitted $\sigma_{\rm POS}$ starts to increase as we include stars with errors beyond $\epsilon_{\rm lim} \gtrsim 0.024$ mas yr^{-1}, roughly equal to the intrinsic velocity dispersion of the galaxy. Those are fainter/high-magnitude stars that likely have underestimated errors. Hence, for further analysis we removed all stars whose PM uncertainties exceed the threshold of $\epsilon_{\rm lim} = 0.024$ mas yr^{-1}. In Section 5.3 we further test the impact of this choice.

Given the possible issues related to stars with large PMs, we decided to also impose a 3σ cut on our PM sample. This can jointly remove unwanted interlopers and remaining stars with underestimated errors. Comparison of the solid and opaque lines in Figure 9 shows that this does not strongly change the inferred $\sigma_{\rm POS}$. Nonetheless, the downside of any velocity cut is that it yields a slight underestimate of the true velocity dispersion (essentially, ∼99.73% of the true value for a 3σ cut of a Gaussian). To assure that this does not bias our dynamical modeling, which depends in part on comparison of LOS and PM kinematics, we performed the same cut in our LOS data set (on top of the previous membership probability cut). This did not significantly change the LOS data set, which already had a cut within a few σ owing to the larger fraction of interlopers.

Our final data set is the most complete and accurate PM catalog of a dSph to date, comprising 364 well-measured stars. Comparatively, Figure 10 shows that it comprises nearly 10 times more stars than in Massari et al. (2020; orange squares), it comprises twice as many stars as in del Pino et al. (2022; blue squares), and it reaches much deeper magnitudes than both data sets could ever do given their necessity for Gaia measurements. Moreover, the uncertainties in our PM measurements are all below the local PM dispersion (see dashed gray line), compared to no such stars in both previous studies. This improvement is particularly important because it is difficult to accurately constrain the PM dispersion of a galaxy based on individual PM measurements with uncertainties that do not resolve this dispersion (which is further compounded by known Gaia systematics; e.g., Fardal et al. 2021; Vasiliev & Baumgardt 2021).

2.3.6. Proper-motion Dispersion Profiles

Having constructed the PM catalog, we proceed to construct velocity dispersion profiles that will be used throughout the next sections. As in Section 2.2.3, all our computations of the dispersion σ of a given random variable follow the recipe presented in van der Marel & Anderson (2010), also recently employed in Vitral et al. (2023a). This consists of a maximum likelihood fit of a Gaussian distribution to the data, aiming to recover the respective standard deviation of the fit. The bias and uncertainty of such an estimate (e.g., Kenney & Keeping 1951) are corrected in a Monte Carlo sense, where we analyze numerous pseudo–data sets in the same fashion as the real data (see Appendix A from van der Marel & Anderson 2010 for details).

For spherically symmetric models of Draco, the velocity dispersion profile can be written as a function of the projected distance to the galaxy's center, R. We thus create logarithmically separated data bins in R whenever we need to visualize σ. In practice, our spherical modeling deals with discrete data (see Section 3.1), such that the bin choices we use to visualize our results do not actually matter for the fitting procedure. For the axisymmetric case, however, σ will also depend on the projected angle ξ of the data bin, defined as the angle between a given point and the projected major axis of the galaxy. Besides, our fitting approach in this case is frequentist (see Section 3.2), such that the binning process requires more attention.

Our LOS kinematics are based on fits of all position angles along an annulus, with the rotation amplitude in Figure 4 pertaining to the value on the kinematic major axis.[24] Instead, for the PMs we have measurements only for specific fields (see Figure 6) that span a small range of position angles. Therefore, whenever sampling velocity moments as explained in Section 3.2, these moments are averaged over all sky angles for the LOS, while we take the mean sky angle of each radial bin for the POS directions (namely, POSr for the POS radial direction and POSt for the POS tangential direction). We use the major axis (i.e., $\xi = 0°$) for comparison to the rotation amplitude.

The inferred velocity dispersion profiles in the three orthogonal directions are shown in Figure 11, together with the LOS rotation curve, all with similar x- and y-scales.[25] This provides, for the

[24] While our further modeling assumes an oblate dSph with maximum rotation on the equatorial plane, the data points used in our fits pertain to the kinematic major axis found in Section 2.2.2, which did not align exactly with Draco's major projected axis but was consistent within the uncertainties.

[25] The adopted galaxy distance (used to transform mas yr^{-1} to km s^{-1}) and the model predictions in this figure will be discussed in Section 4.2.2. The distance $74.98^{+3.95}_{-3.21}$ kpc is from the model fit and is close to the RR Lyrae estimate from Bonanos et al. (2004, namely $D = 75.8 \pm 5.4$ kpc).

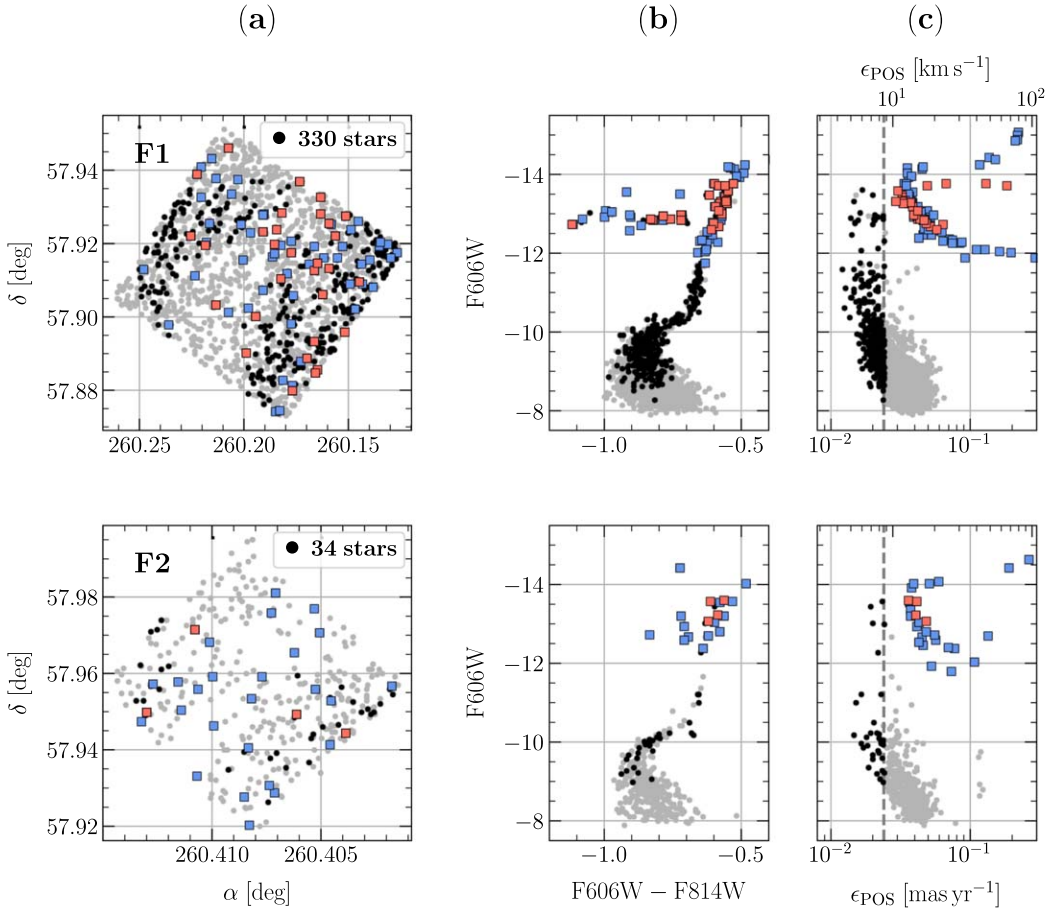

Figure 10. Data overview. (a) Zoom-in of studied fields. Black circles indicate stars that have PM accuracies good enough to fulfill our scientific goals (i.e., $\epsilon_{\rm POS} \lesssim \sigma_{\rm POS}$) and that were used in our dynamical modeling, while the gray circles indicate stars that do not, or that failed our multiple steps of data cleaning. The HST +Gaia sample compiled with GAIAHUB (del Pino et al. 2022) is marked in blue, while the orange stars depict the data set used by Massari et al. (2020). (b) CMDs based on HST data (using instrumental magnitudes), using the same symbols as in panel (a). (c) F606W magnitude as a function of the 1D PM error (calculated as in Equation (B2) from Lindegren et al. 2018), using the same symbols as in panel (a), with a dashed line representing the $\epsilon_{\rm POS}$ threshold we use. Our data set comprises hundreds of stars with $\epsilon_{\rm POS} \lesssim \sigma_{\rm POS}$, compared to zero such stars in previous PM analyses of Draco.

first time, radially resolved 3D velocity dispersion profiles for any dwarf galaxy. Focusing on the observations, we note that the radial PM dispersion is considerably higher than the tangential PM dispersion. Averaged over all radii probed, $\langle\sigma_{\rm POSt}\rangle/\langle\sigma_{\rm POSr}\rangle = 0.80 \pm 0.08$. The ratio of the LOS dispersion to the PM dispersion is somewhat closer to unity, $\langle\sigma_{\rm LOS}\rangle/\langle\sigma_{\rm POS}\rangle = 1.08 \pm 0.09$, where $\sigma_{\rm POS}$ represents an average over both PM directions. The first ratio is independent of galaxy distance, while the second is inversely proportional to it.

The tight observational constraints on ratios like these enable dynamical models of the kinds discussed in Section 3 below to strongly constrain the structure of Draco. To understand why, consider first the ratio $\sigma_{\rm POSt}/\sigma_{\rm POSr}$, which is a measure of the projected velocity dispersion anisotropy in the POS. In spherical geometry, Leonard & Merritt (1989) and van der Marel & Anderson (2010) both showed that there is a direct relation between this projected anisotropy and the intrinsic 3D velocity dispersion anisotropy. In Appendix B.3 we use scale-free dynamical models of the type discussed in de Bruijne et al. (1996) to show that the same is expected to hold in axisymmetric geometry. The details of the relation depend on quantities that are constrained by observational data, such as the projected axial ratio of the system, the position angle of the tracers on the sky, the radial profiles of the luminous and DM densities, the viewing inclination of the galaxy, etc. But in essence, $\sigma_{\rm POSt}/\sigma_{\rm POSr}$ is a diluted measure (i.e., brought closer to unity owing to projection effects) of the intrinsic 3D ratio $\sigma_{\rm tan}/\sigma_r$ (see Figure 16 in Appendix B). Thus, the observed $\langle\sigma_{\rm POSt}\rangle/\langle\sigma_{\rm POSr}\rangle$ implies the presence of radial velocity dispersion anisotropy in Draco. With suitable dynamical modeling, quantitative constraints are obtained on the shape of the 3D velocity dispersion tensor. This then breaks the mass–anisotropy degeneracy that plagues modeling of LOS velocities alone (Binney & Mamon 1982), so that the DM density profile can be determined. And with the 3D anisotropy known, the ratio $\sigma_{\rm LOS}/\sigma_{\rm POS}$ allows a kinematic determination of the galaxy distance (as done previously for globular clusters; e.g., Watkins et al. 2015b).

3. Methods

Various techniques have been used to model the velocity dispersion profiles of dSphs and thus constrain their mass density profiles (see Section 1). In this work, we employ multiple techniques to exploit our data set. This helps us to understand any modeling uncertainties and makes best use of the different codes available in the literature. We summarize them below.

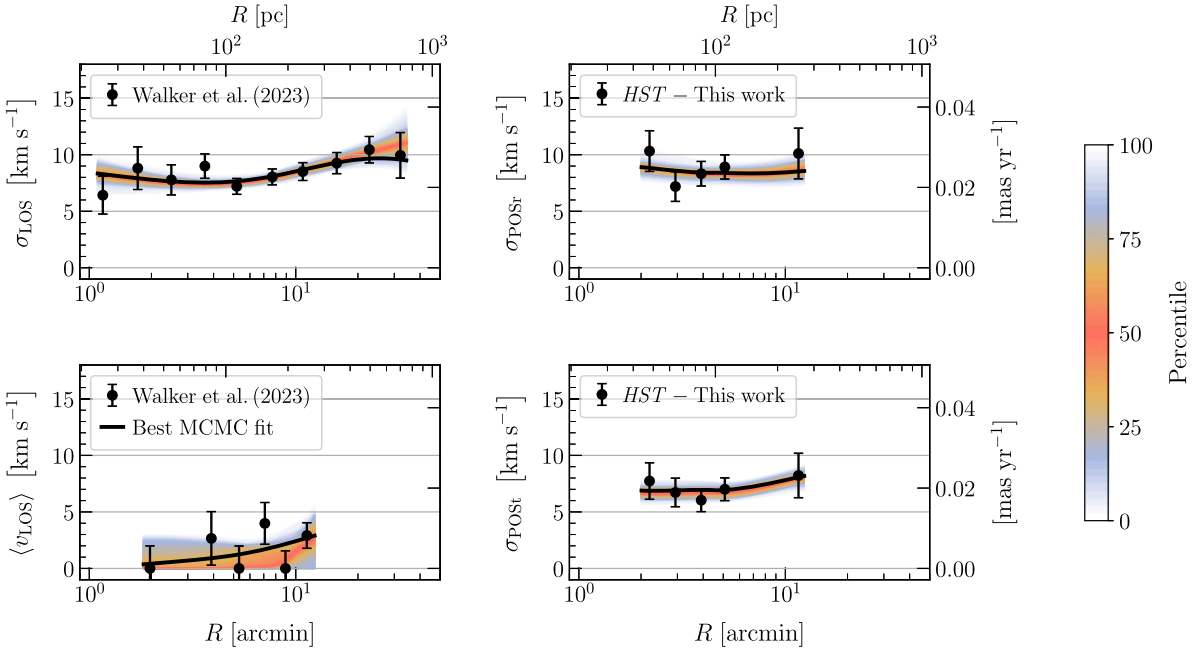

Figure 11. Observations and model comparison. The quantities showcased are velocity dispersion in the LOS (top left), POS velocity dispersion in radial direction (top right), LOS rotation amplitude (bottom left), and POS velocity dispersion in tangential direction (bottom right). The black circles and error bars represent the data, computed from the catalog of Walker et al. (2023) in the left panels and from our HST program in the right panels. Model predictions and the adopted galaxy distance (which we use to convert mas yr^{-1} to km s^{-1} in the right panel) are from our axisymmetric JAMPY MCMC fits for $i = 57.^{\circ}1$ as discussed in Section 4.2.2. Our best fit (as defined in Section 3.2) is depicted as a black solid line, which we interpolate with respect to projected radius R from the actual data R values (this is done for visualization purposes, since there is also a dependence with the projected angle ξ). The percentiles from our MCMC chains are color coded as in the color bar scheme, on the right.

3.1. Spherical Jeans Modeling: MAMPOSST-PM

Although Draco, like many other dSphs, is a flattened system (e.g., Table 1 and Figure 6), previous studies have, in general, considered spherical models to fit its internal kinematics (e.g., Read et al. 2018; Massari et al. 2020). Hence, it is useful to perform a similar kind of modeling if one wants to better interpret and compare previous results that assumed sphericity.

We perform spherical mass modeling with the Bayesian code MAMPOSST-PM (G. A. Mamon & E. Vitral, in preparation), which is an extension of MAMPOSST (Mamon et al. 2013), to handle PMs in addition to LOS velocities. MAMPOSST-PM is briefly described in Section 2 of Vitral & Mamon (2021) and was tested by Read et al. (2021), who showed that MAMPOSST-PM reproduced well the radial profiles of mass density and velocity anisotropy of mock dSphs. MAMPOSST-PM is also a faster code than its mass-modeling counterparts (see Table 2 from Read et al. 2021), which allows us to probe a wide range of dynamical models in less time, which is useful when defining priors and fitting boundaries (see Section 3.2).

3.1.1. General Formalism of MAMPOSST-PM

MAMPOSST-PM fits models for the radial profiles of total mass and the velocity anisotropy of the visible stars to the distribution of these stars in projected phase space. The local velocity ellipsoid is assumed to be an anisotropic Gaussian, whose axes are aligned with the spherical coordinates. The sizes of the axes are obtained by solving the spherical Jeans equation (Binney 1980)

$$\frac{d(\nu\sigma_r^2)}{dr} + 2\,\frac{\beta_{\rm B}(r)}{r}\,\nu(r)\sigma_r^2(r) = -\nu(r)\frac{G\,M(r)}{r^2}, \qquad (4)$$

assuming a given mass profile $M(r)$ and velocity anisotropy profile $\beta_{\rm B}(r)$, for a previously determined mass density profile $\nu(r)$ for the kinematic tracers (here stars). The term $\nu\,\sigma_r^2$ is the dynamical pressure that counteracts gravity.[26] The Binney (1980) velocity anisotropy ("anisotropy" for short) is defined as

$$\beta_{\rm B} \equiv 1 - \frac{\langle v_\theta^2\rangle + \langle v_\phi^2\rangle}{2\langle v_r^2\rangle}, \qquad (5)$$

where the $\langle v_i^2\rangle$ are the second-order velocity moments in spherical polar coordinates. In both spherical and axisymmetric geometry, the first moments $\langle v_r\rangle = \langle v_\theta\rangle = 0$, so that the corresponding velocity dispersions satisfy $\sigma_r^2 = \langle v_r^2\rangle$ and $\sigma_\theta^2 = \langle v_\theta^2\rangle$. In addition, in spherical geometry $\langle v_\theta^2\rangle = \langle v_\phi^2\rangle$. The first azimuthal moment $\langle v_\phi\rangle$ need not generally be zero, so that in general $\langle v_\phi^2\rangle = \sigma_\phi^2 + \langle v_\phi\rangle^2$. Our spherical models are constructed to have $\langle v_\phi\rangle = 0$, but in the axisymmetric models that we present later we do allow for the possibility of mean rotation.

In MAMPOSST-PM, the likelihood is written as

$$\mathcal{L} = \prod_i p(\boldsymbol{v}_i\,|R_i), \qquad (6)$$

[26] The Jeans equation (Equation (4)) is a consequence of the collisionless Boltzmann equation, which considers the incompressibility in phase space of the six-dimensional (6D) DF. Expressing the DF in terms of 6D number, mass, or luminosity density implies that the term ν in the Jeans equation is the number, mass, or luminosity density. For the present case of a dSph made of stars, it makes more physical sense to reason with mass density. For such systems, the mass density is proportional to the number density given the lack of substantial mass segregation, so the mass density profile is obtained from deprojecting the observed surface number density profile and multiplying it by a constant factor.

where the conditional probability of measuring a velocity $\mathbf{v_i}$ at projected radius $R_\mathbf{i}$ is the mean of the local velocity DF, $h(\mathbf{v} \mid R, r)$, integrated along the LOS

$$p(\boldsymbol{v} \mid R) = \frac{2}{\Sigma(R)} \int_R^\infty h(\boldsymbol{v} \mid R, r)\nu(r) \frac{r}{\sqrt{r^2 - R^2}}\, dr. \quad (7)$$

MAMPOSST-PM determines the marginal distributions of the free parameters and their covariances by running the MCMC routine COSMOMC[27] (Lewis & Bridle 2002).

In practice, we use six MCMC chains run in parallel and stop the exploration of parameter space after one of the chains reaches a number of steps $N_{\rm steps} = 10{,}000 N_{\rm free}$, where $N_{\rm free}$ is the number of free parameters of the model. We discard the first $3000 N_{\rm free}$ steps of each MCMC chain, which are associated with a burn-in phase. From the resulting chain values, we assign uncertainties to our best likelihood parameters using the 16th and 84th percentiles of the respective posterior distribution. If the fit lay below (above) those boundaries, we extended the uncertainty down to (up to) the minimum (maximum) chain value.

3.1.2. Parameterizations and Priors of MAMPOSST-PM

MAMPOSST-PM is a parametric code that fits discrete data. The motivations for our choices of parameterization are described further below. Our choice of priors, on the other hand, is performed to maximize the entropy of the posterior probability distribution. This can be done by assigning flat priors whenever we assume no previous knowledge on a specific parameter or Gaussian priors whenever we trust a previous measurement, from a different data set, with reported mean and uncertainty.

The anisotropic runs of MAMPOSST-PM use the generalization (hereafter gOM) of the Osipkov–Merritt model (Osipkov 1979; Merritt 1985; Equation (8a)) or the generalization (hereafter gTiret) of the Tiret et al. (2007; Equation (8b)) model for the velocity anisotropy profile:

$$\beta_{\rm B,gOM}(r) = \beta_0 + (\beta_\infty - \beta_0)\, \frac{r^2}{r^2 + r_\beta^2}, \quad (8a)$$

$$\beta_{\rm B,gTiret}(r) = \beta_0 + (\beta_\infty - \beta_0)\, \frac{r}{r + r_\beta}, \quad (8b)$$

where r_β is the anisotropy radius. We fit β_0 and β_∞ using flat priors from -1.99 to 1.99 to the symmetrized quantity[28] $\beta_{\rm sym} = \beta/(1-\beta/2)$, while fixing r_β to the scale radius of the luminous tracers.[29] Notice that a constant-anisotropy case is obtained by fixing $\beta_\infty = \beta_0$.

The mass density of the luminous tracers is chosen to be a Plummer model, similarly to what was adopted in Massari et al. (2020) and Hayashi et al. (2020), and also supported by our previous fits of Gaia ERD3 data (see Figure 1 in Section 2.1.2). We fit the Plummer r_{-2} radius[30] with Gaussian priors, using the mean and uncertainty from our fits. The total luminous mass of Draco, $M_\star$, was estimated by Martin et al. (2008) from its CMD, by assuming either a Kroupa et al. (1993) or a Salpeter (1955) initial mass function (IMF). We estimate the mean and variance of $\log M_\star$ from both of those values and use them as a Gaussian prior, which encompasses both estimates within 1σ.

We test numerous parameterizations for the DM density profile, including

1. a generalized Plummer model, which is a special case of the $\alpha\beta\gamma$ model by Zhao (1996), with $\alpha = 2$ and $\beta = 5$;
2. the Kazantzidis et al. (2004b) model, which is motivated by N-body simulations of tidally stripped cuspy DM halos;
3. the generalized Navarro–Frenk–White (NFW) profile, motivated by cosmological simulations by Navarro et al. (1997); and
4. the Einasto (1965) profile, which was used recently by Jiao et al. (2023) to fit the MW DM halo.

These density models are all listed in Appendix A and depend on three quantities: a scale radius,[31] a total DM mass ($M_{\rm dark}$, or M_{-2} for the generalized NFW model), and an inner slope γ (n index for the Einasto model). We assume flat priors for all these variables:

1. $\gamma \in [-2, 2]$, which encompasses both cuspy ($\gamma = -1$) and cored ($\gamma = 0$) cases (respectively, $n \in [0.1,\ 10]$). We allow for positive slopes so as to not rule out possible physical mechanisms unforeseen by ΛCDM.
2. $\log(M_{\rm dark}/[M_\odot]) \in [6,\ 12]$. Read et al. (2017) extrapolated classical $M_\star - M_{200}$ relations to lower-mass dSphs, such that Draco, with a luminous mass $\sim 5 \times 10^5\, M_\odot$, is predicted to have $M_{200} \sim 10^9\, M_\odot$. Hence, our priors largely encompass that range.
3. $\log(r_{-2}/[\rm kpc]) \in [-1,\ 1]$. Given the expected M_{200} mass from Read et al. (2017),[32] the concentration relation from Dutton & Macciò (2014) yields $r_{-2,\rm NFW} \sim 1$ kpc. Our prior thus encompasses this range within an order of magnitude.

Finally, we set Gaussian priors for the bulk $v_{\rm LOS}$ of Draco, using our estimate depicted in Table 1, and we also set Gaussian priors for the distance modulus, defined as $\mu_0 = 5\log(D/[\rm kpc]) + 10$. The mean and uncertainty on the distance modulus are derived by propagating the value and respective uncertainty on the RR Lyrae estimate from Bonanos et al. (2004), which yields $\mu_0 = 19.398 \pm 0.156$.

3.2. Axisymmetric Jeans Modeling: JAMPY

To model our data set under the assumption of an oblate axisymmetric galaxy, we use the publicly available code JAMPY (Cappellari 2008, 2020), tailored to the analysis of axisymmetric systems. This software was shown to reproduce well the dynamics of mock oblate dSphs with rotation (Sedain & Kacharov 2023) and has been applied in Zhu et al. (2024) to recover DM structural parameters of thousands of galaxies.

3.2.1. General Formalism of JAMPY

We use the version of JAMPY in which the velocity ellipsoid is aligned with spherical coordinates, given that we assume a

[27] https://cosmologist.info/cosmomc/

[28] $\beta_{\rm sym}$ runs from -2 for a model with only circular orbits to $+2$ for a model with only radial orbits, given that β ranges between $-\infty$ and $+1$, respectively, for these cases.

[29] This choice has been show to provide a better fitting convergence in Vitral & Mamon (2021).

[30] This is defined as the radius where $\frac{d\log\rho}{d\log r} = -2$.

[31] While Appendix A and Table 2 display the usual scale radii for those parameterizations, MAMPOSST-PM fits the r_{-2} quantity.

[32] We kindly thank Justin Read for sharing his algorithm to compute the precise M_{200} value expected for Draco.

spherical global potential, which would be only minimally altered by Draco's luminous axisymmetric component. This configuration of JAMPY considers the Jeans equations for rotating oblate systems (e.g., Bacon et al. 1983)

$$\frac{\partial(\nu\langle v_r^2\rangle)}{\nu\,\partial r} + \frac{(1+\beta_{\rm J})\langle v_r^2\rangle - \langle v_\phi^2\rangle}{r} = -\frac{\partial\Phi}{\partial r}, \tag{9a}$$

$$\frac{(1-\beta_{\rm J})}{\nu}\left[\frac{\partial(\nu\langle v_r^2\rangle)}{\partial\theta} + \frac{\nu\langle v_r^2\rangle}{\tan\theta}\right] - \frac{\langle v_\phi^2\rangle}{\tan\theta} = -\frac{\partial\Phi}{\partial\theta}, \tag{9b}$$

where Φ is the gravitational potential, the symbol $\langle . \rangle$ indicates the DF-averaged quantity, and, finally, $\beta_{\rm J}$ is defined as

$$\beta_{\rm J} \equiv 1 - \frac{\langle v_\theta^2\rangle}{\langle v_r^2\rangle} = 1 - \frac{\sigma_\theta^2}{\sigma_r^2}, \tag{10}$$

where the second equality assumes that $\langle v_\theta\rangle = \langle v_r\rangle = 0$, as in MAMPOSST-PM. Because of symmetry and continuity, axisymmetric models always have $\langle v_\phi\rangle = 0$ and $\langle v_\phi^2\rangle = \langle v_\theta^2\rangle$ along the symmetric axis. Hence, along the symmetry axis, β as defined by Equation (5) equals $\beta_{\rm J}$. Models with $\beta_{\rm J} = 0$ yield the same predicted second velocity moments as models in which the DF $f(E, Lz)$ does not depend on a third integral. Such models have been widely used for fitting data of axisymmetric systems (e.g., van der Marel 1991). Away from the symmetry axis, models with $\beta_{\rm J} = 0$ do not have an isotropic velocity dispersion tensor.

Beyond the nonsphericity, another main difference between MAMPOSST-PM and JAMPY is that the latter allows us to model Draco's rotation, by not imposing $\langle v_\phi\rangle = 0$ throughout the whole system (this equality was also imposed in many previous analyses such as Read et al. 2018; Hayashi et al. 2020; Massari et al. 2020). The first moment in the ϕ-direction relates to the second-order moment in the radial direction through

$$\langle v_\phi\rangle^2 = \langle v_\phi^2\rangle - \sigma_\phi^2, \tag{11a}$$

$$\sigma_\phi^2 = (1-\Omega)\langle v_r^2\rangle, \tag{11b}$$

where the rotation parameter Ω is introduced. This parameter was named γ in Cappellari (2020), but we change this notation to avoid confusion with the inner slope of the DM mass density, which uses the same symbol.

From those equations, JAMPY samples projected velocity moments that we use to compute the respective quantities in the LOS and POS directions. To fit our data set, we employ an MCMC chain using the EMCEE routine that minimizes the χ^2, defined as

$$\chi^2 = \chi^2_{\rm LOS} + \chi^2_{\rm POSr} + \chi^2_{\rm POSt} + \chi^2_{\rm rot}, \tag{12}$$

where $\chi^2_x = \sum(x_{\rm data} - x_{\rm model})_i^2/\epsilon^2_{x,i}$. In Equation (12), the four χ^2 terms pertain to the LOS, POSr, and POSt velocity dispersions at a given (R, ξ)[33] point, while the last term pertains to the first-order moment of the LOS velocity on the major axis.[34]

We then maximize the log-probability of our data set with the set of parameters Θ, defined as $\ln \Pr\{\Theta\} = -\chi^2/2$, along with respective priors defined further in Section 3.2.2. Our MCMC routine sets a maximum of 10,000 iterations per fit, which we run in parallel in 64 CPUs. In those configurations, each run takes ∼3 days to complete, and we perform it for a different set of possible inclinations from the pdf derived in Section 2.1.3. We discard the burn-in phase by removing the first 5000 steps of the chain and visually checking that the chains remain stable further on.

3.2.2. Parameterizations and Priors of JAMPY

As mentioned above, the timescales to run converging JAMPY models are drastically longer than respective MAMPOSST-PM runs, which can be explained by both the software languages employed in each code (PYTHON vs. FORTRAN, respectively) and the choice of parameterizations: while MAMPOSST-PM uses analytical parameterizations for a set of different models, JAMPY assumes multi-Gaussian expansions (MGEs) to model both the potential and the stellar distribution, which allows for more general density profiles at the expense of more time.

Therefore, we use our results from the spherical Jeans modeling to assist our fitting choices with JAMPY. For instance, since we observed no significant preference for a particular DM density parameterization in our spherical modeling results (see Section 4.1), we here decide to use only the generalized Plummer profile, as its analytical expressions are more easily handled when building MGEs in PYTHON. In the absence of external constraints on the geometrical shape of Draco's DM halo, we continue to assume that it is spherical,[35] even when the luminous density is chosen to be axisymmetric (so as to fit the observed projected shape of Draco). Although we use the same priors as MAMPOSST-PM for the DM density parameters[36] and Draco's distance, we fix the stellar density parameters,[37] as we observed no departure from the mean MAMPOSST-PM Gaussian priors. In addition, we assume the $\beta_{\rm J}$ parameter to be a constant,[38] since we show in Section 4 that the data do not prefer more general profiles such as the Osipkov–Merrit generalization of Equation (8a). Indeed, because MAMPOSST-PM assumes no rotation and spherical symmetry, such that $\sigma_\theta = \sigma_\phi$, its velocity anisotropy parameter is equal to JAMPY's $\beta_{\rm J}$.

Finally, we observed that when assuming a spatially constant rotation parameter Ω we could not fit well enough Draco's observed rotation curve. Hence, we assume the more general behavior

$$\Omega(r) = \Omega_0 + (\Omega_\infty - \Omega_0)\frac{1}{1+(r_\Omega/r)^2}. \tag{13}$$

We fit $(\Omega_0, \Omega_\infty)$ by assuming flat priors from −1.99 to 1.99 to the symmetrized quantity $\Omega_{\rm sym} = \Omega/(1-\Omega/2)$, while fixing r_Ω to the luminous scale radius.

As a consistency check that the fitting strategies of the different Jeans modeling algorithms we use do not strongly

[33] R is the projected radius, and ξ is the respective position angle in the POS.
[34] Our PM analysis methodology does not allow us to measure any mean streaming (see Section 2.3.3), so it is not included in the χ^2.
[35] While cosmological simulations tend to favor generally triaxial DM halos (e.g., Jing & Suto 2000; Kazantzidis et al. 2004a), recent observational studies of the MW DM halo support a quasi-spherical potential within the inner ∼30 kpc (Wegg et al. 2019; Hattori et al. 2021).
[36] With exception of the DM scale radius, to which we allow a larger prior toward higher radii.
[37] This means that we fix $M_\star = 4.7\times10^5\,M_\odot$ and the major axis of the projected density as the value fitted in Section 2.1.1, namely $a_{{\rm maj},\star} = 9\farcm1$, where $a_{\rm maj}$ is the respective Plummer major axis (see Appendix A).
[38] A similar assumption is also present in Hayashi et al. (2020), who base their choices on cosmological simulations by Vera-Ciro et al. (2014). For robustness purposes, we also ran a test with a gOM-like parameterization for $\beta_{\rm J}$ and observed no departure from the constant case.

Table 2
Main Results of the MAMPOSST-PM Spherical Jeans Modeling

ID	$\rho_{\rm dark}$	Test	D (kpc)	β_0	β_∞	$r_\star$ (10^2 pc)	$M_\star$ ($10^5\,M_\odot$)	$r_{\rm dark}$ (10^2 pc)	$M_{\rm dark}^{R_{\rm max}}$ ($10^8\,M_\odot$)	$\gamma_{\rm dark}$ or n	$\Gamma_{\rm dark}$	ΔAICc
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
1	GKAZ	$\rho_{\rm dark}$	$75.76^{+3.23}_{-4.22}$	$0.45^{+0.07}_{-0.19}$	⋯	$1.74^{+0.10}_{-0.11}$	$4.25^{+3.16}_{-1.16}$	$1.93^{+12.17}_{-0.27}$	$1.34^{+3.79}_{-0.14}$	$0.76^{+0.32}_{-1.43}$	$-0.11^{+0.28}_{-0.79}$	0.00
2	EIN	$\rho_{\rm dark}$	$76.73^{+3.03}_{-4.28}$	$0.40^{+0.06}_{-0.23}$	⋯	$1.77^{+0.09}_{-0.12}$	$5.05^{+2.24}_{-1.97}$	$4.91^{+19.38}_{-4.91}$	$1.45^{+4.49}_{-1.45}$	$0.70^{+7.80}_{-0.60}$	$-0.31^{+0.26}_{-0.34}$	0.11
3	GPLU	$\rho_{\rm dark}$	$75.99^{+2.89}_{-4.52}$	$0.39^{+0.13}_{-0.14}$	⋯	$1.76^{+0.08}_{-0.13}$	$5.07^{+2.06}_{-2.07}$	$8.68^{+14.83}_{-4.04}$	$1.41^{+0.68}_{-0.43}$	$-0.24^{+1.25}_{-0.55}$	$-0.41^{+0.88}_{-0.42}$	0.17
4	GNFW	$\rho_{\rm dark}$	$74.74^{+4.76}_{-2.65}$	$0.41^{+0.09}_{-0.18}$	⋯	$1.75^{+0.11}_{-0.09}$	$4.30^{+2.99}_{-1.34}$	$2.47^{+15.46}_{-0.06}$	$0.80^{+1.62}_{-0.01}$	$1.44^{+0.01}_{-2.00}$	$-0.24^{+0.01}_{-0.72}$	0.34
5	GPLU	$\beta_{\rm gOM}$	$74.85^{+4.61}_{-3.10}$	$0.48^{+0.08}_{-0.21}$	$-0.98^{+1.78}_{-0.92}$	$1.74^{+0.11}_{-0.10}$	$4.12^{+3.29}_{-1.10}$	$4.94^{+16.52}_{-0.88}$	$1.14^{+1.54}_{-0.12}$	$0.43^{+0.75}_{-1.22}$	$-0.17^{+0.45}_{-0.69}$	1.99
6	GPLU	$\beta_{\rm gTiret}$	$75.87^{+3.51}_{-4.04}$	$0.52^{+0.19}_{-0.23}$	$-0.21^{+0.95}_{-1.26}$	$1.75^{+0.10}_{-0.10}$	$4.70^{+2.69}_{-1.67}$	$6.47^{+18.15}_{-2.25}$	$1.25^{+1.05}_{-0.32}$	$-0.03^{+1.10}_{-0.81}$	$-0.35^{+0.71}_{-0.54}$	1.76
7	GPLU	Cusp	$77.27^{+3.72}_{-3.34}$	$0.25^{+0.11}_{-0.17}$	⋯	$1.78^{+0.11}_{-0.10}$	$5.04^{+2.10}_{-2.06}$	$28.42^{+83.80}_{-7.71}$	$1.26^{+0.12}_{-0.10}$	$\gamma=-1$	$-1.00^{+0.00}_{-0.00}$	1.52
8	GPLU	Core	$74.57^{+3.97}_{-3.20}$	$0.40^{+0.12}_{-0.09}$	⋯	$1.72^{+0.11}_{-0.09}$	$4.33^{+3.02}_{-1.28}$	$6.92^{+1.93}_{-0.82}$	$1.37^{+0.37}_{-0.21}$	$\gamma=0$	$-0.29^{+0.10}_{-0.07}$	−2.16
9	GPLU	PM	$75.68^{+4.93}_{-5.65}$	$0.53^{+0.07}_{-0.18}$	⋯	$1.73^{+0.14}_{-0.13}$	$4.49^{+2.96}_{-1.45}$	$6.04^{+47.73}_{-2.31}$	$1.40^{+2.11}_{-0.70}$	$0.34^{+0.70}_{-1.13}$	$-0.09^{+0.27}_{-0.81}$	⋯
10	GPLU	ϵ	$73.59^{+5.20}_{-2.62}$	$0.56^{+0.04}_{-0.21}$	⋯	$1.69^{+0.14}_{-0.07}$	$5.08^{+2.19}_{-2.06}$	$5.79^{+8.53}_{-1.30}$	$1.61^{+1.04}_{-0.37}$	$0.70^{+0.64}_{-1.24}$	$0.18^{+0.33}_{-0.85}$	⋯

Note. Column (1): model ID. Column (2): DM parameterization: "GPLU" for a generalized Plummer (1911) model with free inner slope, "GKAZ" for a generalized Kazantzidis et al. (2004b) model with free inner slope, "GNFW" for a generalized Navarro et al. (1997) model with free inner slope, and "EIN" for the Einasto (1965) model. Column (3): test type: "$\rho_{\rm dark}$" when testing different parameterizations for the DM density profile, "$\beta_{\rm gOM}$" for a generalized Osipkov (1979)–Merritt (1985) parameterization of the velocity anisotropy profile, "$\beta_{\rm gTiret}$" for a generalized Tiret et al. (2007) parameterization of the velocity anisotropy profile, "Cusp" when forcing an inner density slope of −1 for the DM, "Core" when forcing a cored model for the DM, "PM" when only using PMs (no LOS data), and "ϵ" when using a lower PM error threshold. Column (4): heliocentric distance, in kpc. Column (5): anisotropy value at $r=0$. Column (6): anisotropy value at infinity (only for models with variable anisotropy). Column (7): Plummer scale radius of the stellar component, in 10^2 pc. Column (8): total mass of the stellar component, in $10^5\,M_\odot$. Column (9): DM scale radius, in 10^2 pc. Column (10): DM mass at the maximum projected data radius, in $10^8\,M_\odot$. Column (11): DM asymptotic density slope or Einasto index n. Column (12): DM density slope averaged over the spatial range where PMs are available. Column (13): difference in AICc relative to model 1. Listed uncertainties are based on the 16th and 84th percentiles of the marginal distributions, unless the maximum likelihood solution was outside that boundary, in which case the uncertainties are related to the minimum or maximum value of the MCMC chain. We did not consider the AICc diagnostic when the data set was different from the respective standard model.

diverge from each other, we compared two constant-anisotropy runs from MAMPOSST-PM and JAMPY for a spherical geometry and confirmed that both the inferred velocity anisotropy and the DM density slope differ by much less than their respective 1σ uncertainties.[39]

3.3. Practical Quantities

Given the choices of different DM parameterizations to compare with and the fact that our data are not complete at all radii, we define here practical quantities to help us interpret our results. For example, the total DM mass (or M_{-2} for NFW profiles) is not a well-constrained quantity, since our data do not really allow us to constrain in detail the shape of the DM density at large radii. Hence, a more suitable parameter to display and use for comparison purposes is the DM mass up to a fiducial radius. We do so by displaying further in Tables 2 and 3 the variable $M_{\rm dark}(r=R_{\rm max})$—i.e., the total mass of DM up to the maximum projected radius in our LOS+PM data set, namely $R_{\rm max}=900$ pc. To aid in the interpretation of this quantity, we also compute, at this same radius, the circular velocity

$$v_{\rm circ}(r)=\sqrt{\frac{G\,M(r)}{r}}, \tag{14}$$

which depends on the total[40] cumulative mass up to a certain radius r and on the gravitational constant G.

Similarly, due to the restricted spatial extent of our data, the inner DM slope parameter $\gamma_{\rm dark}$, or the respective Einasto index $n_{\rm dark}$, may not reflect accurately our fits and uncertainties of the DM slope where we are actually able to constrain it—i.e., where we have both PM and LOS data. We therefore define an effective density slope parameter as

$$\Gamma_{\rm dark}\equiv\frac{\int_{r_{\rm min}}^{r_{\rm max}}\frac{d\log\rho}{d\log r}\,\rho(r)\,dr}{\int_{r_{\rm min}}^{r_{\rm max}}\rho(r)\,dr}, \tag{15}$$

where $\rho(r)$ is the DM density. The $r_{\rm min}$ variable is defined as the minimum projected radius in the data where PM information is available. We define $r_{\rm max}=\min(R_{\rm max,\,PM},\,r_{\Lambda{\rm CDM}}/3)$, where $R_{\rm max,\,PM}$ is the maximum projected radius in the data where PM information is available and $r_{\Lambda{\rm CDM}}$ is the scale radius of a NFW profile as expected for DM halos in ΛCDM simulations[41] assuming low-mass stellar components (Read et al. 2017).

In practice, the radial limits over which we average the logarithmic DM density slope are $r_{\rm min}=42$ pc and $r_{\rm max}=297$ pc, which translate to roughly $r_{\rm min}=1\farcm9$ and $r_{\rm max}=22\farcm5$. If one considers the scale radius $r_{\Lambda{\rm CDM}}$, the conversion between $\Gamma_{\rm dark}$ and $\gamma_{\rm dark}$ in our PM radial range is such that cusp ($\gamma_{\rm dark}=-1$) and cored ($\gamma_{\rm dark}=0$) values translate to $\Gamma_{\rm dark}=-1.2$ and $\Gamma_{\rm dark}=-0.38$, respectively, for a generalized NFW profile. For a generalized Plummer profile such as used in our axisymmetric fits, the respective numbers are $\Gamma_{\rm dark}=-1.07$ and $\Gamma_{\rm dark}=-0.14$ for cusp and cored models, thus providing a subtler difference.

[39] Precisely, we measure $\Delta\beta=0.04$ and $\Delta\gamma=0.06$, while the uncertainty on each parameter for the spherical case is of the order of ∼0.15 and ∼1, respectively.

[40] The total mass is a sum of the luminous and dark components.

[41] We choose this value as a reference because we wish to compare our observables to what is predicted from theory, while the 1/3 factor is added with the intent of removing the part of the predicted density profile that has a cuspier drop due to the transition from the inner to the outer density profile.

Table 3
Main Results of the JAMPY Axisymmetric Jeans Modeling

i (deg) (1)	Ω_0 (2)	Ω_∞ (3)	D (kpc) (4)	β_J (5)	$r_{\rm dark}$ (10^2 pc) (6)	$M_{\rm dark}^{R_{\rm max}}$ ($10^8\,M_\odot$) (7)	$\gamma_{\rm dark}$ (8)	$\Gamma_{\rm dark}$ (9)	$\bar{\beta}_B$ (10)	$v_{\rm circ}^{R_{\rm max}}$ (km s^{-1}) (11)
43.0	$0.36^{+0.64}_{-6.42}$	$-2.14^{+2.95}_{-371.34}$	$81.32^{+4.34}_{-4.63}$	$0.95^{+0.02}_{-0.02}$	$9.43^{+102.82}_{-3.64}$	$1.91^{+1.37}_{-0.20}$	$-0.66^{+0.71}_{-0.24}$	$-0.77^{+0.47}_{-0.15}$	$-0.02^{+0.26}_{-0.34}$	$29.21^{+9.14}_{-1.34}$
47.7	$0.46^{+0.02}_{-7.22}$	$-1.95^{+2.55}_{-1.45}$	$76.83^{+3.39}_{-4.37}$	$0.78^{+0.03}_{-0.18}$	$6.22^{+131.43}_{-0.44}$	$1.41^{+0.81}_{-0.18}$	$-0.24^{+0.11}_{-0.80}$	$-0.51^{+0.10}_{-0.55}$	$-0.01^{+0.09}_{-0.67}$	$25.80^{+6.57}_{-1.64}$
52.4	$0.39^{+0.21}_{-6.67}$	$-2.42^{+3.10}_{-1.43}$	$73.33^{+5.16}_{-1.85}$	$0.51^{+0.18}_{-0.16}$	$6.94^{+74.25}_{-1.65}$	$1.06^{+0.72}_{-0.02}$	$-0.63^{+0.72}_{-0.51}$	$-0.79^{+0.46}_{-0.35}$	$-0.34^{+0.43}_{-0.42}$	$22.91^{+6.63}_{-0.44}$
57.1	$0.28^{+0.32}_{-8.10}$	$-2.49^{+3.11}_{-1.06}$	$74.98^{+3.95}_{-3.21}$	$0.35^{+0.20}_{-0.22}$	$7.01^{+295.92}_{-4.56}$	$0.88^{+0.67}_{-0.50}$	$-0.79^{+0.22}_{-0.46}$	$-0.94^{+0.22}_{-0.31}$	$-0.41^{+0.35}_{-0.46}$	$20.72^{+6.77}_{-7.22}$
61.8	$0.45^{+0.54}_{-8.20}$	$-1.59^{+2.33}_{-0.91}$	$73.19^{+5.51}_{-2.71}$	$0.45^{+0.04}_{-0.33}$	$10.71^{+289.80}_{-4.11}$	$1.16^{+0.31}_{-0.28}$	$-0.78^{+0.36}_{-0.45}$	$-0.84^{+0.25}_{-0.39}$	$-0.11^{+0.06}_{-0.59}$	$24.00^{+2.73}_{-3.27}$
66.5	$0.32^{+0.68}_{-8.32}$	$-1.46^{+2.13}_{-1.17}$	$73.42^{+4.59}_{-3.45}$	$0.39^{+0.07}_{-0.46}$	$10.16^{+276.71}_{-2.40}$	$1.02^{+0.34}_{-0.18}$	$-0.80^{+0.17}_{-0.45}$	$-0.87^{+0.13}_{-0.38}$	$-0.11^{+0.11}_{-0.77}$	$22.50^{+3.50}_{-2.17}$
71.2	$-5.23^{+4.91}_{-3.19}$	$0.55^{+0.23}_{-2.58}$	$75.99^{+2.30}_{-5.91}$	$0.24^{+0.21}_{-0.35}$	$70.10^{+209.14}_{-64.08}$	$1.30^{+0.01}_{-0.49}$	$-1.08^{+0.72}_{-0.22}$	$-1.08^{+0.50}_{-0.22}$	$-0.32^{+0.35}_{-0.51}$	$24.90^{+0.40}_{-5.07}$
75.9	$0.19^{+0.81}_{-6.54}$	$-1.66^{+2.32}_{-0.49}$	$72.44^{+5.44}_{-1.79}$	$0.24^{+0.20}_{-0.29}$	$12.74^{+381.65}_{-6.02}$	$0.94^{+0.34}_{-0.17}$	$-0.98^{+0.53}_{-0.31}$	$-1.03^{+0.37}_{-0.28}$	$-0.26^{+0.29}_{-0.42}$	$21.77^{+3.26}_{-2.23}$
80.6	$-4.87^{+4.59}_{-2.66}$	$0.39^{+0.32}_{-2.49}$	$73.07^{+5.56}_{-2.52}$	$0.17^{+0.23}_{-0.34}$	$24.04^{+260.36}_{-14.83}$	$1.13^{+0.13}_{-0.37}$	$-1.04^{+0.32}_{-0.32}$	$-1.05^{+0.24}_{-0.32}$	$-0.34^{+0.34}_{-0.48}$	$23.77^{+1.02}_{-4.45}$
85.3	$-5.70^{+5.34}_{-2.09}$	$0.81^{+0.19}_{-2.92}$	$74.32^{+4.53}_{-3.93}$	$0.19^{+0.19}_{-0.38}$	$173.12^{+141.10}_{-164.65}$	$1.29^{+0.63}_{-0.50}$	$-1.03^{+0.35}_{-0.28}$	$-1.03^{+0.24}_{-0.29}$	$-0.31^{+0.30}_{-0.53}$	$25.10^{+7.18}_{-5.55}$
90.0	$-0.05^{+0.06}_{-6.60}$	$-2.18^{+2.93}_{-318.09}$	$75.92^{+2.44}_{-5.28}$	$-0.01^{+0.39}_{-0.12}$	$16.82^{+294.31}_{-7.06}$	$0.87^{+0.39}_{-0.11}$	$-1.21^{+0.47}_{-0.14}$	$-1.23^{+0.42}_{-0.12}$	$-0.56^{+0.54}_{-0.19}$	$20.48^{+4.43}_{-1.13}$
$\langle . \rangle_i$	$\mathbf{0.15^{+0.24}_{-7.24}}$	$\mathbf{-1.14^{+1.85}_{-1.81}}$	$\mathbf{75.37^{+4.73}_{-4.00}}$	$\mathbf{0.56^{+0.25}_{-0.42}}$	$\mathbf{11.18^{+191.03}_{-4.89}}$	$\mathbf{1.20^{+0.79}_{-0.28}}$	$\mathbf{-0.71^{+0.44}_{-0.49}}$	$\mathbf{-0.83^{+0.32}_{-0.37}}$	$\mathbf{-0.20^{+0.28}_{-0.53}}$	$\mathbf{24.19^{+6.31}_{-2.97}}$

Note. Column (1): inclination, in degrees (90° is edge-on). Column (2): rotation parameter at $r = 0$ (see Equation (13)). Column (3): rotation parameter at $r \to \infty$. Column (4): heliocentric distance, in kiloparsecs. Column (5): β_J velocity anisotropy parameter, as defined in Equation (10). Column (6): DM scale radius, in 10^2 pc. Column (7): DM mass at maximum projected data radius, in $10^8\,M_\odot$. Column (8): DM asymptotic density slope. Column (9): DM density slope averaged over the spatial range where PMs are available. Column (10): globally averaged β_B velocity anisotropy, as defined in Equation (5). Column (11): circular velocity at maximum projected data radius, in kilometers per second. The uncertainties are based on the 16th and 84th percentiles of the marginal distributions, unless the maximum likelihood solution was outside that boundary, in which case the uncertainties are based on the minimum or maximum value of the MCMC chain. The last row (highlighted in bold) displays the integrated estimate for each parameter, averaged over the inclination probability distribution of Draco, as described in Section 4.2.1. The baryonic mass and respective Plummer major axis are fixed in all cases to $4.7 \times 10^5\,M_\odot$ and $9'.1$.

3.4. Statistical Tools

We employ Bayesian evidence to compare our different MAMPOSST-PM mass–anisotropy models and correct for over- and underfitting. This model selection involves comparing the maximum log posteriors using Bayesian information criteria. We use the corrected Akaike information criterion (derived by Sugiura 1978 and independently by Hurvich & Tsai 1989, who demonstrated its utility for a wide range of models)

$$\mathrm{AICc} = \mathrm{AIC} + 2\,\frac{N_{\rm free}\,(1 + N_{\rm free})}{N_{\rm data} - N_{\rm free} - 1}, \tag{16}$$

where AIC is the original Akaike information criterion (Akaike 1998)

$$\mathrm{AIC} = -2\ln\mathcal{L}_{\rm MLE} + 2\,N_{\rm free}, \tag{17}$$

and where $\mathcal{L}_{\rm MLE}$ is the maximum likelihood estimate found when exploring the parameter space, $N_{\rm free}$ is the number of free parameters, and $N_{\rm data}$ is the number of data points. We prefer AICc to the other popular simple Bayesian evidence model, the Bayes information criterion (BIC; Schwarz 1978), because AICc is more robust for situations where the true model is not among the tested ones (e.g., our choice of a Plummer density profile for the stellar component is purely empirical and not theoretically motivated), in contrast with BIC (Burnham & Anderson 2002).

The likelihood (given the data) of one model relative to a reference one is (Akaike 1983)

$$\exp\left(-\frac{\mathrm{AIC} - \mathrm{AIC}_{\rm ref}}{2}\right), \tag{18}$$

and we use it to infer likelihood probabilities. In general, ΔAICc differences $\gtrsim$4 (i.e., a confidence level $\gtrsim$85%, according to Equation (18) above) are required to prefer one model over another. It is also important to mention that such diagnostics are purely statistical and do not account for intrinsic astrophysical phenomena that might favor or disfavor a particular model.

4. Results

4.1. Spherical Modeling

We first present the results of spherical Jeans modeling with MAMPOSST-PM. Key outcomes are listed in Table 2. Our velocity dispersion goodness of fits for the spherical case were very similar to the ones presented in Figure 11 for the case of axisymmetric models, so we do not show the spherical fits separately (caveat: our spherical modeling neglects rotation, so the spherical model predictions in the bottom left panel are zero). Below, we detail our results.

4.1.1. Dark Matter Density Parameterization

The first four lines of Table 2 address the comparison between the four DM density parameterizations we use: Kazantzidis et al. (2004b) listed as GKAZ, Einasto (1965) listed as EIN, generalized Plummer (1911) with free inner slope listed as GPLU, and, finally, the generalized Navarro et al. (1997) with free inner slope listed as GNFW. We refer the reader back to Section 3.1.2 for the motivations of each parameterization and now focus on the practical fitting results.

A comparison of these models' AICc yields a modest preference for the GKAZ profile, followed by EIN, GPLU, and GNFW. However, the AICc differences reach, at most, 0.34 between the GKAZ and NFW profiles, which translates to the GKAZ model being $\sim$1.2 times more likely than GNFW (from Equation (18)). This is definitely not enough to robustly distinguish those models, meaning that they all fit the data equally well. This is true even though the GPLU yields a

slightly more cuspy DM profile, but not significantly so given the high uncertainties associated with this parameter when using spherical models. Therefore, given the analytical simplicity of the GPLU profile and its physical meaning,[42] we choose to use this model for further tests, as well as when modeling Draco with JAMPY further on.

4.1.2. Velocity Anisotropy Parameterization

Models 5 and 6 from Table 2 display our tests with different velocity anisotropy parameterizations, specifically the gOM and gTiret generalizations with free inner and outer anisotropy values. The results show that although the inner velocity anisotropy tends to agree with the constant-anisotropy case (i.e., preferring radial anisotropy), the velocity anisotropy at infinity prefers a more tangential behavior. Nevertheless, the uncertainties associated with this parameter are extremely large and basically encompass very radially anisotropic cases as well. Since the anisotropy at large radii is not well constrained by the data, the improvement in the fit is not enough for us to actually prefer more general models such as this (i.e., AICc is higher). Hence, we use the constant-anisotropy case here, and when performing axisymmetric modeling, we use it as our standard model. More general anisotropy parameterizations could overfit the data.

4.1.3. Cusp versus Core under the Spherical Assumption

One of the main goals of our dynamical modeling is to constrain the DM slope of Draco. From our fits that leave this slope as a free variable, we observe a general preference midway between a cored and a cuspy profile where the PM data are present (i.e., column (12)). A different test is then to directly compare two mass models, one with a fixed inner cusp (i.e., $\gamma = -1$) and another with a fixed core (i.e., $\gamma = 0$). We perform those runs and display them in Table 2 as models 7 and 8.

The cusped and cored models prefer velocity anisotropies that are consistent with each other at the 1σ levels, with cuspy models preferring slightly lower values. The AICc comparison between these two models shows a significant preference for the cored case, which is $\sim$6 times more likely than the cuspy counterpart, or, equivalently, rules out the latter with nearly 85% probability (from Equation (18)). This is opposite to what was found by the spherical modeling from Massari et al. (2020), although the much better completeness and accuracy of our data set (see Figure 10) can easily account for such differences. In the next section, we move on to test this result under the more suitable geometric assumption of axisymmetry.

4.1.4. Velocity Anisotropy under the Spherical Assumption

Along with the measurement of the DM slope, our study is the first to consistently constrain the velocity anisotropy of Draco. As discussed in Section 2.3.6, the projected anisotropy is directly measured by the data, independently of any model assumption. However, to recover the intrinsic 3D value of β, assumptions are required; here we analyze this problem under the consideration of spherical symmetry.

Throughout all models in Table 2, the velocity anisotropy β_B remains within the 1σ range of $\sim$0.2−0.6. This means that under spherical assumptions our data consistently constrain the orbital shapes in Draco to be radially anisotropic. Previous PM studies of Draco (Massari et al. 2020; del Pino et al. 2022) hint toward a similar behavior, although error bars are too high to discard tangential orbits at a 1σ level in the former work. Given the unprecedentedly low PM uncertainties of our data set, we are able to rule out negative values of β_B in model 3 with 98.5% confidence, if one assumes sphericity.

4.2. Axisymmetric Modeling

Having explored spherical models, we now move on to more realistic axisymmetric models. Because JAMPY is much slower than the equivalent MAMPOSST-PM MCMC routines, we are forced here to reduce our set of dynamical models. We thus use the priors and assumptions specified in Section 3.2.2. Specifically, informed by the results of the spherical models, we focus on axisymmetric luminous models with a Plummer density distribution and constant β_J as a function of radius, embedded in a spherical dark halo with a generalized Plummer profile.

4.2.1. Inclination Dependence

We ran JAMPY over a set of 11 inclinations linearly spaced from 43° to 90° (edge-on), which encompasses Draco's inclination pdf computed in Section 2.1.3, and display the results in Table 3. In addition to the parameter β_J defined by Equation (10), we also list the globally averaged Binney anisotropy $\bar{\beta}_B$. The latter was obtained by first calculating the mass-weighted second velocity moments averaged over the entire system and then substituting those into Equation (5).

Overall, the results for Ω_∞ and r_{dark} slightly increase with inclination, while other parameters tend to decrease with i. In particular, models that are closer to edge-on prefer more cuspy DM profiles, lower velocity anisotropies β_J (still positive, but closer to $\beta_J = 0$), and lower dynamical distances.

In spite of these correlations, we can provide a global estimate and uncertainty for every model parameter, integrated over the inclination distribution that we calculated for Draco in Section 2.1.3. That is, for a parameter Θ, we apply

$$p(\Theta) = \int_{i_{min}}^{i_{max}} p(\Theta|i)\, p(i)\, di, \tag{19}$$

where $p(\Theta)$ is the final posterior distribution integrated over all inclinations, $p(\Theta|i)$ is the posterior distribution obtained from the MCMC chain, and $p(i)di$ is the probability of falling into a certain inclination range, which we compute empirically from the distribution computed in Section 2.1.3. In practice, the integral above is a discrete sum over the inclination ranges probed by the 11 values we calculated. We use the same procedure to obtain the most probable value of every parameter, i.e., we average the results in Table 3 weighted by the respective inclination probabilities. Finally, we define the confidence regions as the 16th and 84th percentiles of the numerically computed $p(\Theta)$.

4.2.2. Data–Model Comparison

The parameter estimates thus marginalized over all inclinations are listed in the last row of Table 3. They resemble the estimates for the cases $i = 52.^\circ4$ and $i = 57.^\circ1$. We choose the case of $i = 57.^\circ1$ as a canonical model for display, as it represents a value between the median ($56.^\circ2$) and the mean ($58.^\circ9$) of the inclination pdf. In Figure 11 we display, for this

[42] Models such as GNFW do not have a finite mass, while EIN models are not able to reproduce centrally decreasing density profiles.

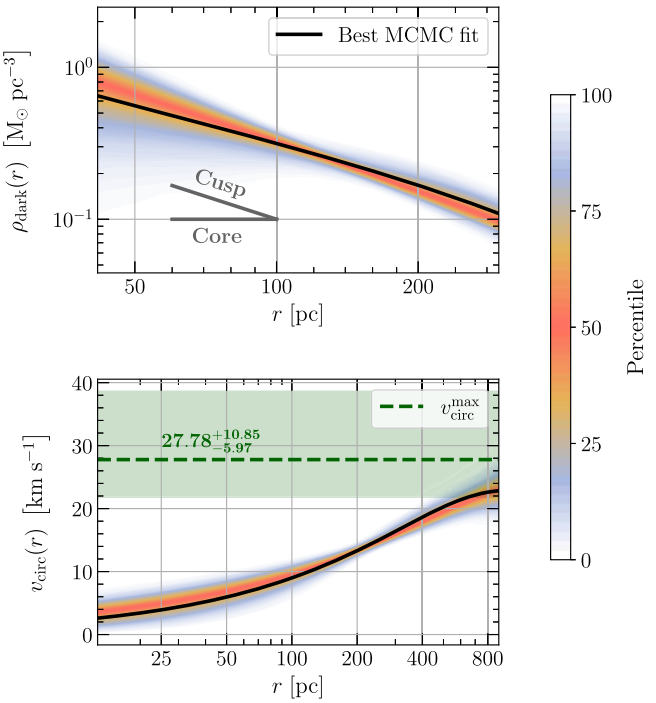

Figure 12. DM density and circular velocity. The top panel shows the DM density profiles from our JAMPY MCMC chains for $i = 57.^\circ1$, color coded according to their respective percentiles. We display the most probable solution (as defined in Section 3.2) as a black line. The bottom panel shows the circular velocity curves, defined by Equation (14), for the same models, as well as their maximum value in green. The radial extent of the top panel covers the extent of our PM data (where $\Gamma_{\rm dark}$ was computed), while the bottom panel covers the extent of our entire data set.

inclination, the data–model comparison for the three velocity dispersion components and the LOS rotation velocity amplitude. Figure 12 shows the DM density profile we estimate over the range where we calculate its slope, as well as the circular velocity over the entire data range.

Figure 11 shows a very satisfactory fit from JAMPY, which helps to strengthen both our parameterization choices and our conclusions further on. Figure 12 shows that both the circular velocity and DM cusp (down to ∼100 pc) are well constrained by our models.

We verified that our results do not depend sensitively on the adopted parameterization for $\rho_\star$. For this we ran a comparison $i = 57.^\circ1$ model in which the stellar density was taken to follow a Sérsic (Sérsic 1963, 1968) profile,[43] instead of the canonical Plummer one. This yielded a considerably higher χ^2 statistic, driven in part by a poorer fit to Draco's rotation profile. Nonetheless, the inferred results for, e.g., $\Gamma_{\rm dark}$ and $\bar{\beta}_{\rm B}$ (namely, $-0.70^{+0.64}_{-0.14}$ and $-1.04^{+0.66}_{-0.40}$, respectively) agree within 1σ with what is listed for $i = 57.^\circ1$ in Table 3.

4.2.3. Inferred Quantities

We display in Figure 13 the posterior distributions and correlation between the parameters $\Theta = \{\Gamma_{\rm dark}, \bar{\beta}_{\rm B}\}$, marginalized over all inclinations as described in Section 4.2.1. The figure shows that models with a core generally require a more radial velocity anisotropy to fit the data, consistent with what has been found in other contexts (e.g., Figure 3 in van der

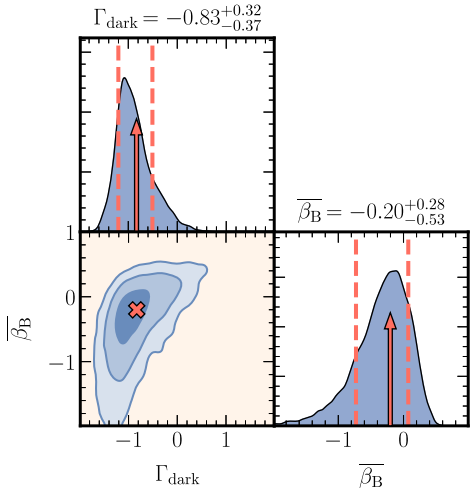

Figure 13. Averaged DM slope and velocity anisotropy: final posterior probability distributions of Draco's DM density slope, $\Gamma_{\rm dark}$, averaged over the range where we have PM data, and the globally averaged $\beta_{\rm B}$ velocity anisotropy (see Equation (5)). The distributions were marginalized over inclinations, as described in Section 4.2.1, and further smoothed for simple visualization purposes. The values highlighted correspond to the best likelihood estimates (title texts, arrows, and cross) and the respective 16th and 84th percentiles (dashed lines). The results imply a classic ΛCDM-like slope in the Draco Sph galaxy and a marginal preference for overall tangential velocity anisotropy.

Marel et al. 2000). Our averaged slope estimate, $\Gamma_{\rm dark} = -0.83^{+0.32}_{-0.37}$, is consistent with a classic ΛCDM slope, even though the posterior distribution we derive has a long tail toward larger (including even positive) $\Gamma_{\rm dark}$ values. The globally averaged Binney $\bar{\beta}_{\rm B}$ is well below the parameter $\beta_{\rm J}$, defined by Equation (10) for all inclinations. This is because $\beta_{\rm B}$ depends on $\langle v_\phi^2 \rangle$, while $\beta_{\rm J}$ does not. Axisymmetric models generally have $\langle v_\phi^2 \rangle$ increasing from the symmetry axis toward the equatorial plane (see Appendix B.4). So while $\beta_{\rm B} = \beta_{\rm J}$ on the symmetry axis, instead $\beta_{\rm B} \leqslant \beta_{\rm J}$ in the equatorial plane. While Table 3 shows that our best-fit models have $\beta_{\rm J}$ positive and increasing with decreasing inclination, the inferred value of $\bar{\beta}_{\rm B}$ depends less on inclination to within the statistical uncertainties. The overall anisotropy marginalized over inclinations is $\bar{\beta}_{\rm B} = -0.20^{+0.28}_{-0.53}$. Hence, our best-fit models are radially anisotropic on the symmetry axis and tangentially anisotropic in the equatorial plane. When integrated over the entire meridional (R, z) plane, they are tangentially anisotropic but still statistically consistent with isotropy.

Similarly, we plot in Figure 14 an equivalent case for the parameters $\Theta = \{\gamma_{\rm dark}, \log r_{\rm dark}\}$, with $r_{\rm dark}$ in pc. As expected from our conversions between $\Gamma_{\rm dark}$ and $\gamma_{\rm dark}$ in Section 3.3, one has a remarkable agreement between the peak of $\gamma_{\rm dark}$'s pdf and a classic ΛCDM slope. Besides, our uncertainties on $\gamma_{\rm dark}$ are consistent with what is expected from PM data sets having a similar number of stars to ours, as argued in Guerra et al. (2023). More importantly, this figure allows us to probe the core radius that our 3D data are able to constrain: while negative asymptotic slopes agree with a large set of DM scale radii, positive slopes require that the respective core (or even a drop in the density) be limited within $\lesssim$1 kpc. Indeed, upon analysis of our MCMC chains, cores larger than 487, 717, and 942 pc are ruled out at 1σ, 2σ, and 3σ confidence,[44] respectively. For reference, the scale

[43] We assigned the structural parameters provided by Odenkirchen et al. (2001) for this specific parameterization.

[44] These numbers are derived upon selecting the elements of the corner plot in Figure 14 that correspond to $\gamma_{\rm dark} \geqslant 0$ and retrieving the values that encompasses 68%, 95%, and 99.7% of the respective $\log r_{\rm dark}$ distribution.

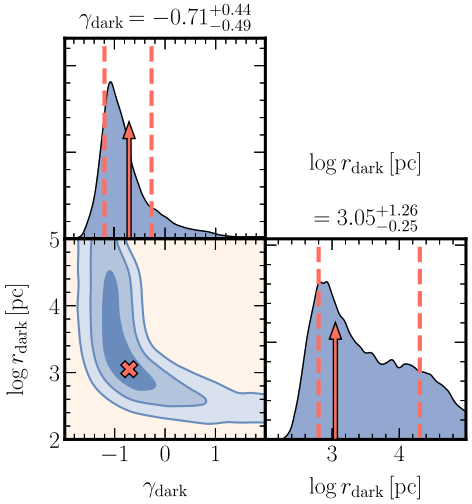

Figure 14. Asymptotic DM slope and scale radius: final posterior probability distributions of Draco's asymptotic DM density slope, $\gamma_{\rm dark}$, and the logarithm of the DM scale radius, $\log r_{\rm dark}$, with $r_{\rm dark}$ in pc. The plotting details are similar to the ones in Figure 13. The results rule out a core larger than 942 pc at 3σ confidence.

radius predicted by ΛCDM, previously mentioned in Section 3.3, equals 1.06 kpc.

The circular velocity at our outermost data point in our best-fit models is $v_{\rm circ}^{R_{\rm max}} = 24.19^{+6.31}_{-2.97}$ km s^{-1}. Similarly, the maximum value of the circular velocity is $v_{\rm circ}^{\rm max}(r) = 27.78^{+10.85}_{-5.97}$ km s^{-1}. This $v_{\rm circ}^{\rm max}(r)$ measurement is generally higher than most previous calculations, namely Strigari et al. (2007; 15–35 km s^{-1}), Martinez (2015; $18.2^{+3.2}_{-1.6}$ km s^{-1}), and Massari et al. (2020; 10.2–17.0 km s^{-1}). Correspondingly, the dark mass in our models is higher as well. It is difficult to precisely determine where such differences could come from, but one could speculate that this relates to different completeness of the respective data sets used in each work. Indeed, the circular velocity values measured by the likewise axisymmetric modeling from Hayashi et al. (2020, Figure 9) over similar radial ranges lie closer to ours (i.e., ∼25–30 km s^{-1}).

From Table 3, one sees that higher heliocentric distances of Draco are usually related to lower inclinations and more cored models, and vice versa. Our estimate of Draco's distance, $D = 75.37^{+4.73}_{-4.00}$ kpc, provides the first dynamical distance for this dwarf. Comparatively, Bonanos et al. (2004) measured 75.8 ± 5.44 kpc using a set of 146 RR Lyrae stars, Aparicio et al. (2001) found 80 ± 7 kpc from analyses of the magnitude of the horizontal branch at the RR Lyrae instability strip, and Muraveva et al. (2020) reported 80.5 ± 2.6 kpc when using 285 RR Lyrae stars. Our measurement is thus comparable to and competitive with other literature results based on stellar population methods. Thus, high-quality astrometric data also provide a valuable validation of standard distance determination techniques.

4.2.4. Spherical versus Axisymmetric Models

Until recently, there were no PM dispersion profiles available for internal mass modeling of dSphs. Hence, methods employed to analyze LOS velocities had to make substantial assumptions to remove degeneracies in the data. Among other things, models usually assumed spherical geometry (e.g., Wilkinson et al. 2002; Read et al. 2018; Massari et al. 2020), with the important exception of Hayashi et al. (2020). The velocity moments that are derived from the Jeans equations then depend only on the projected radius to the system's center. Instead, observed quantities in axisymmetric models depend on the position angle on the sky. We have found that for the new PM data set presented here axisymmetric models yield substantially different results from spherical models. This is true especially for the quantities most of interest, namely the DM cusp slope and the velocity anisotropy. Our axisymmetric models imply lower anisotropy $\bar{\beta}_{\rm B}$ and higher cusp slope $\Gamma_{\rm dark}$. Hence, it is critically important to construct axisymmetric models that properly take position-angle dependencies into account.

In Appendices B.3 and B.4, we again use the scale-free dynamical models of the type discussed in de Bruijne et al. (1996) to explain this result. Figure 16 in Appendix B shows that there is a tight monotonic relation between the PM anisotropy $\sigma_{\rm POSt}/\sigma_{\rm POSr}$ integrated over the sky and the globally integrated intrinsic $\bar{\beta}_{\rm B}$. This relation is very similar for spherical and axisymmetric models. However, for Draco we have not measured $\sigma_{\rm POSt}/\sigma_{\rm POSr}$ over the entire projected image of the galaxy, but only for two fields along the major axis (see Figure 6). Figure 17 in Appendix B shows that in axisymmetric geometry $\sigma_{\rm POSt}/\sigma_{\rm POSr}$ is not constant with position angle on the sky. Instead, it is much lower on the major axis than on the minor axis. Hence, a spherical model that assumes that $\sigma_{\rm POSt}/\sigma_{\rm POSr}$ is the same everywhere as measured on the major axis will overestimate the radial anisotropy $\beta_{\rm B}$. Hence, while our axisymmetric models imply that $\bar{\beta}_{\rm B} = -0.20^{+0.28}_{-0.53}$, our best-fit canonical spherical model (model 3 in Table 2) has instead $\beta_{\rm B} = 0.39^{+0.13}_{-0.14}$. As mentioned in Section 4.1.4, the latter is consistent with previous studies that assumed sphericity (e.g., Massari et al. 2020; del Pino et al. 2022). The higher $\beta_{\rm B}$ in spherical models translates to a shallower DM slope, given the correlation in Figure 13.

Systematic biases of this type are lessened, but not necessarily erased, when using more spatially complete data sets. This is something to keep in mind as we enter a new era of galactic PMs: although spherical models are less costly and thus helpful to understand general choices of parameterization and priors, robust results and respective conclusions can only be obtained when considering more complex models that take into account the real shape of the galaxy (this has also been argued previously by Genina et al. 2018).

Even in our case, there are still degeneracies that are not fully taken into account. Specifically, as highlighted in Cappellari (2020), the deprojection of the surface brightness to obtain the intrinsic luminosity density is not unique unless the axisymmetric galaxy is seen edge-on (Rybicki 1987; Kochanek & Rybicki 1996). The degeneracy increases considerably when the galaxy is seen at low inclinations (Romanowsky & Kochanek 1997; van den Bosch 1997; Magorrian 1999). Our results for $\Gamma_{\rm dark}$ do not depend strongly on the assumed inclination (see Table 3). But it should be kept in mind that at low inclination other deprojections of the luminous density may be possible that differ from the Plummer models assumed here.

5. Robustness of Modeling Results

5.1. The Effect of Binaries

As explained in Section 2.2.3, recent works have proposed that the LOS velocity dispersion from dwarf galaxies could be

inflated owing to the presence of unresolved binaries from single-epoch exposures. In particular, such inflation could translate to mass overestimation and could eventually bias our DM measurements. Beyond the tests performed in Section 2.2.3, which have shown that the LOS velocity dispersion remains consistent when using data with different numbers of epochs, another independent test is to compare mass modeling either with LOS+PM data (case (i)) or with PM data alone (case (ii)). Due to the lack of high-quality PM data, such a test has never been performed before. Our new state-of-the-art PM catalog can thus shed light on this question.

Model 9 in Table 2 depicts the results from spherical Jeans modeling with PM data alone, but with the same priors as in our preferred model 3 (LOS+PM). Despite the slight increase of Poisson uncertainties, due to the fact that the PM subset alone has about half as many stars as the LOS+PM subset, the overall predictions of model 9 agree within 1σ with all the predictions from model 3. Therefore, we conclude from both the tests performed in the current section and those from Section 2.2.3 that our dynamical modeling results for Draco do not have significant biases due to the presence of unresolved binaries. Although we can only constrain this for the specific galaxy analyzed here, this is an important result that sets the stage for future interpretations on the impact of unresolved binaries on the velocity dispersion profiles of other dwarf galaxies.

5.2. The Effect of Tides

To derive mass estimates and density profile shapes, our analyses assume that Draco is in dynamical equilibrium. However, a number of recent studies have suggested that the excessive mass-to-light ratios measured in dSphs could be due to out-of-equilibrium dynamics, which in turn inflate the velocity dispersion (e.g., Klessen & Kroupa 1998; Hammer et al. 2018). Thus, it is of interest to gauge the possibility that our conclusions could be biased by MW tidal effects.

On this subject, the past literature strongly supports that Draco is a galaxy with no clear signs of tidal disruption: Both Odenkirchen et al. (2001) and Ségall et al. (2007) have found no evident stellar streams, asymmetric disturbances, or density breaks that are characteristic of a tidally perturbed system. Orbital analyses from Sohn et al. (2017) support that the last pericenter passage of Draco, when accounting for a massive Large Magellanic Cloud (see their Table 5), happens $\sim$4 Gyr ago, at an average closest MW distance of $\sim$100 kpc, while Pace et al. (2022) report a pericenter of $r_{\rm peri} = 58.0^{+11.4}_{-9.5}$ kpc and Battaglia et al. (2022) report $r_{\rm peri} = 51.7^{+4.1}_{-6.1}$ kpc ($37.6^{+4.2}_{-4.4}$ kpc) for a lighter (heavier) MW, mostly from Gaia DR3 data.

For the lower pericenter distances found by Battaglia et al. (2022) and Pace et al. (2022), one could envision the gas expansion scenario proposed by Hammer et al. (2024) and Wang et al. (2024), which could produce higher velocity dispersion profiles than the stellar component alone (although still with typically $\lesssim$half of the values observed in Figure 11). However, Sohn et al. (2017) also provide arguments (see their Figure 4) against a mean radial expansion of the stellar component. Hence, there is no indication that strong tidal effects could be biasing our results, thus making Draco an ideal galaxy for our equilibrium-based dynamical modeling.

5.3. Error Threshold

As discussed in Section 2.3.5, dynamical modeling should be careful when including data points with PM uncertainties higher than the system's intrinsic velocity dispersion. The larger the PM uncertainties, the more important it is that they are known very accurately. This is difficult to guarantee, since not all sources of observational uncertainty are always known or easily quantified. If the PM uncertainties are underestimated, then the galaxy PM dispersion will be overestimated, which biases the dynamical modeling results. Hence, in our analysis we have only included stars for which the observational PM uncertainty is smaller than 0.024 mas yr^{-1}, similar to the intrinsic PM dispersion of the galaxy (Section 2.3.5). In addition, previous works with globular cluster PM data have often considered only low PM error stars when deriving PM dispersion profiles from Gaia (Baumgardt et al. 2019; Vitral et al. 2022) and HST (Bellini et al. 2014; Watkins et al. 2015a).

To further test our choice of PM error threshold, we also performed, for comparison, a MAMPOSST-PM analysis with only the subset of stars meeting a lower threshold of 0.022 mas yr^{-1}. We display this run as model 10 in Table 2, which should be compared to its counterpart with the higher error threshold, model 3. Model 10 has 239 stars with measured PMs, compared to 364 in model 3. This implies a decrease of 35% in statistical completeness and results in somewhat larger uncertainties for all inferred model parameters. Within 1σ all parameters agree between these two models. This further supports that our results are robust to our choice of PM error threshold and that there is no indication of biases due to underestimated errors in our standard data set.

5.4. Higher-order Velocity Moments

The shape of the LOS velocity distribution can in principle be used to obtain a constraint on the velocity anisotropy that is independent of our Jeans results (van der Marel & Franx 1993). However, detailed modeling of this shape requires more complicated modeling techniques than those presented here (e.g., Chanamé et al. 2008), which is outside the scope of the present paper. Nonetheless, we describe in Appendix B.5 that approximate modeling is again possible with the scale-free dynamical models of the type discussed in de Bruijne et al. (1996). Figure 18 in Appendix B shows that plausible axisymmetric models exist that both (a) have values of $\bar{\beta}_{\rm B}$ consistent with those derived from our Jeans models and (b) predict a kurtosis for the LOS velocity distribution that is consistent with the observed value. Hence, we would not expect that future detailed modeling of higher-order moments would change our conclusions about Draco's velocity anisotropy and hence its mass distribution. However, per Figure 18, it might be helpful to constrain Draco's viewing inclination, as well as details of its phase-space DF.

6. Cosmological Implications and Future Work

Our newly measured asymptotic DM slope for Draco is $\gamma_{\rm dark} = -0.71^{+0.44}_{-0.49}$, consistent with the behavior expected in ΛCDM (Navarro et al. 1997), especially when comparing the most likely results from our MCMC chains in Figure 14 (in other words, the peak of the respective posterior pdf). Instead, various early studies of dSph galaxies found that observations favored shallow inner DM density profile slopes, consistent with a constant-density "core" at the center (e.g., Battaglia et al. 2008;

Walker & Peñarrubia 2011; Amorisco & Evans 2012; Brownsberger & Randall 2021),[45] and inconsistent with predictions of standard ΛCDM cosmology, from DM-only simulations. Hence, if future PM studies of other dSph galaxies were to support our findings, then this supports the standard cosmological hypothesis that the DM in the Universe behaves as a cold particle. The cold aspect of the DM particles would cause them to clump toward the deeper regions of the potential well and form diverging cusps such as the one we measure.

Early suggestions of DM cores in dwarf galaxies inspired some studies to propose fundamental changes in the nature of DM, such as WDM, e.g., sterile neutrinos and gravitinos, which predict lower central DM densities and cored profiles (Dalcanton & Hogan 2001), or SIDM, for which DM particles in the central region thermalize via collisions and thereby form a cored profile (e.g., Sameie et al. 2020). The fact that our study can rule out cored profiles larger than 942 pc at a 3σ level thus imposes useful constraints on the SIDM cross section.[46]

Independent evidence for the presence of DM cores in gas-rich star-forming dwarf galaxies exists on the basis of H I rotation curve modeling (e.g., Flores & Primack 1994; Moore 1994; Burkert 1995). A likely explanation for this in the context of standard ΛCDM cosmology is the impact of baryons on the DM density profile. This may transform DM cusps into cores by transferring energy and mass to the outer parts of the halos, e.g., via supernova feedback (Read & Gilmore 2005; Pontzen & Governato 2012; Brooks & Zolotov 2014), or star formation events (Read et al. 2018). Our findings in the present paper suggest that these processes have not been important for Draco. However, that is not unexpected. Draco is below the limit where stellar feedback as implemented in current cosmological simulations should still produce a small core (Fitts et al. 2017). Moreover, its star formation shut down long ago (∼10 Gyr; Aparicio et al. 2001). Hence, our work does not provide insight into how the DM profiles of higher-mass or star-forming galaxies may be impacted by baryonic physics.

To further probe how the variety of proposed theoretical mechanisms to form a core compare to observations, it will be essential to measure and model the 3D dynamics of dwarfs with different characteristics than Draco, for example, dSphs having more troubled dynamical pasts, those having suffered close encounters with other satellites or the MW itself, or instead dSphs with star formation histories that shut down more recently. Efforts in these directions are already underway—we have acquired multiepoch data sets of the Sculptor dSph galaxy from the same HST programs described here (Sohn et al. 2021) and are in the process of also acquiring analogous HST data for the Ursa Minor dSph galaxy (Vitral et al. 2023b). Moreover, we have ongoing programs on JWST to further extend our time baselines for both Draco and Sculptor (van der Marel et al. 2023). The Nancy Grace Roman Space Telescope will provide the opportunity to further extend such studies over large fields of view (Han et al. 2023). This will help lift various degeneracies that have so far complicated a full exploration of existing discrepancies between observations and theoretical predictions for dwarf galaxies.

[45] Notice that such analyses considered the Sculptor and Fornax dwarfs, which have a higher stellar content, from whence it is not completely excluded that a cuspy DM density profile could have been lowered by baryonic effects.

[46] Specific constraints on how the core radius relates to the DM cross section and particle mass can be found in Read et al. (2018) and Macciò et al. (2013; Macciò et al. 2012's erratum), respectively.

7. Conclusions

For many decades, 3D velocity data sets of the internal kinematics of dwarf galaxies were only conceivable through numerical simulations. Thanks to HST's record time in operation and its exquisite astrometric capabilities, we can now study other galaxies using POS velocities that are not simulated but observed. With four epochs of HST observations of the Draco dSph spanning an 18 yr temporal baseline, we measure precise PMs for hundreds of stars, with uncertainties below the intrinsic velocity dispersion of the galaxy. This provides the most precise PM data set of Draco to date. We make this data set publicly available as online material,[47] so that it can also be used for other studies than those described here.

By combining the PMs with existing LOS velocities, we derive for the first time radially resolved 3D velocity dispersion profiles for any dwarf galaxy. With suitable modeling, these directly constrain the intrinsic velocity anisotropy of the galaxy and resolve the mass–anisotropy degeneracy that often plagues dynamical modeling.

To fit the measurements and infer the radial mass profile, we solve the Jeans equations in both spherical and axisymmetric geometries. The latter provides the first axisymmetric modeling of any pressure-supported external galaxy that is observationally constrained by all three orthogonal components of the stellar velocity field.

The viewing inclination of the galaxy is not constrained by the data. Hence, we marginalize our modeling results over all possible inclinations, informed by the overall distribution of projected galaxy shapes for elliptical galaxies in the nearby Universe. None of our conclusions depend sensitively on the actual inclination.

Below, we summarize our main findings for the Draco dSph galaxy, as well as the cosmological implications of our study.

1. We provide new estimates of the galaxy center, based on elliptical Bayesian fits to the Gaia EDR3 stellar counts (see Table 1).
2. We determine the LOS rotation curve of the galaxy (see Figure 4) and find that it has measurable rotation. Over the radial region covered by the LOS data set, $\langle v/\sigma \rangle = 0.22 \pm 0.09$.
3. We show that the impact of unresolved binaries on the LOS data is negligible and does not significantly alter the dynamical modeling results.
4. For the PMs in our HST fields along the projected major axis, we measure averaged observed velocity dispersion ratios of $\langle \sigma_{\rm POSt} \rangle / \langle \sigma_{\rm POSr} \rangle = 0.80 \pm 0.08$ and $\langle \sigma_{\rm LOS} \rangle / \langle \sigma_{\rm POS} \rangle = 1.08 \pm 0.09$, where $\sigma_{\rm POS}$ represents an average over both PM directions. The first ratio is independent of galaxy distance, while the second is inversely proportional to it.[48] The tight observational constraints on ratios like these enable dynamical models to strongly constrain the structure of Draco.
5. Axisymmetric models imply that the velocity dispersion tensor of the galaxy is radially anisotropic along the symmetry axis, with $\beta_{\rm J} = 0.56^{+0.25}_{-0.42}$. This is similar to the anisotropy everywhere in our best-fit spherical model ($\beta_{\rm B} = 0.39^{+0.13}_{-0.14}$). However, the best-fit axisymmetric

[47] The data set is available at doi:10.5281/zenodo.11111113.

[48] We assume $D = 75.37^{+4.73}_{-4.00}$ kpc.

models are tangentially anisotropic in the equatorial plane, as required to maintain hydrostatic equilibrium in an oblate system. The globally averaged anisotropy is $\bar{\beta}_{\rm B} = -0.20^{+0.28}_{-0.53}$, so Draco is tangentially anisotropic but still statistically consistent with isotropy.

6. Construction of axisymmetric models is essential for flattened galaxies. This is particularly important for PM data sets such as those presented here, which do not cover all position angles. Spherical models then yield biased estimates for both the velocity anisotropy and the inferred cusp slope.
7. We infer a DM density slope averaged over the spatial range for which we have PM measurements, $\Gamma_{\rm dark}$, of $-0.83^{+0.32}_{-0.37}$ and an asymptotic DM density slope of $\gamma_{\rm dark} = -0.71^{+0.44}_{-0.49}$. Cores larger than 487, 717, and 942 pc are ruled out at 1σ, 2σ, and 3σ confidence, respectively. The data do not have the constraining power to distinguish between different plausible parameterizations for the cusped DM density profile (see Table 2).
8. The measured slope is in good agreement with ΛCDM predictions,[49] given that our measurements fall well within the break radius of the DM density profile predicted by cosmological simulations. Our best likelihood results corroborate the idea that DM is formed by some sort of cold particle. An asymptotic core is marginally inconsistent with the data at 89.5% confidence, when marginalized over all other quantities, arguing against modified DM scenarios such as warm DM or SIDM. Nonetheless, a small asymptotic core cannot be effectively ruled out.
9. The measured cusp slope provides no evidence that it has been lowered through baryonic feedback processes, although this cannot be ruled out either. However, this is not unexpected given Draco's low mass and ancient star formation history. This provides no insight into how the DM profiles of higher-mass or star-forming galaxies may be impacted by baryonic physics.
10. We measure Draco's DM halo to have a mass of $1.20^{+0.79}_{-0.28} \times 10^8\,M_\odot$ at the outermost data point ($R_{\rm max} = 900$ pc), where the circular velocity reaches $v_{\rm circ} = 24.19^{+6.31}_{-2.97}$ km s^{-1}. The maximum circular velocity in our best-fit models is $v^{\rm max}_{\rm circ}(r) = 27.78^{+10.85}_{-5.97}$ km s^{-1}.
11. We infer a dynamical distance of $75.37^{+4.73}_{-4.00}$ kpc. This is consistent with estimates obtained in the literature using other methods. Our ∼6% distance uncertainty is competitive with the uncertainties inherent to galaxy distances based on stellar evolution and the cosmic distance ladder. This is similar to the situation for globular clusters (Watkins et al. 2015b).

We have obtained here one of the most reliable constraints to date on the DM density profiles of dwarf galaxies. The results lessen the tension around the "cusp–core" problem and give further credence to standard ΛCDM cosmology. Our study is just a first step into the realm of 3D axisymmetric dynamics of dSphs. The methods we used can be largely generalized to other systems, and further studies on both Draco and other galaxies are already underway. Hence, many substantial advances are likely to be made in this area in the coming years.

[49] For reference, a traditional NFW profile has an asymptotic cusp slope of −1.

Acknowledgments

Support for this work was provided by NASA through grants for programs GO-12966 and GO-16737 from the Space Telescope Science Institute (STScI), which is operated by the Association of Universities for Research in Astronomy (AURA), Inc., under NASA contract NAS5-26555, and through funding to the JWST Telescope Scientist Team (PI: M. Mountain) through grant 80NSSC20K0586. M.G.W. acknowledges support from National Science Foundation (NSF) grants AST-1909584 and AST-2206046. We thank Jorge Peñarrubia for thoughtful suggestions concerning our models and Justin Read for sharing his PYTHON algorithm to compute $M_\star$–M_{200} relations as published in Read et al. (2017).

The Digitized Sky Surveys were produced at the Space Telescope Science Institute under US government grant NAG W-2166. The images of these surveys are based on photographic data obtained using the Oschin Schmidt Telescope on Palomar Mountain and the UK Schmidt Telescope. The plates were processed into the present compressed digital form with the permission of these institutions.

This work has made use of data from the European Space Agency (ESA) mission Gaia (https://www.cosmos.esa.int/gaia), processed by the Gaia Data Processing and Analysis Consortium (DPAC, https://www.cosmos.esa.int/web/gaia/dpac/consortium). Funding for the DPAC has been provided by national institutions, in particular the institutions participating in the Gaia Multilateral Agreement.

This project is part of the High-resolution Space Telescope PROper MOtion (HSTPROMO) Collaboration (https://www.stsci.edu/~marel/hstpromo.html), a set of projects aimed at improving our dynamical understanding of stars, clusters, and galaxies in the nearby Universe through measurement and interpretation of PMs from HST, Gaia, and other space observatories. We thank the collaboration members for the sharing of their ideas and software.

Data Availability

The data that we used to construct PMs were obtained from the Mikulski Archive for Space Telescopes at the Space Telescope Science Institute. They relate to HST observations from programs GO-10229, GO-10812, GO-12966, and GO-16737, which can be accessed via doi:10.17909/f6wq-1098. We also share our final data products on Zenodo at doi:10.5281/zenodo.11111113, which consist of PMs, 3D velocity dispersion profiles, rotation profile, and the catalog ID of the stars from Walker et al. (2023) used in this work.

Facilities: HST, Gaia, MMT

Software: PYTHON (Van Rossum & Drake 2009), BALROGO (Vitral 2021), JAMPY (Cappellari 2020), MAMPOSST-PM (G. A. Mamon & E. Vitral, in preparation), EMCEE (Foreman-Mackey et al. 2013), SCIPY (Virtanen et al. 2020), NUMPY (van der Walt et al. 2011), MATPLOTLIB (Hunter 2007), APLPY (Robitaille & Bressert 2012), SCALEFREE (de Bruijne et al. 1996), GEOGEBRA (Hohenwarter 2002), Mathematica version 12.0 (Champaign, IL: Wolfram Research, Inc.).

Appendix A
Density Profiles

We present here the analytical forms of the luminous and DM density profiles considered throughout our work. Those

are, respectively, the Plummer (1911) density profile,

$$\rho_{\rm PLU}(r) = \frac{3M_\infty}{4\pi a^3}\left(1 + \frac{r^2}{a^2}\right)^{-\frac{5}{2}}, \tag{A1}$$

where M_∞ is the total tracer mass at infinity and a is a scale radius (the same one depicted in Table 2). In projected geometry, we have the Plummer axisymmetric surface density profile, which we use in Section 2.1.1:

$$\Sigma_{\rm PLU,axi}(R, \xi) = \frac{M_\infty}{\pi a^2(1-\epsilon)} \times \left[1 + \frac{R^2}{a^2}\left(\cos^2\xi + \frac{\sin^2\xi}{(1-\epsilon)^2}\right)\right]^{-2}. \tag{A2}$$

where ξ is the projected angle on the sky ($\xi = 0$ on the projected major axis) and ϵ is defined as $\epsilon \equiv 1 - b/a$ (b and a are minor and major axes, respectively).

Back to the spherical profiles used for Jeans modeling, there is the generalized Plummer model,

$$\rho_{\rm GPLU}(r) = (3+\gamma)\frac{M_\infty}{4\pi a^3}\left(\frac{r}{a}\right)^{\gamma}\left(1 + \frac{r^2}{a^2}\right)^{-\frac{(\gamma+5)}{2}}, \tag{A3}$$

where the new parameter $\gamma < 3$ measures the inner density slope. The Kazantzidis et al. (2004b) profile is

$$\rho_{\rm GKAZ}(r) = \frac{1}{\Gamma(3+\gamma)}\frac{M_\infty}{4\pi a^3}\left(\frac{r}{a}\right)^{\gamma}\exp\left(-\frac{r}{a}\right). \tag{A4}$$

The generalized (Navarro et al. 1997; NFW) profile is

$$\rho_{\rm GNFW}(r) = \frac{M_{-2}/(4\pi a^3)}{{}_2F_1(1,\,1;\,4+\gamma;\,-2-\gamma)} \times \left(\frac{3+\gamma}{2+\gamma}\right)^{3+\gamma}\left(\frac{r}{a}\right)^{\gamma}\left(1 + \frac{r}{a}\right)^{-(3+\gamma)}, \tag{A5}$$

where M_{-2} stands for the cumulative mass within the radius where the condition $d\log\rho/d\log r = -2$ is satisfied. The Einasto (1965) profile is

$$\rho_{\rm EIN}(r) = \frac{1}{n\,\Gamma(3n)}\frac{M_\infty}{4\pi a^3}\exp\left[-\left(\frac{r}{a}\right)^{\frac{1}{n}}\right], \tag{A6}$$

where n is the Einasto index. We also use the Sérsic (1963, 1968) model,

$$\Sigma_{\rm SER}(R) = \frac{M_\infty\, b_n^{2n}}{2\pi R_e^2\, n\,\Gamma(2n)}\exp\left[-b_n\left(\frac{R}{R_e}\right)^{\frac{1}{n}}\right], \tag{A7}$$

where R is the projected distance to the source center, R_e is the effective radius containing half of the projected luminosity, and n is the Sérsic index. The term b_n is a function of n, obtained by solving the equation

$$\Gamma(2n)/2 = \gamma(2n, b_n), \tag{A8}$$

where $\gamma(a, x) = \int_0^x t^{a-1}e^{-t}dt$ is the lower incomplete gamma function. This model does not have an exact analytical deprojection into volume density (see Graham & Driver 2005 for a review). Instead, we use a combination of the most precise analytical approximations for different domains of Sérsic indices and radii. This combination is described thoroughly in Appendix A from Vitral & Mamon (2021) and essentially uses the methods from Lima Neto et al. (1999) plus Simonneau & Prada (2004) and Vitral & Mamon (2020).

Appendix B
Scale-free Reference Models

B.1. Model Parameters and Distribution Functions

The Jeans models in Section 3 are largely based on numerical solution of ordinary differential equations. This means that it is somewhat complicated to quickly or approximately explore how the intrinsic properties or predicted observables of the models depend on its key parameters. For the latter purpose we have found it more convenient to resort to a simpler class of models, namely the scale-free models introduced by de Bruijne et al. (1996).

These models consider the case of an axisymmetric luminous density with axial ratio q and radial dependence $\rho_\star \propto r^{-\kappa}$,[50] embedded in a spherical gravitational potential with circular velocity $v_{\rm circ} \propto r^{-\delta/2}$. The case of a logarithmic potential corresponds to $\delta = 0$, while a Kepler potential corresponds to $\delta = 1$.

For these potentials, there are different classes of phase-space DFs that can yield hydrostatic equilibrium in the same geometry and with the same globally averaged anisotropy, but with a different variation of the anisotropy in the meridional (R, z) plane. Specifically, de Bruijne et al. (1996) defined distinct Case I and Case II DFs. In the Case I DFs (for which $f(E, L_z)$ models form a subset) the quantity $\beta_{\rm B}$, as defined in Equation (5), is more tangential in the equatorial plane than on the symmetry axis. In the Case II DFs, instead, $\beta_{\rm B}$ is constant throughout the system. The intrinsic and projected quantities of interest for these DFs can be expressed semianalytically, so that they can be computed quickly.

A software package called SCALEFREE originally written by one of us (RvdM) for the de Bruijne et al. (1996) paper is available for this purpose. The scale-free models have the additional advantage that the full DF is known, so that higher-order moments can be calculated in addition to the second-order moments constrained by the Jeans equations (see Appendix B.5).

The scale-free models exactly reproduce the asymptotic large-radius limit of our Jeans models in Section 3, at intrinsic and projected radii well in excess of the scale radii $r_\star$ and $r_{\rm dark}$ of the luminous and DM (see, e.g., Table 2). But they also provide a reasonable approximation at the intermediate radii where we have PM measurements, if the power-law slopes κ and δ are set to reproduce the average slopes of the luminous density and circular velocity over those radii.

From exploration of the known properties of our best-fit Jeans models in Section 3, we found that the combination of $\kappa = 2$ and $\delta = 0$ (i.e., an axisymmetric isothermal luminous density in a spherical isothermal potential) is reasonable over the relevant radial range in Draco. We show below some predictions of such models as a function of the geometry (spherical, or axisymmetric of a given axial ratio and inclination), of the type of DF, and of the globally averaged $\bar{\beta}_{\rm B}$. The latter can be calculated numerically for given β_p,

[50] We relabel the parameter γ used in de Bruijne et al. (1996) here as κ, to avoid confusion with the central cusp slope of the DM density defined in Appendix A.

where β_p is a parameter that enters into the analytical DFs and controls the amount of anisotropy in the models.[51]

B.2. Evaluation of Proper Motions

The de Bruijne et al. (1996) paper addressed the evaluation of projected LOS velocities of the scale-free models, but not the calculation of PMs, which were not observationally accessible at the time. We therefore extended their formalism to POS velocities.[52]

Figure 15 shows the general geometry of the problem, with the same conventions as in Evans & de Zeeuw (1994). The Cartesian frame (x, y, z) is aligned with the axisymmetric luminous density, so that the z-axis is its symmetry axis. The (r, θ, ϕ) are spherical coordinates in this system. A second Cartesian frame (x', y', z') is introduced so that the (x', y') frame corresponds to the POS, with $x' \equiv y$ along the projected major axis of the luminous density and z' along the LOS direction. The viewing inclination i is the angle between the z- and z'-axes. The coordinates are related through

$$x' = y, \tag{B1a}$$

$$y' = -x\cos i + z\sin i, \tag{B1b}$$

$$z' = x\sin i + z\cos i. \tag{B1c}$$

The velocity along any coordinate direction in the (x', y', z') frame can be expressed as a combination of radial, tangential, and azimuthal components,

$$v_{\rm dir} = C_r v_r + C_\theta v_\theta + C_\phi v_\phi. \tag{B2}$$

For the LOS direction dir $= z'$, the C_i terms are given by Equation (49) in de Bruijne et al. (1996):

$$C_r = \sin\theta\sin i\cos\phi + \cos\theta\cos i, \tag{B3a}$$

$$C_\theta = \cos\theta\sin i\cos\phi - \sin\theta\cos i, \tag{B3b}$$

$$C_\phi = -\sin i\sin\phi. \tag{B3c}$$

The velocities in the radial and tangential directions on the POS can be written as

$$v_{\rm POSr} = \frac{x' v_{x'} + y' v_{y'}}{\sqrt{x'^2 + y'^2}}, \tag{B4a}$$

$$v_{\rm POSt} = \frac{y' v_{x'} - x' v_{y'}}{\sqrt{x'^2 + y'^2}}. \tag{B4b}$$

From these relations it follows[53] that the C_i terms for the POSr and POSt directions are, respectively,

1. $v_{\rm POSr}$:

$$C_r = \sqrt{\Delta^2}, \tag{B5a}$$

$$C_\theta = [\sin(2\theta)(\cos^2 i - \sin^2 i\cos^2\phi) - \cos(2\theta)\sin(2i)\cos\phi]/(2\sqrt{\Delta^2}), \tag{B5b}$$

$$C_\phi = \sin\phi\sin i[\sin\theta\sin i\cos\phi + \cos\theta\cos i]/\sqrt{\Delta^2}. \tag{B5c}$$

2. $v_{\rm POSt}$:

$$C_r = 0, \tag{B6a}$$

$$C_\theta = \sin i\sin\phi/\sqrt{\Delta^2}, \tag{B6b}$$

$$C_\phi = [\cos\theta\sin i\cos\phi - \sin\theta\cos i]/\sqrt{\Delta^2}. \tag{B6c}$$

Above, the variable Δ^2 is defined as

$$\Delta^2 \equiv \sin^2\theta\sin^2\phi + (\cos\theta\sin i - \sin\theta\cos i\cos\phi)^2. \tag{B7}$$

With these equations one can calculate the nth-order intrinsic and projected moments for dir = (LOS, POSr, POSt) starting from Equation (50) in de Bruijne et al. (1996),

$$\rho\langle v_{\rm dim}^n\rangle = \sum_{j=0}^{n}\sum_{k=0}^{n-j}\binom{n}{j}\binom{n-j}{k} \times C_r^j C_\theta^k C_\phi^{n-j-k}\rho\langle v_r^j v_\theta^k v_\phi^{n-j-k}\rangle, \tag{B8}$$

as described in their paper. All the quantities $\rho\langle v_r^j v_\theta^k v_\phi^{n-j-k}\rangle$ are uniquely determined by the DF and the model parameters.

B.3. Anisotropy Dependence of Observables

Leonard & Merritt (1989) and van der Marel & Anderson (2010) both showed that in spherical geometry there is a direct relation between the projected anisotropy $\sigma_{\rm POSt}/\sigma_{\rm POSr}$ and the intrinsic 3D velocity dispersion anisotropy $\sigma_{\rm tan}/\sigma_r$, where $\sigma_{\rm tan} \equiv [(\sigma_\theta^2 + \sigma_\phi^2)/2]^{1/2}$. The scale-free models discussed above allow us to demonstrate that a similar relation holds in axisymmetric geometry as well.

Figure 16 shows the relation for three different geometries, namely (a) spherical; (b) axisymmetric and edge-on, with the same axial ratio as observed in projection of the sky (listed in Table 1); and (c) axisymmetric and at viewing inclination $i = 57.^\circ1$, as in Section 4.2.2, which is close to the median

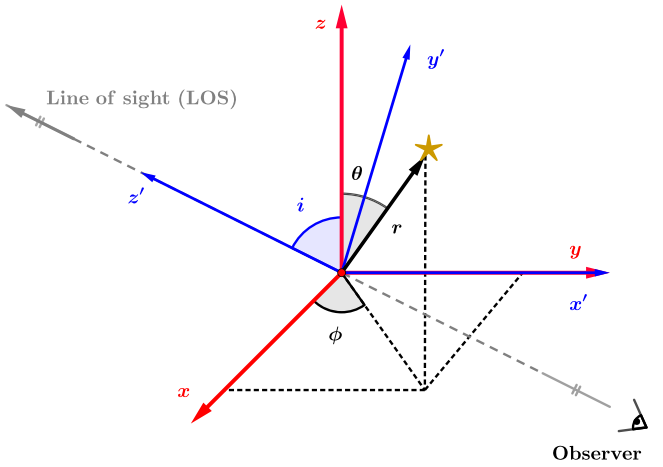

Figure 15. Geometry of the system: set of coordinate definitions. The astronomical source is inclined by an angle i, and the LOS direction is the same as the z'-direction depicted above.

[51] We relabel the parameter β in de Bruijne et al. (1996) here as β_p, to set it apart from the distinct quantities $\beta_{\rm B}$ and $\beta_{\rm J}$ already defined by Equations (5) and (10).

[52] The corresponding software can be found at https://gitlab.com/eduardo-vitral/scalefree.

[53] We computed the relations in Equations (B5a)–(B7) using the MATHEMATICA software.

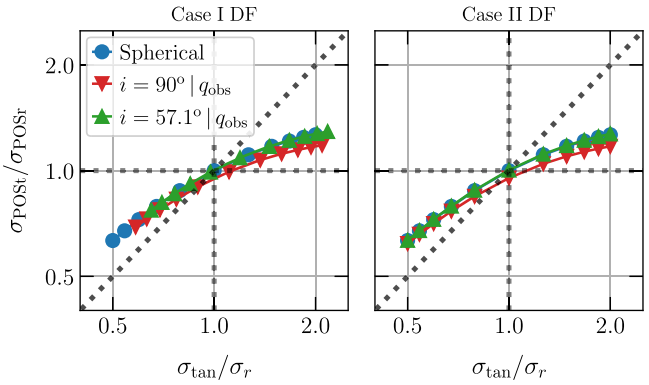

Figure 16. Anisotropy dependence of kinematics: projected anisotropy $\sigma_{POSt}/\sigma_{POSr}$ as a function of intrinsic 3D velocity dispersion anisotropy σ_{tan}/σ_r. Here σ is used as shorthand for $\langle v^2\rangle^{1/2}$. Mass-weighted averaged predictions over all position angles are shown for scale-free models with $\kappa = 2$ and $\delta = 0$ and for Case I (left) and Case II (right) DFs, respectively. Each panel shows predictions for three different geometries, namely (a) spherical, (b) axisymmetric and edge-on, and (c) axisymmetric and at viewing inclination $i = 57.^\circ1$. The axisymmetric models all have the same projected axial ratio as Draco. The projected anisotropy in the POS is tightly related to the intrinsic anisotropy.

expected inclination given Draco's projected axial ratio. The intrinsic axial ratio of the latter model is $q = 0.608$. The figure shows mass-weighted averages over all position angles (either on the POS or in the meridional (R, z) plane, respectively), given that in axisymmetric models all quantities generally vary as a function of position angle (see Appendix B.4). The two different panels show the predictions for Case I and Case II DFs, respectively. In the figures, σ is used as shorthand for $\langle v^2\rangle^{1/2}$; these quantities are independent of whether or not the system has mean rotation, but the quantities are equal to the dispersion only if there is no rotation.

In all cases shown, the projected $\sigma_{POSt}/\sigma_{POSr}$ is simply a diluted measure (i.e., brought closer to unity owing to projection effects) of the intrinsic 3D ratio σ_{tan}/σ_r, with only subtle variations depending on the assumed geometry and the type of DF. The slope of the relation does vary with the radial slope κ of the luminous tracer density distribution (not shown here, but see van der Marel & Anderson 2010), but the relation remains monotonic and one-to-one. This explains why, with suitable modeling, our new PM measurements have the constraining power to determine the intrinsic velocity dispersion anisotropy of Draco, even when not using LOS data.

B.4. Position-angle Dependence

For this paper, we have obtained new HST PM data primarily along the projected major axis. Thus, while Figure 16 shows the quantity $\sigma_{POSt}/\sigma_{POSr}$ integrated over all position angles, it is important to assess also how the kinematics vary with position angle in the models. The three rows of Figure 17 show SCALEFREE model predictions for the same three geometries as in Figure 16. For each of the geometries we show models with two different types of DF (Case I and Case II), but always with $\bar{\beta}_B = -0.20$, consistent with the best fit from our axisymmetric Jeans models in Table 3. The left two panels show the variations of the internal kinematics in the meridional (R, z) plane, while the right two panels show the variations of the projected kinematics in the POS.

The main difference in the internal kinematics predictions for the axisymmetric models between the two types of DF is in the variation of σ_r with θ' ($\equiv 90^\circ - \theta$). For either DF, σ_ϕ increases from the symmetry axis to the major axis (as before, we use in this discussion σ as shorthand for $\langle v^2\rangle^{1/2}$). This is a direct consequence of the tensor virial theorem, which requires that overall the system has more dynamical pressure parallel to the equatorial plane than perpendicular to it, so as to support its flattened shape. The right two panels show that in projection σ_{POSr} decreases from the major to the minor axis, while σ_{POSt} increases from the major to the minor axis. These behaviors are nearly independent of inclination (at fixed projected axial ratio) and of the specific type of DF. These variations are ignored when a spherical model is constructed because the projected quantities are then independent of position angle. This can introduce important biases, as discussed in Section 4.2.4.

While Figure 17 pertains to a specific set of parameter combinations, we verified that different parameter combinations yield qualitatively similar conclusions. For example, when considering flatter models with lower axial ratio (either intrinsically or in projection), the variations from major to minor axis become more pronounced than in Figure 17 but otherwise remain qualitatively similar.

B.5. Higher-order Moments

Our analysis in this paper has been based on the second (and first) velocity moments that enter into the Jeans equations of hydrostatic equilibrium. However, for the observed LOS velocities we have enough measurements, with small enough error bars, to make it possible to also determine higher-order moments. Specifically, we calculated both the fourth-order Gauss-Hermite moment h_4 (as in van der Marel & Franx 1993) and the kurtosis $\overline{K_{LOS}}$ of the overall LOS velocity distribution. The kurtosis is defined here as

$$\overline{K_{LOS}} = [\langle v_{LOS}^4\rangle / \langle v_{LOS}^2\rangle^2] - 3, \tag{B9}$$

where, as before, $\langle v^k\rangle$ indicates a kth-order velocity moment. We use all position angles on the sky and center the LOS velocity distribution on the bulk systemic velocity listed in Table 1. Hence, odd moments are generally statistically consistent with zero. The observations then imply that $h_4 = 0.065 \pm 0.036$ and $\overline{K_{LOS}} = 0.21 \pm 0.25$. Both of these quantities are positive, indicating that the LOS velocity distribution is more centrally peaked than a Gaussian.

The second-order Jeans equations cannot be used to interpret these measurements. However, the scale-free DF models predict all higher-order moments $\langle v^k\rangle$ for any given set of model parameters. Reconstructing the full velocity distribution and Gauss–Hermite moments from the higher-order moments is not a straightforward inversion problem (de Bruijne et al. 1996), so we restrict data–model comparisons here to the kurtosis. Figure 18 shows the predicted kurtosis as a function of the globally averaged Binney $\bar{\beta}_B$, defined by Equation (5). As in Figure 16, we show predictions for different geometries, with Case I DFs in the left panel and Case II DFs in the right panel. We now also include a flatter axisymmetric model with true axial ratio $q = 0.45$, which at viewing inclination $48.^\circ3$ projects to the observed axial ratio for Draco. The horizontal orange-gray band shows the observed kurtosis with its uncertainty. For reference, the vertical green band shows the 68% confidence band on $\bar{\beta}_B$ for Draco from our axisymmetric Jeans models (bottom row of Table 3).

The predictions in the figure show that more radially anisotropic models predict more peaked LOS velocity

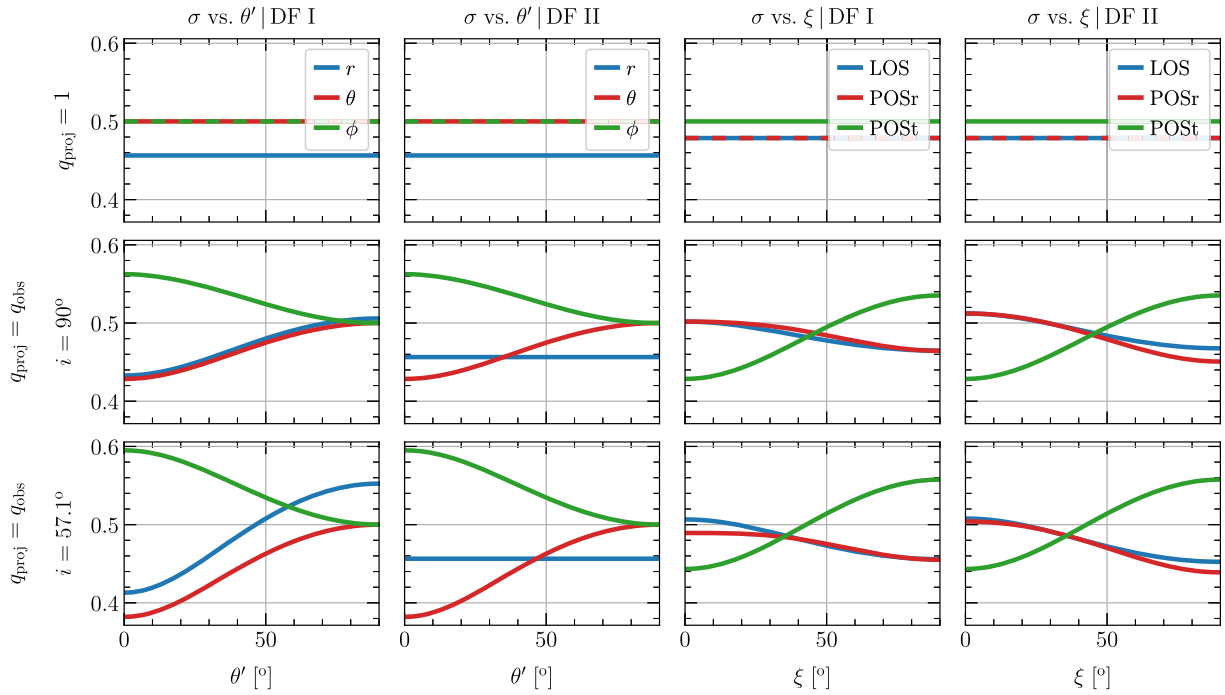

Figure 17. Position-angle dependence of kinematics. Rows show the predictions of scale-free models for the same three geometries as in Figure 16, in each case chosen to have globally averaged anisotropy $\bar{\beta}_B = -0.20$. The left two panels show the variations of σ_r, σ_θ, and σ_ϕ with angle θ' in the meridional (R, z) plane ($\theta' \equiv 90° - \theta = 0$ in the equatorial plane). The right two panels show the variations of the projected σ_{LOS}, σ_{POSr}, and σ_{POSt} with the angle ξ in the POS ($\xi = 0$ on the projected major axis). Here σ is used as shorthand for $\langle v^2 \rangle^{1/2}$. Velocities are expressed in the dimensionless units defined in Section 2.1 of de Bruijne et al. (1996). The first and third columns show Case I DFs, while the second and fourth columns show Case II DFs. In axisymmetric geometry there are important variations of the kinematical observables with position angle on the sky, which are ignored when observations are interpreted with spherical models. This can introduce biases in model results.

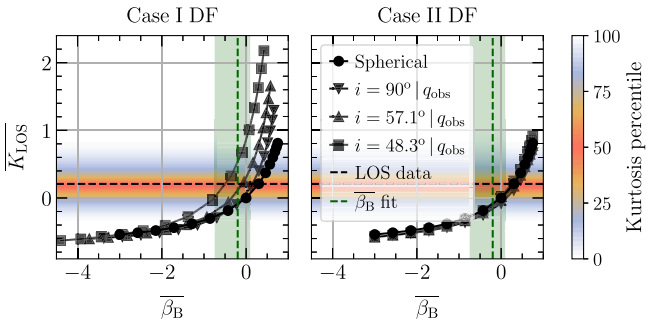

Figure 18. Kurtosis: kurtosis of the grand-total LOS velocity distribution as a function of the globally averaged velocity anisotropy $\overline{\beta_B}$. As in Figure 16, different curves show predictions of scale-free models for different geometries, for Case I (left panel) and Case II (right panel) DFs, respectively. We now also include a flatter axisymmetric model with true axial ratio $q = 0.45$, which at viewing inclination 48.°3 projects to the observed axial ratio for Draco. We show in dashed black the observed kurtosis along with the percentiles of its distribution, color coded as indicated on the right. The $\overline{\beta_B}$ inferred from our axisymmetric Jeans models, with its 68% confidence region, is depicted in green. Plausible models exist that match all constraints (i.e., have predictions within the intersection of the orange and green bands).

distributions (higher kurtosis), as has been previously established (e.g., van der Marel & Franx 1993). For the Case II DFs, the predictions are more or less independent of geometry. But for the Case I DFs, flatter axisymmetric models (seen at lower inclinations) predict higher kurtosis values. Hence, higher-order moments can help to determine both the viewing inclination and the full structure of the DF. More sophisticated dynamical modeling is required to fully exploit this information (e.g., Chanamé et al. 2008). Nonetheless, we note here that an axisymmetric Case I scale-free model at the median inclination of 57.°1 predicts the observed $\overline{K_{LOS}} = 0.21$ when $\overline{\beta_B} = -0.07$. The latter is fully consistent with the constraints that we have obtained from our axisymmetric Jeans models. Hence, there is no reason to expect that detailed modeling of higher-order moments would alter the conclusions about Draco's velocity anisotropy, and hence its mass distribution, that we have drawn in this paper.

ORCID iDs

Eduardo Vitral https://orcid.org/0000-0002-2732-9717
Roeland P. van der Marel https://orcid.org/0000-0001-7827-7825
Sangmo Tony Sohn https://orcid.org/0000-0001-8368-0221
Mattia Libralato https://orcid.org/0000-0001-9673-7397
Andrés del Pino https://orcid.org/0000-0003-4922-5131
Laura L. Watkins https://orcid.org/0000-0002-1343-134X
Andrea Bellini https://orcid.org/0000-0003-3858-637X
Matthew G. Walker https://orcid.org/0000-0003-2496-1925
Gurtina Besla https://orcid.org/0000-0003-0715-2173
Marcel S. Pawlowski https://orcid.org/0000-0002-9197-9300
Gary A. Mamon https://orcid.org/0000-0001-8956-5953

References

Akaike, H. 1998, in Selected Papers of Hirotugu Akaike, ed. E. Parzen, K. Tanabe, & G. Kitagawa (New York: Springer), 199
Akaike, H. 1983, 44th World Statistics Congress (The Hague: International Statistical Institute), 277
Amorisco, N. C., & Evans, N. W. 2012, MNRAS, 419, 184

Anderson, J. 2022, One-Pass HST Photometry with hst1pass, Instrument Science Report ACS 2022-02, STScI
Anderson, J., Bedin, L. R., Piotto, G., Yadav, R. S., & Bellini, A. 2006, A&A, 454, 1029
Aparicio, A., Carrera, R., & Martínez-Delgado, D. 2001, AJ, 122, 2524
Armandroff, T. E., Olszewski, E. W., & Pryor, C. 1995, AJ, 110, 2131
Bacon, R., Simien, F., & Monnet, G. 1983, A&A, 128, 405
Battaglia, G., Helmi, A., Tolstoy, E., et al. 2008, ApJL, 681, L13
Battaglia, G., Taibi, S., Thomas, G. F., & Fritz, T. K. 2022, A&A, 657, A54
Baumgardt, H., Hilker, M., Sollima, A., & Bellini, A. 2019, MNRAS, 482, 5138
Bellini, A., Anderson, J., van der Marel, R. P., et al. 2014, ApJ, 797, 115
Bellini, A., Libralato, M., Bedin, L. R., et al. 2018, ApJ, 853, 86
Bianchini, P., Norris, M. A., van de Ven, G., et al. 2016, ApJL, 820, L22
Binney, J. 1980, MNRAS, 190, 873
Binney, J., & Mamon, G. A. 1982, MNRAS, 200, 361
Binney, J., & Tremaine, S. 1987, Galactic Dynamics (Princeton, NJ: Princeton Univ. Press)
Bonanos, A. Z., Stanek, K. Z., Szentgyorgyi, A. H., Sasselov, D. D., & Bakos, G. Á 2004, AJ, 127, 861
Breddels, M. A., Helmi, A., van den Bosch, R. C. E., van de Ven, G., & Battaglia, G. 2013, MNRAS, 433, 3173
Brooks, A. M., & Zolotov, A. 2014, ApJ, 786, 87
Brownsberger, S. R., & Randall, L. 2021, MNRAS, 501, 2332
Burkert, A. 1995, ApJL, 447, L25
Burnham, K. P., & Anderson, D. R. 2002, Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach (2nd ed.; New York: Springer)
Cappellari, M. 2008, MNRAS, 390, 71
Cappellari, M. 2020, MNRAS, 494, 4819
Chanamé, J., Kleyna, J., & van der Marel, R. 2008, ApJ, 682, 841
Dalcanton, J. J., & Hogan, C. J. 2001, ApJ, 561, 35
de Bruijne, J. H. J., van der Marel, R. P., & de Zeeuw, P. T. 1996, MNRAS, 282, 909
del Pino, A., Libralato, M., van der Marel, R. P., et al. 2022, ApJ, 933, 76
Dutton, A. A., & Macciò, A. V. 2014, MNRAS, 441, 3359
Einasto, J. 1965, TrAlm, 51, 87
Evans, N. W., & de Zeeuw, P. T. 1994, MNRAS, 271, 202
Fardal, M. A., van der Marel, R., del Pino, A., & Sohn, S. T. 2021, AJ, 161, 58
Fitts, A., Boylan-Kolchin, M., Elbert, O. D., et al. 2017, MNRAS, 471, 3547
Flores, R. A., & Primack, J. R. 1994, ApJL, 427, L1
Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306
Genina, A., Benítez-Llambay, A., Frenk, C. S., et al. 2018, MNRAS, 474, 1398
Genina, A., Read, J. I., Frenk, C. S., et al. 2020, MNRAS, 498, 144
Gilmore, G., Randich, S., Worley, C. C., et al. 2022, A&A, 666, A120
Graham, A. W., & Driver, S. P. 2005, PASA, 22, 118
Guerra, J., Geha, M., & Strigari, L. E. 2023, ApJ, 943, 121
Hammer, F., Wang, J., Mamon, G. A., et al. 2024, MNRAS, 527, 2718
Hammer, F., Yang, Y., Arenou, F., et al. 2018, ApJ, 860, 76
Han, J. J., Dey, A., Price-Whelan, A. M., et al. 2023, arXiv:2306.11784
Hargreaves, J. C., Gilmore, G., Irwin, M. J., & Carter, D. 1996, MNRAS, 282, 305
Hattori, K., Valluri, M., & Vasiliev, E. 2021, MNRAS, 508, 5468
Hayashi, K., Chiba, M., & Ishiyama, T. 2020, ApJ, 904, 45
Hohenwarter, M. 2002, Master's thesis, Paris Lodron Univ., Salzburg
Hunter, J. D. 2007, CSE, 9, 90
Hurvich, C. M., & Tsai, C.-L. 1989, Biometrika, 76, 297
Jiao, Y., Hammer, F., Wang, H., et al. 2023, A&A, 678, A208
Jing, Y. P., & Suto, Y. 2000, ApJL, 529, L69
Kazantzidis, S., Kravtsov, A. V., Zentner, A. R., et al. 2004a, ApJL, 611, L73
Kazantzidis, S., Mayer, L., Mastropietro, C., et al. 2004b, ApJ, 608, 663
Kenney, J., & Keeping, E. 1951, Mathematics of Statistics Part 2 (Princeton, NJ: Van Nostrand)
Kenney, J., & Keeping, E. 1963, Mathematics of Statistics Part 1 (Princeton, NJ: Van Nostrand)
Klessen, R. S., & Kroupa, P. 1998, ApJ, 498, 143
Kleyna, J., Wilkinson, M. I., Evans, N. W., Gilmore, G., & Frayn, C. 2002, MNRAS, 330, 792
Kochanek, C. S., & Rybicki, G. B. 1996, MNRAS, 280, 1257
Kowalczyk, K., del Pino, A., Łokas, E. L., & Valluri, M. 2019, MNRAS, 482, 5241
Kozhurina-Platais, V., Borncamp, D., Anderson, J., Grogin, N., & Hack, M. 2015, ACS/WFC Revised Geometric Distortion for DrizzlePac, Instrument Science Report ACS/WFC 2015-06, STScI
Kroupa, P., Tout, C. A., & Gilmore, G. 1993, MNRAS, 262, 545
Lambas, D. G., Maddox, S. J., & Loveday, J. 1992, MNRAS, 258, 404
Lazar, A., & Bullock, J. S. 2020, MNRAS, 493, 5825
Leonard, P. J. T., & Merritt, D. 1989, ApJ, 339, 195
Lewis, A., & Bridle, S. 2002, PhRvD, 66, 103511
Li, T. S., Koposov, S. E., Erkal, D., et al. 2021, ApJ, 911, 149
Libralato, M., Bellini, A., Bedin, L. R., et al. 2014, A&A, 563, A80
Libralato, M., Bellini, A., van der Marel, R. P., et al. 2018, ApJ, 861, 99
Libralato, M., Bellini, A., Vesperini, E., et al. 2022, ApJ, 934, 150
Lima Neto, G. B., Gerbal, D., & Márquez, I. 1999, MNRAS, 309, 481
Lindegren, L., Hernández, J., Bombrun, A., et al. 2018, A&A, 616, A2
Macciò, A. V., Paduroiu, S., Anderhalden, D., Schneider, A., & Moore, B. 2012, MNRAS, 424, 1105
Macciò, A. V., Paduroiu, S., Anderhalden, D., Schneider, A., & Moore, B. 2013, MNRAS, 428, 3715
Magorrian, J. 1999, MNRAS, 302, 530
Mamon, G. A., Biviano, A., & Boué, G. 2013, MNRAS, 429, 3079
Martin, N. F., de Jong, J. T. A., & Rix, H.-W. 2008, ApJ, 684, 1075
Martinez, G. D. 2015, MNRAS, 451, 2524
Martínez-García, A. M., del Pino, A., Aparicio, A., van der Marel, R. P., & Watkins, L. L. 2021, MNRAS, 505, 5884
Martínez-García, A. M., del Pino, A., Łokas, E. L., van der Marel, R. P., & Aparicio, A. 2023, MNRAS, 526, 3589
Massari, D., Helmi, A., Mucciarelli, A., et al. 2020, A&A, 633, A36
Massari, D., Posti, L., Helmi, A., Fiorentino, G., & Tolstoy, E. 2017, A&A, 598, L9
McConnachie, A. W. 2012, AJ, 144, 4
McGaugh, S. S., & Wolf, J. 2010, ApJ, 722, 248
Merritt, D. 1985, ApJ, 289, 18
Moore, B. 1994, Natur, 370, 629
Muraveva, T., Clementini, G., Garofalo, A., & Cusano, F. 2020, MNRAS, 499, 4040
Navarro, J. F., Frenk, C. S., & White, S. D. M. 1997, ApJ, 490, 493
Odenkirchen, M., Grebel, E. K., Harbeck, D., et al. 2001, AJ, 122, 2538
Osipkov, L. P. 1979, SvAL, 5, 42
Pace, A. B., Erkal, D., & Li, T. S. 2022, ApJ, 940, 136
Pianta, C., Capuzzo-Dolcetta, R., & Carraro, G. 2022, ApJ, 939, 3
Plummer, H. C. 1911, MNRAS, 71, 460
Pontzen, A., & Governato, F. 2012, MNRAS, 421, 3464
Pryor, C., & Kormendy, J. 1990, AJ, 100, 127
Read, J. I., & Gilmore, G. 2005, MNRAS, 356, 107
Read, J. I., Iorio, G., Agertz, O., & Fraternali, F. 2017, MNRAS, 467, 2019
Read, J. I., Mamon, G. A., Vasiliev, E., et al. 2021, MNRAS, 501, 978
Read, J. I., & Steger, P. 2017, MNRAS, 471, 4541
Read, J. I., Walker, M. G., & Steger, P. 2018, MNRAS, 481, 860
Robitaille, T., & Bressert, E., 2012 APLpy: Astronomical Plotting Library in Python, Astrophysics Source Code Library, ascl:1208.017
Romanowsky, A. J., & Kochanek, C. S. 1997, MNRAS, 287, 35
Rybicki, G. B. 1987, in IAU Symp. 127, Structure and Dynamics of Elliptical Galaxies, ed. P. T. de Zeeuw (Dordrecht: D. Reidel), 397
Salpeter, E. E. 1955, ApJ, 121, 161
Sameie, O., Yu, H.-B., Sales, L. V., Vogelsberger, M., & Zavala, J. 2020, PhRvL, 124, 141102
Schwarz, G. 1978, AnSta, 6, 461
Sedain, A., & Kacharov, N. 2023, arXiv:2305.11256
Ségall, M., Ibata, R. A., Irwin, M. J., Martin, N. F., & Chapman, S. 2007, MNRAS, 375, 831
Sérsic, J. L. 1963, BAAA, 6, 41
Sérsic, J. L. 1968, Atlas de Galaxias Australes (Cordoba: Observatorio Astronomico)
Simonneau, E., & Prada, F. 2004, RMxAA, 40, 69
Sohn, S. T., del Pino Molina, A., Besla, G., et al. 2021, HST Proposal, 16737
Sohn, S. T., Patel, E., Besla, G., et al. 2017, ApJ, 849, 93
Spencer, M. E., Mateo, M., Olszewski, E. W., et al. 2018, AJ, 156, 257
Strigari, L. E., Frenk, C. S., & White, S. D. M. 2018, ApJ, 860, 56
Strigari, L. E., Koushiappas, S. M., Bullock, J. S., & Kaplinghat, M. 2007, PhRvD, 75, 083526
Sugiura, N. 1978, Commun. Stat. Theory Methods, 7, 13
Tiret, O., Combes, F., Angus, G. W., Famaey, B., & Zhao, H. S. 2007, A&A, 476, L1
Tolstoy, E., Irwin, M. J., Helmi, A., et al. 2004, ApJL, 617, L119
van den Bosch, F. C. 1997, MNRAS, 287, 543
van der Marel, R. P. 1991, MNRAS, 253, 710
van der Marel, R. P., Alves, D. R., Hardy, E., & Suntzeff, N. B. 2002, AJ, 124, 2639
van der Marel, R. P., & Anderson, J. 2010, ApJ, 710, 1063

van der Marel, R. P., Anderson, J., Bellini, A., et al. 2023, JWST Proposal, 4513
van der Marel, R. P., & Franx, M. 1993, ApJ, 407, 525
van der Marel, R. P., Magorrian, J., Carlberg, R. G., Yee, H. K. C., & Ellingson, E. 2000, AJ, 119, 2038
van der Walt, S., Colbert, S. C., & Varoquaux, G. 2011, CSE, 13, 22
Van Rossum, G., & Drake, F. L. 2009, Python 3 Reference Manual (Scotts Valley, CA: CreateSpace)
Vasiliev, E. 2019, MNRAS, 482, 1525
Vasiliev, E., & Baumgardt, H. 2021, MNRAS, 505, 5978
Vera-Ciro, C. A., Sales, L. V., Helmi, A., & Navarro, J. F. 2014, MNRAS, 439, 2863
Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, NatMe, 17, 261
Vitral, E. 2021, MNRAS, 504, 1355
Vitral, E., Kremer, K., Libralato, M., Mamon, G. A., & Bellini, A. 2022, MNRAS, 514, 806
Vitral, E., Libralato, M., Kremer, K., et al. 2023a, MNRAS, 522, 5740
Vitral, E., & Mamon, G. A. 2020, A&A, 635, A20
Vitral, E., & Mamon, G. A. 2021, A&A, 646, A63
Vitral, E., Sohn, S. T., Bellini, A., et al. 2023b, HST Proposal, 17434
Walker, M. G., Caldwell, N., Mateo, M., et al. 2023, ApJS, 268, 19
Walker, M. G., Mateo, M., Olszewski, E. W., et al. 2007, ApJS, 171, 389
Walker, M. G., Mateo, M., Olszewski, E. W., et al. 2009, ApJ, 704, 1274
Walker, M. G., Olszewski, E. W., & Mateo, M. 2015, MNRAS, 448, 2717
Walker, M. G., & Peñarrubia, J. 2011, ApJ, 742, 20
Wang, J., Hammer, F., Yang, Y., et al. 2024, MNRAS, 527, 7144
Wang, W., Zhu, L., Jing, Y., et al. 2023, ApJ, 956, 91
Watkins, L. L., van der Marel, R. P., Bellini, A., & Anderson, J. 2015a, ApJ, 803, 29
Watkins, L. L., van der Marel, R. P., Bellini, A., & Anderson, J. 2015b, ApJ, 812, 149
Wegg, C., Gerhard, O., & Bieth, M. 2019, MNRAS, 485, 3296
Wilkinson, M. I., Kleyna, J., Evans, N. W., & Gilmore, G. 2002, MNRAS, 330, 778
Wilkinson, M. I., Kleyna, J. T., Evans, N. W., et al. 2004, ApJL, 611, L21
Wilson, A. G. 1955, PASP, 67, 27
Zhao, H. 1996, MNRAS, 278, 488
Zhu, K., Lu, S., Cappellari, M., et al. 2024, MNRAS, 527, 706
Zhu, L., van de Ven, G., Watkins, L. L., & Posti, L. 2016, MNRAS, 463, 1117