Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **453,** 541–560 (2015) doi:10.1093/mnras/stv1622

Detection of satellite remnants in the Galactic halo with *Gaia*– III. Detection limits for ultrafaint dwarf galaxies

T. Antoja,[1★†] C. Mateu,[2,3] L. Aguilar,[2] F. Figueras,[4] E. Antiche,[4] F. Hernández-Pérez,[3] A. G. A. Brown,[5] O. Valenzuela,[6] A. Aparicio,[7] S. Hidalgo[7] and H. Velázquez[2]

[1]*Scientific Support Office, Directorate of Science and Robotic Exploration, European Space Research and Technology Centre (ESA/ESTEC), Keplerlaan 1, NL-2201 AZ Noordwijk, the Netherlands*
[2]*Instituto de Astronomia, Universidad Nacional Autónoma de México, Apartado Postal 877, 22860 Ensenada, B.C., México*
[3]*Centro de Investigaciones de Astronomía, AP 264, Mérida 5101-A, Venezuela*
[4]*Departament d'Astronomia i Meteorologia, Institut de Ciències del Cosmos, Universitat de Barcelona, IEEC, Martí Franquès 1, E-08028 Barcelona, Spain*
[5]*Sterrewacht Leiden, Leiden University, PO Box 9513, NL-2300 RA Leiden, the Netherlands*
[6]*Instituto de Astronomia, Universidad Nacional Autónoma de México, Apartado Postal 70264, 04510 México D.F., México*
[7]*Instituto de Astrofísica de Canarias, Calle Vía Lactea s/n, E-38205 La Laguna, Tenerife, Spain*

Accepted 2015 July 15. Received 2015 July 15; in original form 2015 April 8

ABSTRACT

We present a method to identify ultrafaint dwarf galaxies (UFDGs) candidates in the halo of the Milky Way using the future *Gaia* catalogue and we explore its detection limits and completeness. The method is based on the Wavelet Transform and searches for overdensities in the combined space of sky coordinates and proper motions, using kinematics in the search for the first time. We test the method with a *Gaia* mock catalogue that has the Gaia Universe Model Snapshot as a background, and use a library of around 30 000 UFDGs simulated as Plummer spheres with a single stellar population. For the UFDGs, we use a wide range of structural and orbital parameters that go beyond the range spanned by real systems, where some UFDGs may remain undetected. We characterize the detection limits as function of the number of observable stars by *Gaia* in the UFDGs with respect to that of the background and their apparent sizes in the sky and proper motion planes. We find that the addition of proper motions in the search improves considerably the detections compared to a photometric survey at the same magnitude limit. Our experiments suggest that *Gaia* will be able to detect UFDGs that are similar to some of the known UFDGs even if the limit of *Gaia* is around 2 mag brighter than that of SDSS, with the advantage of having a full-sky catalogue. We also see that *Gaia* could even find some UFDGs that have lower surface brightness than the SDSS limit.

Key words: methods: data analysis – Galaxy: formation – Galaxy: halo – galaxies: dwarf – dark matter – astrometry.

1 INTRODUCTION

The current cosmological cold dark matter paradigm posits the assemblage of large structures in the Universe from smaller ones (Press & Schechter 1974; White & Rees 1978; Springel, Frenk & White 2006). A galaxy like ours must have formed by the merger of a large number of smaller systems, that even today, must be still in the process of being accreted. A discrepancy between the predicted and observed number of galaxy satellites has given rise to the so called 'missing satellite problem' (Klypin et al. 1999; Moore et al. 1999). However, in recent years, an entirely new population of hitherto unknown systems with very low luminosity and surface brightness, dominated by dark matter, the so called 'ultrafaint dwarf galaxies' (UFDGs), has been discovered, opening up the possibility of resolving this problem (e.g. Simon & Geha 2007; Bullock 2010). The knowledge of their structural properties, chemical abundances and stellar populations is also key to understanding fundamental issues (see review by Belokurov 2013) like the process of star formation and the role of feedback in these relatively low-mass environments (Brown et al. 2014); how to distinguish between a

★E-mail: tantoja@cosmos.esa.int
†ESA Research Fellow.

dwarf galaxy and a globular cluster in some extreme cases (e.g. Segue 1 and 2, Willman 1, Boo II and CmB; Forbes & Kroupa 2011); or to what extent UFDGs could have contributed to the stellar population found in the Galactic halo today (e.g. Kirby et al. 2008).

So far, all known UFDGs were discovered as overdensities in deep large-area photometric surveys, the vast majority in SDSS (e.g. Willman et al. 2005a,b; Belokurov et al. 2006, 2007, 2008, 2009, 2010; Zucker et al. 2006b,a), together with recent findings in Pan-STARRS (Laevens et al. 2015) and the Dark Energy Survey (DES; Koposov et al. 2015; Bechtol et al. 2015).

The ESA *Gaia* mission, launched in 2013 December, offers excellent prospects for the discovery of new members of the UFDG population. *Gaia* (de Bruijne 2012; Perryman et al. 2001) will measure accurate positions, parallaxes and proper motions for all stars out to its survey limit of $G = 20$ ($V = 20$–22, depending on the colour of the source), where G is the white light photometric pass-band of *Gaia* (Jordi et al. 2010). Multicolour photometry will be obtained for all stars and radial velocities will be collected for stars at $G_{\rm RVS} < 16$ mag, where $G_{\rm RVS}$ indicates the pass-band of the Radial Velocity Spectrograph on-board. *Gaia* will also provide astrophysical information on all the sources observed, primarily through multicolour photometry. The astrophysical parameters of all *Gaia* sources will be provided as part of the survey data products (Bailer-Jones et al. 2013).

Although the *Gaia* survey is not as deep as SDSS, Pan-STARRS or DES, it is all sky at a spatial resolution comparable to that of *Hubble Space Telescope* (*HST*), and will deliver high-accuracy astrometry (positions and proper motions) for all sources. The combination of these unique features is what makes the comparatively shallow survey of *Gaia* potentially powerful in the search for UFDGs. Here, we aim to exploit this in a technique to identify UFDGs. The combination of positions and kinematics has proven to be most efficient in the search for dark matter subhalos in cosmological simulations (e.g. Onions et al. 2012; Behroozi, Wechsler & Wu 2013). But, to our knowledge, this is the first time that both configuration space and kinematics are included in the search of UFDGs. As we will show, *Gaia* will enable us to probe parts of the UFDG parameter space which have not been covered before, and will allow for a comprehensive study of the spatial distribution around the Milky Way (MW) of this faint galaxy population.

This work continues the series (Brown, Velázquez & Aguilar 2005; Mateu et al. 2011), in which we have assumed the task of building ever more realistic *Gaia* mock catalogues, and used them to test tools that we have introduced to detect and characterize substructure in the stellar halo of our Galaxy.

In Section 2, we introduce our *Gaia* mock catalogue, which serves as our laboratory to study the detectability of UFDGs. This includes a stellar background and our synthetic UFDGs. The details of the *Gaia* selection function and error model used to generate the *Gaia* observables are described as well. In Section 3, we present our detection tool, which consists of a peak identifier that is applied in the sky and proper motion planes, a cross-matcher that identifies peaks with common members in both planes, and a procedure to evaluate the statistical significance of the matched peaks. Section 4 presents our results. Detection limits are shown as a function of astrophysical parameters and of 'effective parameters', namely a combination of the former on which our detection method depends directly. In Section 5, we summarize the limits of our method as well as the assumptions that it is based on. Our conclusions are presented in Section 6.

2 THE *GAIA* MOCK CATALOGUE

Our *Gaia* mock catalogue is the stage where we assess the success, efficiency and limits of our UFDGs detection technique. As such, it represents a controlled, but realistic environment. There are several elements that compose it. First, we need a model of the Galaxy, from which suitable stellar backgrounds can be extracted (Section 2.1). We also need a mass model for our synthetic UFDGs and a stellar population model (Section 2.2). The latter is because our UFDGs are not merely ensembles of particles, but stellar properties must be assigned to them, as they impinge on the value and quality of their *Gaia* observables. The previous elements allow us to assemble an extensive library of UFDGs at various distances and with a wide range of intrinsic parameters, projected against stellar backgrounds at several Galactic latitudes. We then use a *Gaia* selection function and error model to transform the theoretical quantities into realistic *Gaia* observables, as our detection method should work based on them only (Section 2.3). In Section 2.4, we present the filtering method that we use to eliminate foreground stars. Finally, we examine the nature of the UFDGs projections in the sky and proper motion planes, as these are the basic input variables that our method works on (Section 2.5).

2.1 The galactic background model

We use as a Galactic background the Gaia Universe Model Snapshot (GUMS) from Robin et al. (2012), which is a simulated catalogue of the sources expected to be observed by *Gaia*, at a fixed epoch. It includes the simulation of Galactic sources, Solar system and extragalactic objects.

We note here that *Gaia* will observe large numbers (potentially millions) of galaxies and about half a million QSOs (de Souza et al. 2014; de Bruijne et al. 2015a), which will all appear as faint point sources and could thus complicate the search for UFDGs. However, discrete source classification will be part of the published data (Bailer-Jones et al. 2013) and in this work we assume that we can rely on this to filter out galaxies and QSOs (but see Section 5). Therefore, we use only Galactic sources, and restrict the catalogue to a range in latitude of $20^\circ < |b| < 90^\circ$, to avoid the crowding and high extinction expected near the Galactic plane.

The Galactic sources in GUMS are generated based on the Besançon Galactic Model, which includes the Galactic Thin and Thick Discs, Bulge and Halo, based on appropriate density laws, kinematics, star formation histories (SFHs), enrichment laws, initial mass function (IMF) and total luminosities for each of the component populations, described in detail in Robin et al. (2012). Objects are simulated with masses down to the hydrogen burning limit, corresponding to spectral types down to $\sim L5$. Binary and multiple star systems are also simulated (see details in Arenou 2011), introduced with a probability that depends on the mass and evolutionary state of the primary star. The probability distribution for the separations is assumed to be a lognormal with the parameters reported by Duquennoy & Mayor (1991) for (primary) stars down to solar masses, and from Close et al. (2003) for low-mass stars.

2.2 The UFDG model

2.2.1 *The dynamical model*

For our basic synthetic UFDG dynamical model, we use a simple Plummer sphere with isotropic velocity distribution. A particular

realization of this model is uniquely defined by two of the following parameters:

(i) $M_{\rm T}$: total mass
(ii) r_o: core radius
(iii) $r_{\rm h}$: half-mass radius
(iv) $\sigma_{\rm V}$: velocity dispersion (3D).

The total gravitational binding energy of a Plummer sphere is

$$W_{\rm P} = -f\,\frac{M_{\rm T}^2}{r_{\rm o}}, \quad \text{where} \quad f \equiv \frac{3\pi G}{32}.$$

By virial equilibrium, we can establish a relation between the total mass, core radius and velocity dispersion,

$$2K = -W \quad \Longrightarrow \quad \sigma_{\rm V}^2 = \frac{f\,M_{\rm T}}{r_{\rm o}}.$$

In appropriate astronomical units, the previous relation is

$$\left(\frac{\sigma_{\rm V}}{\rm km\,s^{-1}}\right) = 0.035\,59\,\sqrt{\frac{(M_{\rm T}/\,{\rm M}_\odot)}{(r_{\rm o}/\,{\rm pc})}}.$$

Observationally, the scalelength usually reported is the half-light radius. Under the assumption of a position independent mass-to-light ratio (i.e. well mixed), the half-mass radius and its light counterpart coincide. We assume this and use $r_{\rm h}$ indistinctly as the half-light, or the half-mass radius. For a Plummer sphere, the relation between the core and half-mass radius is $r_{\rm h} = 1.304\,77\,r_{\rm o}$. Then, the previous relation between velocity dispersion, mass and radius can be written as

$$\left(\frac{\sigma_{\rm V}}{\rm km\,s^{-1}}\right) = 0.040\,66\,\sqrt{\frac{(M_{\rm T}/\,{\rm M}_\odot)}{(r_{\rm h}/\,{\rm pc})}}. \quad (1)$$

For a model with a given $r_{\rm h}$ and $\sigma_{\rm V}$, the mass-to-light ratio M/L of the UFDG is given by the ratio of its total mass $M_{\rm T}$, derived from equation (1), and the chosen total V-band stellar luminosity L_V. The number of particles in the realization ($N_{\rm s}$) and the total stellar mass ($M_{\rm s}$) are a consequence of the assumed total luminosity, SFH and stellar mass function.

2.2.2 *The stellar population model*

We simulate the stellar population of the UFDGs as a single star formation burst with an age of 12 Gyr and metallicity $Z = 0.0001$, assuming a Chabrier (2003) IMF. We use the HB13 stellar population synthesis code from Hernández-Pérez & Bruzual (2013), which allows for a consistent treatment of isolated and binary stars. The prescriptions assumed in this code are similar, yet not identical, to those used for the statistical orbital properties of binaries in GUMS (see Section 2.1). In HB13, the binary probabilities and orbital parameters are randomly drawn and assigned to each primary star in the population at age zero and the evolution is followed using the Hurley, Tout & Pols (2002) binary evolution code. Binary probabilities are assumed to depend on the mass of the primary using the prescription from Lada (2006), and the distribution of periods, and thus, separations, of Duquennoy & Mayor (1991). The resulting M_G versus intrinsic $V-I$ colour Hess diagram for the stellar population used for the UFDGs is shown in Fig. 1.

2.2.3 *Parameters of the simulated UFDGs*

Each of the simulated UFDG has nine free parameters.

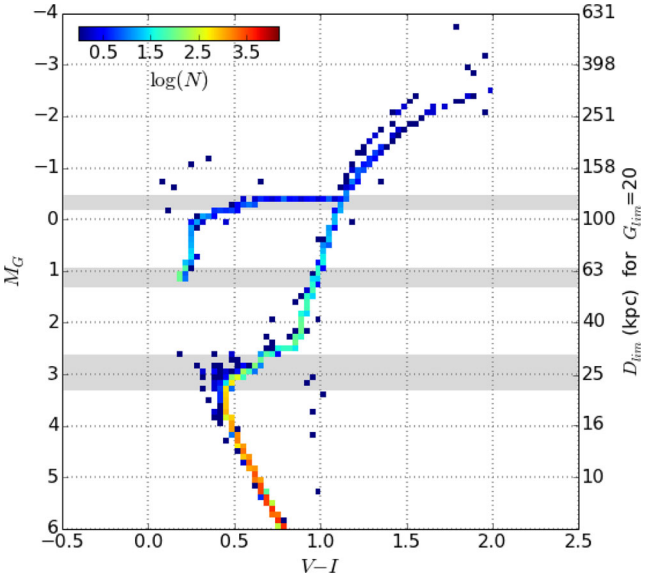

Figure 1. Hess diagram M_G versus intrinsic $V-I$ colour for the UFDG's stellar population model (see the text for details). The colour scale is proportional to the logarithmic number of stars in each bin. The right y-axis indicates the maximum distance up to which a star with a given M_G will be observable by *Gaia*, given the expected magnitude limit of $G_{\rm lim} = 20$ (assuming $A_V = 0$). For the grey bands see discussion at end the of Section 2.3

(i) Intrinsic parameters.

(a) Total V-band luminosity: L_V.
(b) Half-light radius: $r_{\rm h}$.
(c) Velocity dispersion: $\sigma_{\rm V}$.

(ii) Extrinsic parameters.

(a) Heliocentric distance: D.
(b) Position in the sky (l, b).
(c) Galactocentric velocity vector modulus: $V_{\rm gal}$.
(d) Azimuthal and latitudinal orientation angles of the galactocentric velocity vector: $\phi_{\rm V}$, $\theta_{\rm V}$.

We have generated a set of libraries with a total of ∼30 000 UFDGs covering large ranges of the nine parameters (see Table 1). Our main library is generated with the following parameters drawn at random: (i) the number of stars that would be observable by *Gaia* $N_{\rm obs}$, (ii) the heliocentric distance D, (iii) the apparent size of the UFDGs in the sky θ and (iv) in the proper motion plane $\Delta\mu$, and (v) the centre-of-mass velocity. These quantities are described in detail in Sections 2.3 and 2.5. The first four parameters are generated from a uniform distribution in a logarithmic scale. For the last one, the angles $\phi_{\rm V}$ and $\theta_{\rm V}$ and the modulus $V_{\rm gal}$ are generated following a uniform distribution, with $V_{\rm gal}$ between zero and the local escape velocity for the Galaxy.[1] The remaining parameters (namely L_V, $r_{\rm h}$ and $\sigma_{\rm V}$) are obtained from the ones above. We also require $N_{\rm obs}$ to be at least 10 in this library. This library is designed with particular goals described in detail in Sections 4.2 and 4.3.

Fig. 2 illustrates the range explored in half-light radii and velocity dispersion (left-hand panel), as well as in mass-to-light ratio and total V-band luminosity (right). The observed values of these parameters for known UFDGs and classical dwarf spheroidal (dSph)

[1] We compute the escape velocity as $V_{\rm e} = V_{\rm c}\sqrt{2(1-\ln(R_{\rm gal}/r_{\rm t}))}$, with $V_{\rm c} = 200$ km s^{-1} and $r_{\rm t} = 200$ kpc.

Table 1. Ranges of parameters in the simulated UFDGs (first row) and parameters of the UFDG used as our fiducial case. See Sections 2.2.3 and 2.3 for definitions. $N_{\rm obs}$ and M/L can be obtained from the other parameters.

	L_V ($L_\odot$)	$r_{\rm h}$ (pc)	$\sigma_{\rm V}$ (km s^{-1})	D (kpc)	l (°)	b (°)	$V_{\rm gal}$ (km s^{-1})	$\phi_{\rm V}$ (°)	$\theta_{\rm V}$ (°)	M/L ($M_\odot/L_\odot$)	$M_{\rm s}$ ($M_\odot$)	$N_{\rm obs}$
Ranges	86.–7.5 × 10^6	5–4 000	1–500	10–250	0–180	0–90	24–550	0–360	−90–90	0.04–3.9 × 10^7	10–1.6 × 10^6	10–10^3
Fiducial	5 × 10^3	80	10	20	90	30	453	0	0	902	6 × 10^3	94

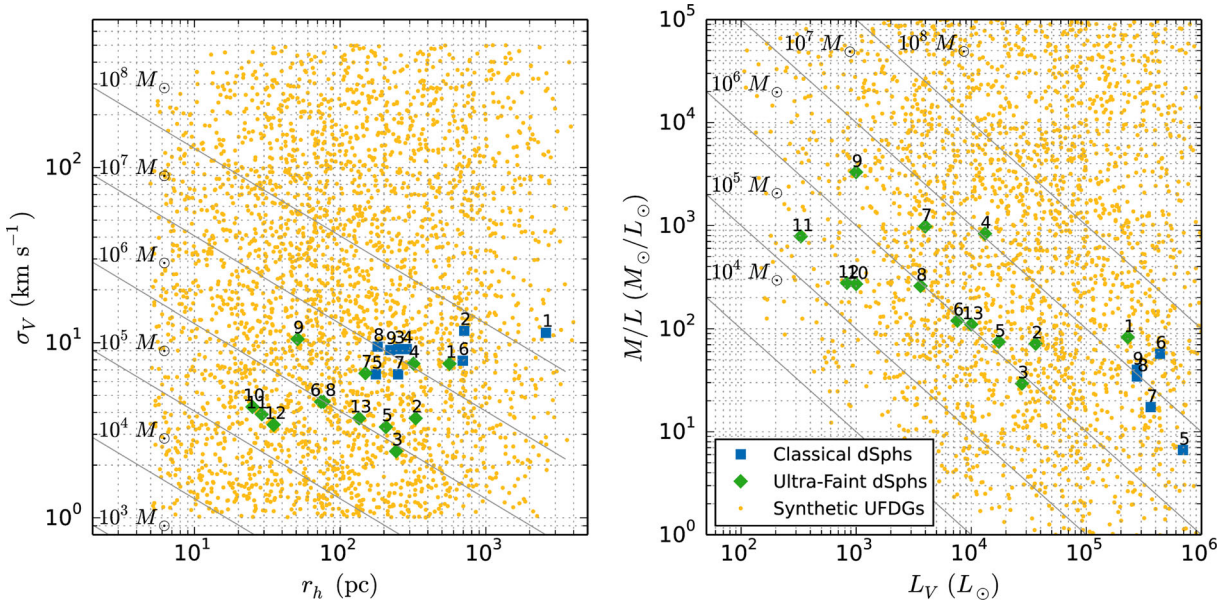

Figure 2. Left: $\sigma_{\rm V}$ versus $r_{\rm h}$. Right: M/L versus L_V. The yellow dots correspond to members of our UFDG library. Note that this library contains only systems with at least 10 stars observable by *Gaia*. Solid lines indicate constant total-mass models. Known UFDGs and classical dSph galaxies are shown with green diamonds and blue squares, respectively (data from McConnachie 2012). The labels correspond to: Sgr (1), For (2), LeoI (3), Scl (4), LeoII (5), Sex (6), Car (7), UMi (8), Dra (9) for blue squares and CVnI (1), Her (2), Boo (3), UMa (4), LeoIV (5), CVnII (6), UMaII (7), CmB (8), BooII (9), WilI (10), SegI (11), SegII (12), LeoV (13) for green diamonds. We do not include LeoT, which would not be observable with *Gaia*, and PscII that lacks measurements on some parameters.

galaxies are shown with green diamonds and blue squares, respectively. Note that the range explored by our synthetic library (yellow dots) is much larger than the observed one for $r_{\rm h}$, $\sigma_{\rm V}$ and M/L. In particular, the large range covered in $\sigma_{\rm V}$ results in a very large range of M/L (the library spans an even larger range of M/L than shown in Fig. 2). We are pushing the limits of the parameter space explored, towards regions where the detection would be observationally more difficult, i.e. towards larger $r_{\rm h}$ and $\sigma_{\rm V}$ (top and right areas of left-hand panel), and lower luminosity and high M/L (top and left areas of right-hand panel). The fact that $N_{\rm obs}$ is generated uniformly, together with the large scatter in luminosity for small $N_{\rm obs}$ due to stochastic effects, produces the diffuse boundary in L_V in the right-hand panel.

2.3 The *Gaia* selection function and error model

Here, we present our model for the *Gaia* observations that includes the selection function and the *Gaia* error model that we apply to the GUMS model and the simulated UFDGs. The *Gaia* observables are the five astrometric parameters (l, b, ϖ, μ_l, μ_b), the radial velocity, the *Gaia* photometry (including the *G Gaia* magnitude and the two broad-band magnitudes $G_{\rm BP}$ and $G_{\rm RP}$). The final *Gaia* catalogue will also provide three atmospheric parameters (metallicity, surface gravity and effective temperature) and extinction. The true values for these observables and parameters are obtained directly from the models. The conversion from the Johnson–Cousins photometric system to *Gaia* magnitudes is done following the transformation given in table 3 from Jordi et al. (2010). We do not consider extinction because all fields used in our study are at relatively high latitudes (at least 30°).

The GUMS model and the simulated UFDGs include binary and multiple systems. To determine which ones will be resolved by *Gaia*, we use a prescription used within the Data Processing and Analysis Consortium.[2] In this model, the minimum angular separation on the sky that *Gaia* can resolve depends on the apparent magnitudes of the stars in the system, with the minimum separation being ∼38 mas. For the unresolved cases, a single detection is considered by computing the total integrated magnitude, averaging positions and taking the atmospheric parameters (such as surface gravity) of the primary star in the system.

As an example, if we take a field[3] of 2° × 2° centred at $l = 90°$ and $b = 30°$, there are initially 25 521 objects, from which 57 per cent are single stars, 13 per cent are stars of resolved multiple systems

[2] http://www.cosmos.esa.int/web/gaia/dpac (Mignard et al. 2008)

[3] In what follows, we always work with 2° × 2° fields. To cover the same solid angle, regardless of latitude, we have converted the Galactic longitude of the stars l, to $l' = (l - l_0)\cos(b_0) + l_0$, where l_0 and b_0 are the longitude and latitude of the centre of the field, respectively. For simplicity, we use l instead of l' hereafter.

and 30 per cent are unresolved systems.[4] For a simulated UFDG at 50 kpc, these fractions are 63, 6 and 31 per cent, respectively.

To simulate *Gaia*-like errors for the GUMS catalogue and the simulated UFDGs, we use the code presented in Romero-Gómez et al. (2015), updated to the post-launch performance[5] as described in de Bruijne, Rygl & Antoja (2015b). Up to date information is available from the *Gaia* web pages.[6] The uncertainties on the astrometry, photometry and spectroscopy are mainly functions of the magnitude and colour. The geometrical factors and the effect of the number of passages due to the scanning law are also taken into account.[7] For the surface gravity, we take a constant error of 0.25 dex, based on table 4 of Bailer-Jones et al. (2013). Lacking a model of the *Gaia* performances for unresolved systems, we use the same prescriptions as for single or resolved stars. Only stars with magnitude $G < 20$ (the *Gaia* magnitude limit) are considered.

From all the *Gaia* astrometric observables, we cannot make use of parallaxes to infer distances to UFDGs stars. The median relative error in parallax of the stars in the UFDGs in the range of distances considered here is at least of 70 per cent and on average 170 per cent, since they can be very faint and distant objects. Besides, radial velocities are not available for most of the cases as 90 per cent of the UFDGs in the range of distances explored here have at most 10 per cent of stars that are brighter than the magnitude limit of the *Gaia* spectrograph ($G_{\rm RVS} = 16$). Therefore, we use as our observables only the two angular positions in the sky (l and b) and the two proper motions ($\mu_{\ell *} \equiv \mu_\ell \cos(b)$ and μ_b). The errors in the angular coordinates in the sky are of the order of 0.05–0.4 mas whereas in proper motion these are about 0.03–0.3 mas yr^{-1}.

The number of UFDG stars seen by *Gaia* $N_{\rm obs}$ (not the same as the total number of stars in the realization $N_{\rm s}$) is determined by the total stellar luminosity L_V of the system and the distance of the UFDG (given an assumed stellar population model). On the right axis of Fig. 1, we indicate the distance limit associated with the *Gaia* deepest magnitude $G = 20$. Note that at distances larger than 25 kpc only giant stars are observed. In Fig. 3, we show the number of UFDG stars observed by *Gaia* as function of luminosity and distance. For instance, UFDGs of luminosity around 1000 $L_\odot$ have no stars bright enough to be observed by *Gaia* beyond ∼40 kpc but they will have around 15 observable stars around 23 kpc. The oscillations with distance present at around 25, 60 and 120 kpc are because the type of stars of the UFDG population that *Gaia* can detect changes as it is observed at different distances, depending on whether or not features like the main-sequence turn-off are observable. These distances have been marked in the Hess diagram of Fig. 1 and they correspond to the main-sequence turn-off, the extreme horizontal branch and the horizontal branch, respectively. Note also the stochasticity around small numbers of observable stars.

The left-hand panel of Fig. 4 shows the range covered by our library in number of stars observable by *Gaia* $N_{\rm obs}$ and distance D to the Sun. To include the known UFDGs and classical dSphs in this plot, we have computed $N_{\rm obs}$ assuming the stellar population model described in Section 2.2.2, and the total luminosity and distance

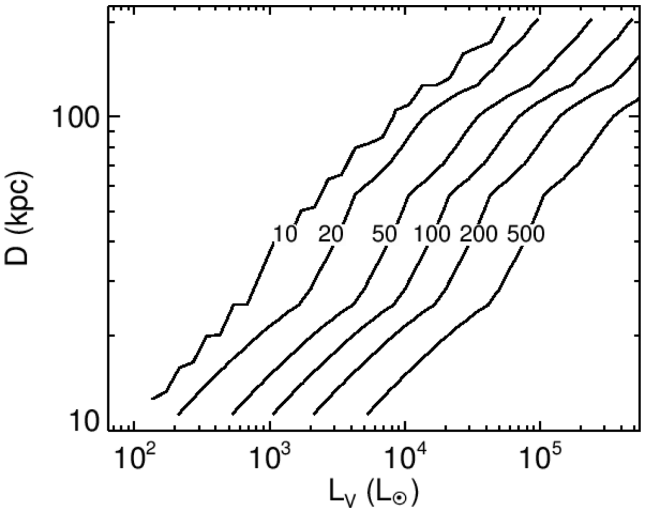

Figure 3. Number of stars observed by *Gaia* $N_{\rm obs}$, as a function of the total V-band luminosity and distance of the UFDG.

reported by McConnachie (2012) for these systems.[8] Here, we can see that there are real systems that go beyond the range covered by our library towards small number of observed stars. We must remember that $N_{\rm obs}$ is the number of stars that would be seen by *Gaia*, which has a ∼2 mag brighter limit than the SDSS[9] used to identify those systems (e.g. Belokurov et al. 2007). This is a limitation imposed by *Gaia* that we cannot get around. Note also that the boundaries of the regions spanned by the library in this panel are sharp by construction (see Section 2.2.3).

2.4 Filtering the foreground

Along a given line-of-sight (LOS), it is important to minimize the number of background stars[10] $N_{\rm BG}$ with respect to the number of stars in the UFDG. We use a parallax cut to filter out foreground disc stars, which have large parallaxes with small errors. Thus, we discard data for stars with $\varpi - e_\varpi > 0.1$ mas, i.e. an observed parallax which, within the errors, corresponds to distances smaller than 10 kpc. We also filter out foreground disc dwarfs with the implementation of a surface gravity $\log g$ cut: we discard stars with $\log g - e_{\log g} > 4$, where $\log g$ is the atmospheric parameter derived from the *Gaia* observables. With these two cuts, we reduce $N_{\rm BG}$ typically by an order of magnitude. For instance, there were 25 521 stars in the GUMS model in our fiducial field ($l = 90^\circ$ and $b = 30^\circ$) and with the cuts we reduce this number to $N_{\rm BG}$=1 413.

We stress that these cuts have been designed to have minimal loss of observable stars from the UFDGs, particularly at relatively large distances ($D > 10$ kpc), for which dwarf stars will not be

[4] After the cuts in parallax and surface gravity (see Section 2.4), these fractions become 54, 5 and 41 per cent, respectively. The relative increase of unresolved systems is because we are selecting large distances and giant stars (dwarf stars have been removed), which have higher binary fractions.
[5] The code was released at the *second Gaia Challenge Workshop* and is publicly available at https://github.com/mromerog/Gaia-errors
[6] http://www.cosmos.esa.int/web/gaia/science-performance
[7] http://www.cosmos.esa.int/web/gaia/table-6

[8] The SFH assumed in the stellar population model is reasonably representative of the SFH of known UFDGs. For simplicity, we assume the same model for the classical dwarfs to get a rough estimate of $N_{\rm obs}$, although these have very different SFHs.
[9] The SDSS survey is 2 mag deeper comparing the r and G bands, or between 1 and 2 mag deeper comparing the g and G bands. This is estimated by taking the two extreme colours of the stars in our simulated UFDGs, that is $V - I$ =0.25 and $V - I$ =1.5, and convert these to $G - r$ and $G - g$ colours following Jordi et al. (2010).
[10] The Galactic sources in GUMS are actually foreground and background stars. We use hereafter 'background' for simplicity.

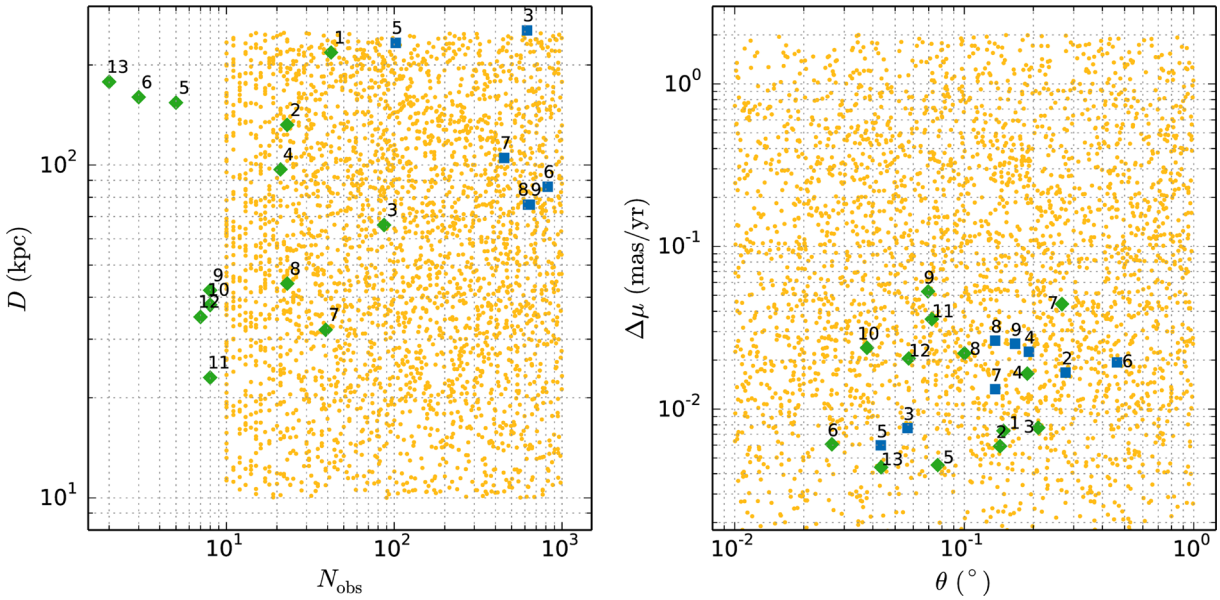

Figure 4. Left: D versus $N_{\rm obs}$. Right: apparent size $\Delta\mu$ in the proper motion plane versus apparent size θ in the sky. The yellow dots show the range spanned by the synthetic UFDG library. Known UFDGs and classical dSph galaxies are shown with green diamonds and blue squares, respectively (data from McConnachie 2012). The labels are as in Fig. 2.

observable by *Gaia*. The fraction of lost stars is up to 70 per cent for nearby UFDGs at 10 kpc. However, it goes down to 30 per cent at ∼18 kpc and is less than 10 per cent for distances larger than 25 kpc. Nevertheless, with the cuts we are maximizing the relative fraction of UFDG stars with respect to the background in all cases, given that the fraction of stars lost from the background is larger than that of the UFDGs. It is our experience that it is difficult to devise a unique algorithm that can identify our target systems at all distances, and so, limits are introduced as a necessary compromise. It is clear that specifically tailored algorithms could be used for nearby cases.

2.5 The sky and proper motion planes

The starting point of our detection procedure (Section 3) is the UFDGs projections in the sky and proper motion planes, thus it is essential to understand the behaviour of these projections of the systems and the background.

Fig. 5 shows the stars in the customary $2^\circ \times 2^\circ$ field of view of our fiducial simulated system in the sky (top) and proper motion planes (bottom). The parameters of this system are listed in Table 1. The stars belonging to the UFDG are coloured in green, while the background stars are in black. This system is hardly seen in the sky plane because it is very diffuse. But note how it is much more compact in the other plane. The compactness of the UFDGs in the proper motion plane is a general characteristic of most of our simulated UFDGs that improves considerably our search (Section 4), being a fundamental advantage of the *Gaia* data.

Note also the very different nature of the background in these planes. In the sky, the background is roughly constant, but becomes markedly non-uniform in the proper motion plane. This requires a special treatment when assigning significance to the peaks (Section 3.3).

The apparent sizes of an UFDG in the sky and proper motion planes are set by its intrinsic size and velocity dispersion, combined with its distance from the Sun. These sizes can span a wide range in both planes. The half-light angular size is given by

$$\theta(^\circ) \sim 0.0573 \frac{r_{\rm h}(\,{\rm pc})}{D(\,{\rm kpc})}. \quad (2)$$

In our synthetic library, $r_{\rm h}$ varies between (see Table 1) 5 and 4000 pc and D between 10 and 250 kpc (this is the approximate distance limit to detect at least ∼10 stars with *Gaia*, for a luminous UFDG with $L_V \sim 7 \times 10^4\,{\rm L}_\odot$). This implies a range of apparent angular sizes of [4 arcsec, 23°].

In the proper motion plane, the apparent size is

$$\Delta\mu(\,{\rm mas\,yr^{-1}}) \sim 0.211 \frac{\sigma_{\rm V}(\,{\rm km\,s^{-1}})}{D(\,{\rm kpc})}. \quad (3)$$

For values of $\sigma_{\rm V}$ in the range [1, 500] km s^{-1} and again D between 10 and 250 kpc, we end up with a range of apparent sizes of [0.0008, 11] mas yr^{-1} (but see below).

The right-hand panel of Fig. 4 shows the sizes spanned in the sky and proper motion planes by the UFDGs in our library. As we will see in Section 3, these two parameters are the most important, together with the number of visible stars in the UFDG, in determining the detectability of the system. We can see here that our library extends well beyond the spread covered by real systems. Note again that the boundaries of the library are sharp in this panel, as it is generated with *apparent* parameters drawn at random from a uniform distribution in a logarithmic scale.

It is also important to note that the apparent size of an UFDG in the proper motion space is greatly influenced by the observational errors. To illustrate this, we use a set of ∼1 300 simulated UFDGs located at different distances and with velocity dispersions between 15 and 25 km s^{-1}. Each black dot in the top panel of Fig. 6 shows the error in $\mu_{\ell*}$ (similar for μ_b) in each simulated UFDG computed as the median of all the individual stars errors in each UFDG. The proper motion error slightly increases with distance, as one would naively expect due to the fainter magnitudes. But the error also oscillates with distance. This is because the *Gaia* performances depend on the magnitude and colour of the star and the type of stars

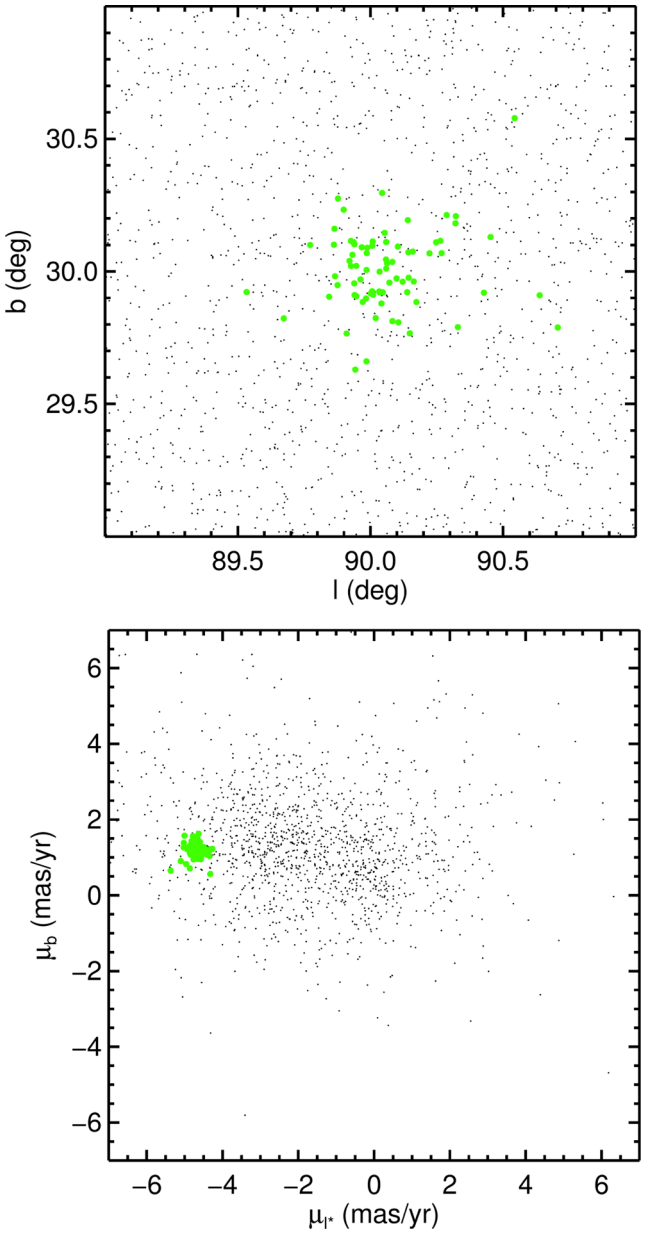

Figure 5. Sky (top) and proper motion plane (bottom) for the field of our fiducial UFDG in Table 1. The stars belonging to the system are shown as green dots while the background stars are in black.

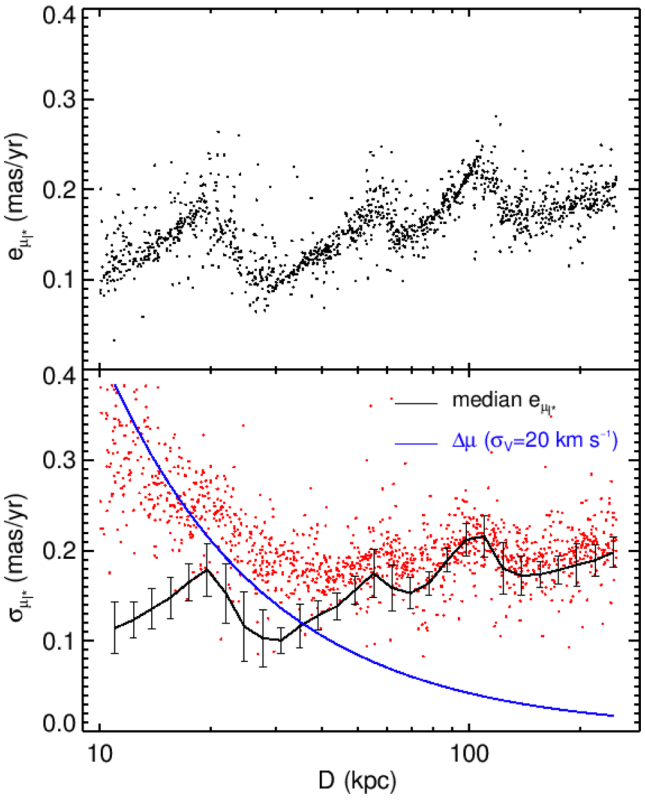

Figure 6. Top: median error $e_{\mu_{\ell *}}$ in $\mu_{\ell *}$ for ∼1300 synthetic UFDG with velocity dispersions around 20 km s^{-1} ($15 < \sigma_V < 25$ km s^{-1}) at different heliocentric distance. Bottom: dispersion $\sigma_{\mu_{\ell *}}$ in the $\mu_{\ell *}$ proper motion of the same set of synthetic UFDG (red dots). The black curve shows the median error in $\mu_{\ell *}$ calculated in logarithmic bins from the top panel. The error bars correspond to the standard deviation. The blue line is the expected size according to equation (3) for $\sigma_V = 20$ km s^{-1}.

in the UFDG that *Gaia* can detect, and thus the fraction of stars with certain magnitudes and colours, changes as it is observed at different distances (as seen in Section 2.3). One can see that the error has minima around 30, 70 and 135 kpc. These are distances slightly larger than the ones at which a sudden increase in the number of stars of certain types occurs. They correspond to the main-sequence turn-off, the extreme horizontal branch and the horizontal branch, as discussed previously (grey shaded stripes in Fig. 1).

The bottom panel of Fig. 6 shows the real size in the proper motion plane (red dots) computed as the standard deviation of the proper motion coordinate $\mu_{\ell *}$ ($\sigma_\mu \equiv \sigma_{\mu_{\ell *}}$) of the stars in each UFDG. We also overplot the error in $\mu_{\ell *}$ ($e_{\mu_{\ell *}}$) at each distance (black curve) taken as the median error in logarithmic bins from the top panel. The blue line in this plot shows the expected size according to equation (3) for a velocity dispersion of 20 km s^{-1}. We see that the sizes of the UFDGs decrease up to ∼40 kpc and for larger distances they follow the oscillations due to the *Gaia* errors. Once the size of the UFDGs is dominated by the observational error, the apparent size oscillates between 0.1 and 0.2 mas yr^{-1}. For smaller velocity dispersions, e.g. ∼5 km s^{-1}, the errors dominate already at a distance of 10 kpc. Therefore, the minimum apparent size of the UFDGs is set by the observational errors. For the range of parameters explored here, this is above 0.1 mas yr^{-1} in 99.5 per cent of the cases. In what follows, we take σ_μ instead of $\Delta\mu$ as a better measure of the apparent size of the UFDG in proper motion space.

As seen in Section 2.3, the errors in the angular coordinates in the sky are around 0.05–0.4 mas, which is negligible compared to the apparent sizes of the UFDGs. For this reason, we do not observe a similar effect in the sky plane.

3 THE DETECTION TOOLS

In this section, we present all the different elements that compose our detection method to identify UFDG candidates against the background.

Our strategy is as follows. We consider fields of view of 2° × 2° in the sky. We first detect overdensities independently in the sky and the corresponding proper motion planes. For this, we use the Wavelet Transform (WT, see Section 3.1). We do this for overdensities of different amplitudes and sizes, and keep the most significant ones. After this, we perform what we call the cross-match

of peaks (Section 3.2). This consists in counting how many stars belong *simultaneously* to a certain peak in the sky and to a certain peak in proper motion space. We do this for *all pairs* of peaks of any size between both spaces. For each cross-match, we finally compute the probability that the observed number of common stars is just a coincidence (Section 3.3–3.5). Cross-matches with a low probability are selected as possible UFDG candidates. Below we detail each of the steps of our method.

One may wonder why this separate treatment for the sky and proper motion planes. After all, what we are looking for is a single peak in the 4D space of positions in the sky and proper motion planes. This is because of the very different nature of these two planes, which results in the impossibility of having a natural metric in the combined space. Any metric will imply the introduction of an arbitrary dimensional scale which will limit the nature of the systems found. This is why we have preferred to work on the sky and proper motion planes separately, and then use the cross match procedure to relate peaks. The peaks that we do identify correspond to single peaks in the combined 4D space, but not necessarily using a unique metric, as our combination of different wavelet scales in both planes allows for a larger range of identified peaks than if using a single metric.

Although the whole detection process might seem complex, it is quite straightforward from the computational point of view. The entire algorithm takes a total of 40 s to run for our fiducial field of $2^\circ \times 2^\circ$ in a single Intel(R) Core(TM) i7-3770 CPU @3.40GHz. This might change depending on the LOS but, as a first approximation, the celestial sphere above $b = 30^\circ$ would require 86 h of CPU time, which in fact can be spread into several CPU for different LOS.

3.1 Wavelet analysis

To detect overdensities in the sky and proper motion planes like the ones of Fig. 5, we use the WT (Starck & Murtagh 2002). This can be thought of as a 'localized' Fourier transform that gives information about certain frequencies and where in the image these frequencies are located. Due to the wide range of apparent sizes of our simulated UFDGs (see Section 2.5), our method needs to be able to detect overdensities of different sizes. In the application here, a discrete set of frequencies (i.e. scales) are probed and we get information about the localization of those particular structures. We use here the *à trous* ('with holes') variant of the WT (Starck & Murtagh 2002) which computes a discrete set of scale-related 'views' of a 2D function or image. We have previously used this technique to detect moving groups in the stellar velocity distribution of the solar neighbourhood and surroundings (Antoja et al. 2008, 2012). To perform the calculations, we use the MR software developed by CEA (Saclay, France) and Nice Observatory.

Although the WT works at a specific scale, it can identify overdensities within some range in size. Nevertheless, it is important to realize that we are probing a discrete set of scales in the images, and therefore, it is critical to choose those scales wisely. We explore four logarithmically spaced scales in each plane within the ranges found in Section 2.5. For the sky, as we are dealing with fields of $2^\circ \times 2^\circ$, we have chosen the scales $0\overset{\circ}{.}05$, $0\overset{\circ}{.}1$, $0\overset{\circ}{.}2$ and $0\overset{\circ}{.}4$. Even though the higher scale puts a limit on the maximum size of an UFDG that can be detected in principle, the innermost parts of the more luminous UFDGs can still be detected, even if they have larger angular sizes. For the proper motion plane, we use scales of 0.12, 0.24, 0.48 and 0.96 mas yr^{-1}. Here, what we are missing are exceptional cases with extremely high velocity dispersion which are very close.

An example of the WT planes in the sky for our fiducial UFDG in Table 1 is shown at the top part of Fig. 7, while the bottom part shows the proper motion plane. In each case, the four scales mentioned are shown. The blue colours are proportional to the values of the WT.

After the WT, we search for relative maxima to detect the overdensities. The algorithm computes the Wavelet Probability (WP), that is the probability that the detected overdensities in the wavelet space are not due to Poisson noise. For this, it uses a model for this type of noise in wavelet space. This is done by first using the Anscombe transform (Anscombe 1948) that converts a signal with Poisson noise into Gaussian noise, for which the treatment in the WT planes is more straightforward (see Starck & Murtagh 2002 and references therein). Here, we will consider only WT peaks that have a WP of being real detections of WP $\geq$ 99.7 per cent (green crosses in Fig. 7), 95.4 $\leq$ WP $<$ 99.7 (orange crosses) and 68.2 $\leq$ WP $<$ 95.4 (red crosses), similar to $>3\sigma$, 2σ–3σ and 1σ–2σ significance levels in the Gaussian case, respectively. The size of the crosses in this figure indicates the size or scale that is being probed in each WT plane (also indicated in the top of the plots). There is an additional condition on the overdensities: they should have at least five stars to be considered a peak.

The UFDGs are optimally detected (i.e. with higher WP) when the scales probed are similar to their apparent sizes. In the example of Fig. 7, a black circle indicates the position and extension of the UFDG in the sky plane and in the proper motion plane. In the sky images, for scales that are smaller than the apparent size of the UFDG (two first panels), some peaks are detected inside the region occupied by the UFDG but with low WP (1σ or 2σ, red and orange crosses). For larger scales, which in this case are similar to the apparent size of the UFDG, the detection is above 3σ (green crosses). In other cases, the detection is always below 3σ or below 2σ because the UFDGs can be very diffuse in this plane, as already highlighted in Section 2.5.

On the other hand, the UFDG is very compact in proper motion space, and it stands out as an overdensity for all scales studied in the left part of the panels. In our example, the fiducial UFDG is always detected above 3σ (green crosses). In other cases, the best detection is for a particular scale that is close to the apparent size of the UFDG.

Note also how in both planes, a number of low-WP random detections appear (most of red crosses in Fig. 7). Because of this, we need to discard false overdensities and keep only good UFDG candidates (Section 3.2). Besides, in the proper motion case, two overdensities with high WP are also detected in the centre of the distribution for the two largest scales. These correspond to the peaks of the background distribution, which as we have seen in Fig. 5, is not uniform. Note that the distribution of background stars in the proper motion plane will be different for each LOS and therefore, its centroid will shift to different positions in this plane. This does not occur for the sky plane which presents a uniform background.

3.2 Cross-matching peaks in the two planes

So far, we have detected peaks in the sky and proper motion planes, separately. However, contrary to false detections, an UFDG is an overdensity in the 4D combined space l-b-$\mu_{\ell*}$-μ_b. This is precisely the feature that we need to exploit to identify UFDGs, beyond what has been currently achieved.

To do this, we list the stars contributing to the peaks identified separately in the sky and proper motion planes. By stars belonging to a certain peak, we mean those that are enclosed in a circle around the peak with the radius of the WT scale in the considered plane. In

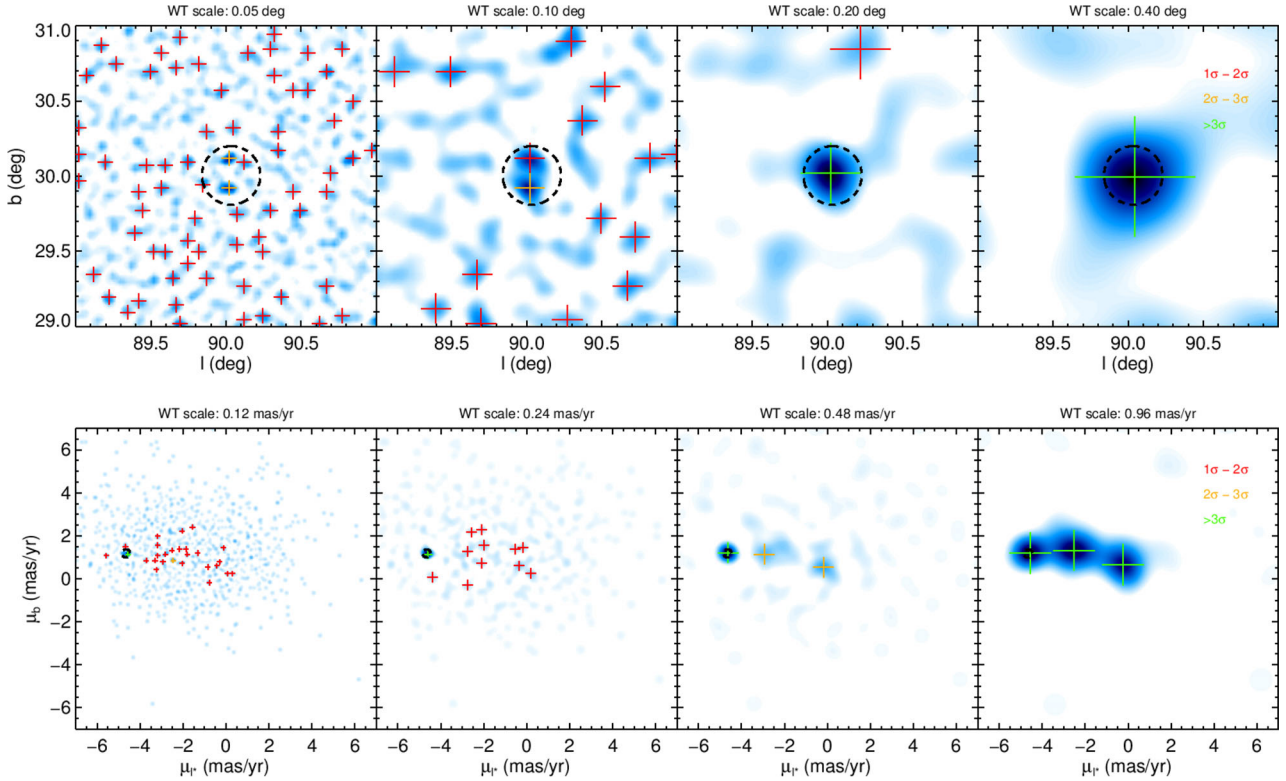

Figure 7. WT at different scales for our fiducial UFDG for the sky (top) and proper motion planes (bottom). The black dashed circle shows the true position and size of the system in each plane. The circle in the proper motion plane is very small but can be seen better in the smaller scales. The position is calculated as the median of the coordinates (positions and proper motions) of the stars in the UFDG that are observed by *Gaia*, while the size of the circle is taken as the maximum between the standard deviation of the coordinates. Red, orange and green crosses indicate peaks at between 1σ and 2σ, between 2σ and 3σ, and $>3\sigma$ significance, respectively.

practice, because *Gaia* is a point source catalogue, we can identify the stars by their *id* number. Then, we see whether a large fraction of these stars belong *simultaneously* to a certain peak in the sky and a peak in the proper motion plane. We call this 'cross-match of peaks'. This cross-match is done for every peak and at every scale in the sky, compared to every peak at every scale in the proper motion plane.

The computation of the probability of having this cluster of common stars occurring by chance is computed as described in Section 3.3. In Section 3.4, we explain how we filter out false detections. Because each UFDG can be detected in more than one scale, we also need to keep only independent detections. This is explained in Section 3.5.

3.3 Assessing the probability of the detections

Here, we describe the statistics machinery that we devised to assess the probability of detection, i.e. compute which detections have a very low probability of occurring by chance.

We are interested in $P(N_{\rm com}|\langle N_{\rm com}\rangle)$, i.e. the probability of observing a certain number of common stars $N_{\rm com}$ in a peak in the sky and a peak in the proper motion plane, given the expected number of common stars $\langle N_{\rm com}\rangle$. This probability[11] is simply given by the Poisson probability distribution function

$$P \equiv P(N_{\rm com}|\langle N_{\rm com}\rangle) = {\rm Poisson}(N_{\rm com}|\langle N_{\rm com}\rangle). \quad (4)$$

[11] Do not confuse this probability for the combined sky and proper motion planes P, with the WP used in Section 3.1

An estimate of $\langle N_{\rm com}\rangle$ is given by

$$\langle N_{\rm com}\rangle = \langle N_{\rm sky}\rangle \int_{A_\mu} \rho(\mu_{\ell*}, \mu_b)\, d\mu_{\ell*} d\mu_b, \quad (5)$$

where $\langle N_{\rm sky}\rangle$ is the expected number of stars in the $l-b$ peak and $\rho(\mu_{\ell*}, \mu_b)$ describes the (normalized) number density of stars in the proper motion plane, both under the assumption that no UFDG is present. A_μ indicates the area of the peak over which we are integrating, which is a circle with a radius given by the WT scale, centred on the $(\mu_{\ell*}, \mu_b)$ coordinates of the peak in question. For convenience, we use hereafter the logarithm of the probability, $\ln P$.

For simplicity, we assume that the background density in the $l-b$ plane is uniform, which is reasonable for the field size used, and therefore $\langle N_{\rm sky}\rangle = N_{\rm BG}(\pi r_{\rm sky}^2)/A_{\rm T}$, where $A_{\rm T}$ is the total area of the field in the sky plane (in our case 4 deg^2), $r_{\rm sky}$ is the wavelet scale in the plane of the sky and $N_{\rm BG}$ is the number of background stars in the field. The latter is computed from the observed data itself, by taking the eight fields adjacent to our problem field, with the same total area. For each of these fields, we compute the total number of stars and we take the median. This is a better estimation of the number of background stars than the total number of stars in the considered field, specially in cases of luminous UFDGs that have a number of observed stars that is not negligible compared to the number of background stars.

In the proper motion plane, however, it is crucial to account for the fact that the density is not constant and this is achieved by the integral term in equation (5), which, multiplied by $\langle N_{\rm sky}\rangle$ gives the number of common stars expected to lie within the area of the detected peak.

The distribution of stars in the proper motion plane $\rho(\mu_{\ell*}, \mu_b)$ is different depending upon the direction on the sky, so it must be computed independently for each field. We do this from the observed data itself, by taking the mentioned eight adjacent fields with the same total area. For each of these fields, we compute the density as a (normalized) 2D histogram in the $\mu_{\ell*} - \mu_b$ plane with a pixel size of 0.8 mas yr^{-1}. Taking these eight 2D histograms, we compute the pixel-by-pixel median density to obtain a statistically reliable estimate in each of the matrix cells, minimizing the effect of outliers. With this, we are assuming that the distribution of background proper motions remains similar among these adjacent fields. This is indeed the case. For instance, variations in the median proper motion in longitude and latitude among the adjacent fields are in general smaller than the pixel size. We numerically evaluate the integral in equation (5) using the trapezoid rule in 2D and bi-linear interpolation on the median density matrix.

It could happen that one or various of the adjacent fields contains UFDGs. This would yield a wrong estimation of $N_{\rm BG}$ and $\rho(\mu_{\ell*}, \mu_b)$. The fact that we use the median of the eight fields helps to alleviate this issue. However, in case very luminous UFDGs are present, our algorithm checks if the number of stars in one of the adjacent fields is significantly larger than in the others. This is done by checking that the dispersion in the number of stars in the eight fields is not larger than 2.5 times the square root of the median. In this case, the algorithm could be re-run without the field in question.

Instead of using the probability of equation (4), we can also use the significance s, defined as the number of times above the expected value of the distribution, scaled to the dispersion of the distribution

$$s = \frac{N_{\rm com} - \langle N_{\rm com}\rangle}{\sqrt{\langle N_{\rm com}\rangle}}. \quad (6)$$

The advantage of using s instead of $\ln P$ is that s is positive and it increases for more relevant detections.

3.4 Setting a threshold probability for detection

We are only interested in those detections which have very low probability (very negative $\ln P$) *and* also a number of common stars[12] $N_{\rm com} > \langle N_{\rm com}\rangle$. However, the central peak(s) of the background in the proper motion plane can appear also as a detection. This is because we estimate its expected number of stars using the adjacent fields and any small fluctuation above this value can give a significant detection, though with a larger $\ln P$. To filter peaks that correspond to the background and not to the UFDGs, we can conservatively select a relatively low (very negative) threshold value for $\ln P$, below which we consider detections to be relevant. However, we must realize that, as we lower the threshold value $\ln P_{\rm thres}$, although we eliminate spurious peaks, we start losing relevant detections, so this is a compromise between false positives and losing bona fide peaks.

We have explored this compromise on various LOS's. In Table 2, we list the percentage of recovered UFDGs per cent$_{\rm rec}$ and of false detections per cent$_{\rm false}$ from the total number of tested UFDGs, as a function of five different values for the threshold and 10 different LOS's. For values above -9.0, there are several fields where the percentage of false detections is above 20 per cent. For a threshold of -12.0, all false detections are at most 1.1 per cent. Although, we could choose a $\ln P_{\rm thres}$ in between those two to make it the most optimal, we conservatively choose $\ln P_{\rm thres} = -12$.

[12] The last condition is required in order to select only overdensities but not underdensities.

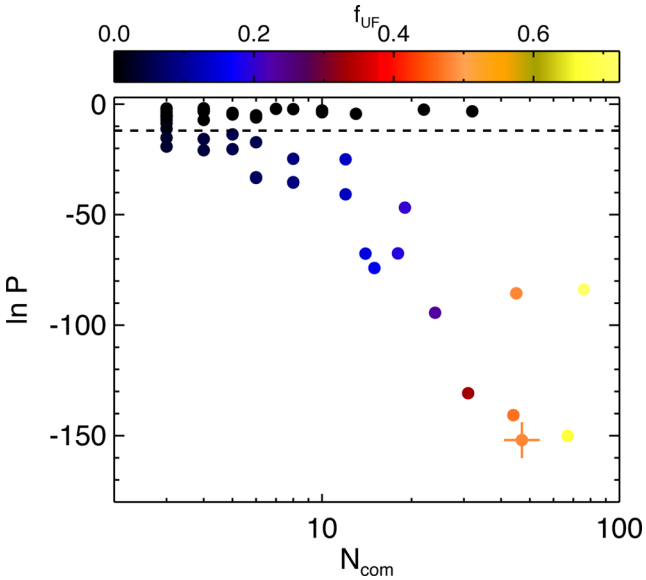

Figure 8. $\ln P$ as a function of the number of common stars $N_{\rm com}$ for our fiducial field and UFDG. The colour scale is proportional to $f_{\rm UF}$, i.e. the fraction of recovered stars from each UFDG. The dashed horizontal line mark the line below which we consider detections as relevant. The detection with the lowest $\ln P$ (marked with a cross) is what we take as the best *independent* detection (see the text for details).

3.5 Independent detections

As explained before, the cross-match of peaks is done for all peaks at all scales and a given UFDG can be detected in more than one scale. This means that we need to select which of the many detections made in a given field, are in fact independent detections, i.e. different objects.

We first organize all detections by increasing $\ln P$, choose the detection with the lowest $\ln P$ and compare its $l - b$ and $\mu_{\ell*} - \mu_b$ coordinates with the remaining detections. Now, we choose the next independent detection as the one with the lowest $\ln P$ that lies, both in the sky and proper motion planes, at a distance *larger* than the sum of the WT scales of the two detections (i.e. they do not overlap). We repeat this procedure until we have gone through all the available (relevant) detections.

To illustrate the behaviour of $\ln P$ and the selection of independent detections, in Fig. 8, we plot for our fiducial UFDG, $\ln P$ as a function of the number of common stars $N_{\rm com}$ for all detections in this field, that is the results of cross-matching all peaks at all scales in the sky and proper motion plane. In the plot, we use a colour scale proportional to $f_{\rm UF}$, defined as the number of stars from those $N_{\rm com}$ that truly belong to the UFDG divided by the total number of stars originally in the UFDG. In other words, $f_{\rm UF}$ is the fraction of recovered stars from each UFDG. Dots below the horizontal dashed line are relevant detections, i.e. with $\ln P$ below $\ln P_{\rm thres}$ (Section 3.4).

There is a correlation of $\ln P$ with $N_{\rm com}$. As expected, detections with larger numbers of common stars have on average lower $\ln P$. There is also a sequence that moves across the plot above the threshold. This corresponds to peaks in the proper motion background (note that they are black points, i.e. with no stars belonging

Table 2. Detections, real and false, as percentage of the number of UFDGs used in each LOS.

$\ln P_{\rm thres} =$		-3.0		-6.0		-9.0		-12.0		-15.0	
l	b	per cent$_{\rm rec}$	per cent$_{\rm false}$	per cent$_{\rm rec}$	per cent$_{\rm false}$	per cent$_{\rm rec}$	per cent$_{\rm false}$	per cent$_{\rm rec}$	per cent$_{\rm false}$	per cent$_{\rm rec}$	per cent$_{\rm false}$
90	30	86.8	150.2	86.7	105.0	85.4	50.3	83.4	0.3	81.7	0.0
90	42	85.7	44.9	85.4	38.7	84.3	1.2	83.0	0.0	81.6	0.0
90	55	84.6	37.4	84.4	35.9	83.6	1.8	82.6	1.1	81.8	0.3
90	68	81.4	39.0	81.2	38.9	80.8	27.8	80.2	0.3	79.3	0.0
90	80	84.6	7.8	84.2	0.5	83.4	0.1	82.8	0.0	82.2	0.0
170	30	87.7	68.6	87.4	67.5	85.9	0.4	84.6	0.1	83.4	0.0
170	42	84.7	6.3	84.4	0.6	83.8	0.2	83.2	0.0	82.6	0.0
170	55	79.9	47.6	79.7	43.8	79.3	0.4	78.8	0.0	77.8	0.0
170	68	82.2	1.7	82.1	0.5	81.7	0.0	81.5	0.0	80.6	0.0
170	80	82.9	3.2	82.6	0.4	81.5	0.0	81.1	0.0	80.8	0.0

to the UFDG). As explained before, these detections are filtered by our threshold.

Also, detections with the largest values of $f_{\rm UF}$ have low $\ln P$, i.e. they are significant detections. However, as the number of stars in common increases, the value of $\ln P$ decreases, reaches a minimum and then increases again. The minimum value occurs for detections at the optimum scales in the sky and proper motion planes. It is in this case that a large fraction of UFDG stars lie in the detected peak inside the WT scale, and the background is sufficiently low, so that the difference between the *observed* and expected number of common stars in the peak is maximal. Increasing the WT scale past the optimum values causes the inclusion of more stars of the UFDG in the peak but also more stars of the background that might not necessarily belong to both peaks in the sky and proper motion plane simultaneously, and hence, this causes $\ln P$ to go back to larger values. This is a very convenient behaviour which allows us to select detections at the optimum WT scales. The detection with the lowest $\ln P$ (marked with a cross) is what we take as the best *(and in this case, only) independent* detection. Finally, notice that in this example no false positives are picked up.

4 RESULTS

As we have seen in Section 2.2.3, we face a 9D parameter space. Even with our library of more than 30 000 different synthetic UFDGs, it is clear that we can cover only a limited amount of this vast hypervolume.

To explore this space with some order, we will rely first on a series of carefully curated ensembles of cases. In each one, all parameters, except two, will be kept fixed (Section 4.1). This allows us to take 2D sections of the original parameter space. Then, we will identify in Section 4.2, a reduced number of combinations of the original parameters that our detection procedure depends on directly, and which we call 'effective parameters'. In Section 4.3, we explore the limits and completeness of our method in the space of effective parameters, as well as in some of the most relevant original parameters. The effect introduced by changing the background level as we look at different LOS's is discussed in Section 4.4.

4.1 The physical parameter space

In the ensembles of tests presented here, we vary only two parameters, keeping the other seven parameters constant. The values for the fixed parameters are shown in Table 3.

Table 3. Values for the fixed parameters in the ensembles of UFDGs shown in Fig. 9.

L_V ($\rm L_\odot$)	$r_{\rm h}$ (pc)	$\sigma_{\rm V}$ ($\rm km\,s^{-1}$)	D (kpc)	l (°)	b (°)	$V_{\rm gal}$ ($\rm km\,s^{-1}$)	$\phi_{\rm V}$ (°)	$\theta_{\rm V}$ (°)
5×10^3	80	5	20	90	30	300	0	0

4.1.1 *The* $\rm r_h$ *versus* $\sigma_{\rm V}$ *plane*

In the first test, we use 625 synthetic UFDGs with varying $r_{\rm h}$ and $\sigma_{\rm V}$. As indicated in Table 3, all UFDGs are located at a fixed position in the sky at $l = 90°$, $b = 30°$, at a heliocentric distance of 20 kpc and have a luminosity of $L_V = 5 \times 10^3\,\rm L_\odot$. These are approximately the mean values for the observed UFDGs. The velocity dispersion varies logarithmically between 3 and 100 $\rm km\,s^{-1}$ and the half-light radius between 5 pc and 4 kpc. Due to stochastic variations in each realization, despite the luminosity and distance being constant, the number of detectable stars $N_{\rm obs}$ varies between 38 and 207. The results of the test are shown in the top panel of Fig. 9. Each symbol (squares and crosses) in this plot corresponds to one simulated UFDG. The colour scale in the panels is proportional to the detection significance (equation 6). Black crosses indicate UFDGs that were not detected. From this plot, we can evaluate the detection limits as a function of $r_{\rm h}$ and $\sigma_{\rm V}$.

UFDGs with $r_{\rm h}$ larger than ∼600 pc are not detected (for this fixed distance and luminosity). This is because their apparent size in the sky is very big, making them extremely diffuse. In fact, $r_{\rm h} = 700$ pc results in an angular size equal to the sky fields that we are using for our analysis ($2° \times 2°$). We also notice that for velocity dispersions below 20 $\rm km\,s^{-1}$ the detection significance depends mainly on the half-light radius (vertical contours). This is because in this regime the apparent size of the UFDG in proper motion is in fact constant and set by the observational errors (Section 2.5). Above this velocity dispersion, the contours bend slightly to the left, meaning that for a given size in the sky, the detection is more significant for lower velocity dispersions. Note that we are exploring velocity dispersions up to ∼100 $\rm km\,s^{-1}$, i.e. significantly larger than the typical velocity dispersion of $\sigma_{\rm V} \sim 5\,\rm km\,s^{-1}$ of known UFDG and classical dSph galaxies (McConnachie 2012).

4.1.2 *The* $\sigma_{\rm V}$ *versus* $V_{\rm gal}$ *plane*

The middle panel of Fig. 9 is a test with 625 UFDGs, where the velocity dispersion $\sigma_{\rm V}$ and the modulus of the velocity vector $V_{\rm gal}$ are varied. In this case, $V_{\rm gal}$ is varied linearly instead of logarithmically. Notice that in this ensemble, we only change the position and spread

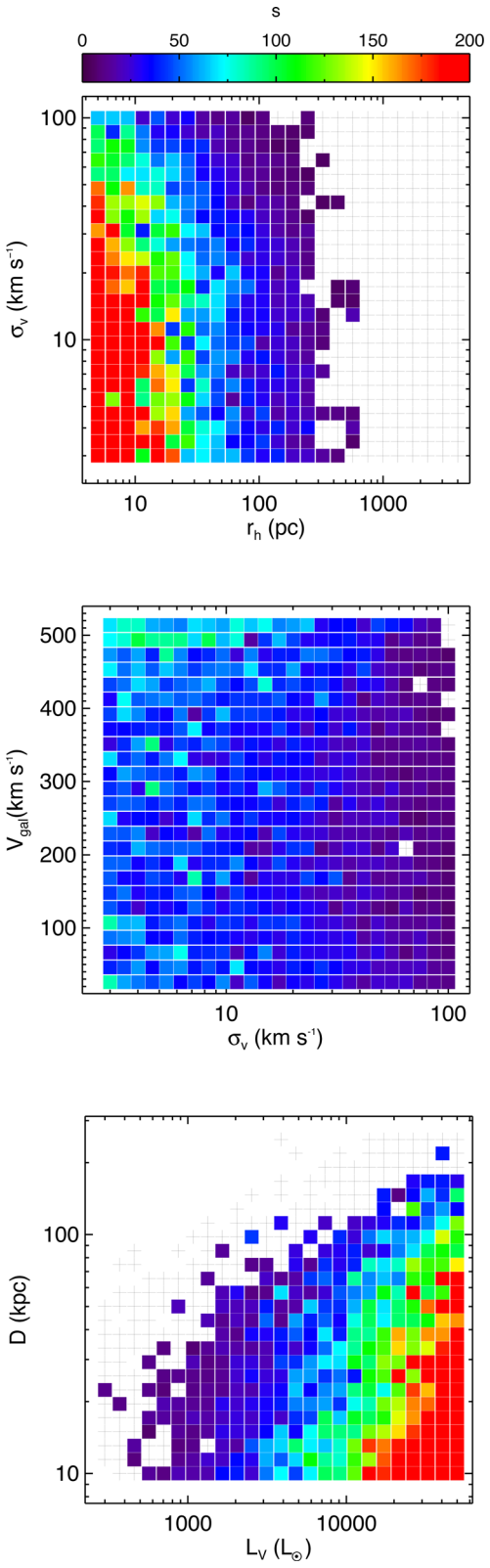

Figure 9. Detectability tests run with several ensembles of UFDGs with only two varying parameters: r_h versus σ_V(top), σ_V versus V_{gal} (middle), L_V versus distance D (bottom). The colour scale indicates the detection significance s. UFDGs with significance over 200 have been plotted with a colour saturated at this value. Black crosses indicate UFDGs that were not detected.

of the UFDG peak in the proper motion plane. In particular, we have chosen the values for the fixed velocity angles (θ_V and ϕ_V), so that the UFDG peak position moves horizontally across the proper motion plane as we vary V_{gal}, covering all possible contrasts between background and UFDG, and coinciding with the background peak for $V_{gal} \sim 150\,\mathrm{km\,s^{-1}}$.

Note how for a fixed value of V_{gal}, the best detections are the ones for lower velocity dispersions, which produce more concentrated peaks. Besides, something that immediately stands out from this plot, compared with the others shown in Fig. 9, is the shallow variation in the detection significance across the entire part of this plane. At the horizontal region around $V_{gal} \sim 150\,\mathrm{km\,s^{-1}}$, we see that the significance is the lowest, as we expected, but this is a very subtle effect. This lack of sensitivity indicates that, although they play a role, these two parameters (and specially V_{gal}) have little effect on the detectability of the UFDGs.

4.1.3 The L_V versus D plane

The bottom panel of Fig. 9 shows the results of the experiment where luminosity and distance were varied between 3×10^2 and $5 \times 10^4\,L_\odot$, and between 10 and 250 kpc, respectively. There are 509 UFDGs in this test. Their mass-to-light ratio M/L is between 20 and 4×10^3. Their observable number of stars N_{obs} varies between 2 and 6000. Note that here we consider a lower value for the minimum N_{obs} than 10 as indicated in Table 1 to sample in detail the detection limit. We find, however, that the minimum number of N_{obs} that gives a positive detection is 5 for this particular example.

In this test, there are several competing effects. For a fixed luminosity, as we increase the distance, the size in the sky and proper motion planes decreases, which favours identification, but on the other hand, the number of visible stars also decreases, which makes identification harder. The first effect scales as $\propto 1/D$, while the second, being an individual star luminosity problem, scales as $\propto 1/D^2$. So, at large distances the latter dominates and we lose the UFDGs, as seen in this panel. For instance, UFDGs with luminosities around $10^4\,L_\odot$ are not detected beyond ~100 kpc. Also, given a fixed distance, more luminous objects are detected with higher significance.

There is an interesting modulation in the colour contours in this panel. There are three leftward indentations of better detections at around 10, 25 and 60 kpc, which are better seen in the red and light-blue colours. These features are not statistical fluctuations, but the result of the effect of the *Gaia* observational errors in measured proper motions. As explained in Section 2.5 (Fig. 6), the size of the UFDGs in proper motion plane changes with distance in a peculiar way, presenting several minima at around the mentioned distances. At these distances, therefore, the UFDGs are slightly more concentrated in proper motion and, hence, easier to detect. The upper-left part of the panel, which does not contain any coloured squares or black crosses, is the region where systems do not have stars that can be observed by *Gaia*.

From this simple tests, one can see that some properties of the UFDGs are more relevant for the detections. In particular, the luminosity and distance, which set the number of observable stars together with the apparent size in the sky, seem to have a larger impact on the significance of our detections, than the size in proper motion space and the position of the peak with respect to the background.

4.2 The 'effective' parameter space

If we look at the essence of our problem devoid of its astronomical context, our task is to identify common peaks in two different planes, subject to a noisy and not necessarily uniform background. Seen as such, the key parameters upon which a successful detection depends are: the height of the peaks compared with the background level and the spread of the peaks, that is the apparent sizes of the UFDGs in the sky θ and proper motion planes σ_μ, and the number of observed stars that they are composed of $N_{\rm obs}$, with respect to the background. The probability will also depend on the projection of the centre of mass velocity in the proper motion plane, since this determines how close the UFDG peak appears to the centre of the proper motions distribution, where the majority of the background contaminants lie.

One can also see that for a certain UFDG that has been detected in the optimal scales (that is almost all stars in the UFDG are enclosed inside the joint peak detection), $N_{\rm com} \sim N_{\rm obs} + N_{\rm BG,\,in}$, where $N_{\rm BG,\,in}$ is the number of background stars that fall inside the joint peak detection. Assuming that the number of background stars inside the joint peak is similar to the expected one, that is $N_{\rm BG,\,in} \sim \langle N_{\rm com} \rangle$, the significance of equation (6) is equivalent to

$$s \sim \frac{N_{\rm obs}}{\sqrt{N_{\rm BG,in}}}. \tag{7}$$

As $N_{\rm BG,in} \propto\sim N_{\rm BG}\theta^2\sigma_\mu^2$, it follows from equation (7) that UFDGs that have the same ratio $N_{\rm obs}/\theta$ and all the rest of the parameters equal (including the LOS), have also approximately the same significance[13]. The same applies to UFDGs with the same ratio $N_{\rm obs}/\sigma_\mu$. For this reason, given a certain LOS (we deal with different LOS in Section 4.4) we can describe our detection problem based on these two quantities $N_{\rm obs}/\theta$ and $N_{\rm obs}/\sigma_\mu$, together with the position of the peak in proper motion space. We call these the 'effective parameters'. The latter has, however, less relevance compared to other properties, as already seen.

These quantities depend in turn on other astronomical parameters that characterize the system and its position with respect to the observer. But the successful detection of an UFDG depends only on a limited number of combinations of them, that is on the effective parameters. The importance of the effective parameters is that they reduce the dimensionality of the parameter space where we need to determine the boundaries of successful detection of our procedure. In particular, while we describe our UFDGs by using nine physical parameters, the detection of these systems depends only on three (and mainly two) effective parameters.

Here, we test that this concept is indeed correct. To evaluate the dependency of the detection limits and significance on these effective parameters, we have built a library of 2000 UFDGs with varying $N_{\rm obs}$, but keeping constant the apparent sizes in the sky and proper motion planes, as well as projected centre of mass velocity. Note that there is no straightforward way of generating UFDGs with the same real size in proper motion space σ_μ because of the effects of observational errors in proper motion (Section 2.5), and therefore we do it approximately by generating UFDGs with constant $\Delta\mu$, assuming that $\Delta\mu \sim \sigma_\mu$. Thus in this exercise, we vary the physical parameters $r_{\rm h}$, $\sigma_{\rm V}$ and D in a way that their combination (equations 2 and 3) result in constant θ and $\Delta\mu$. We also change $V_{\rm gal}$ with distance in order to obtain the same proper motion peak for these UFDGs.

[13] Note also that UFDGs with the same $N_{\rm obs}/\theta^2$ or $N_{\rm obs}/\sigma_\mu^2$ (apparent 'surface' density) do not have the same significance.

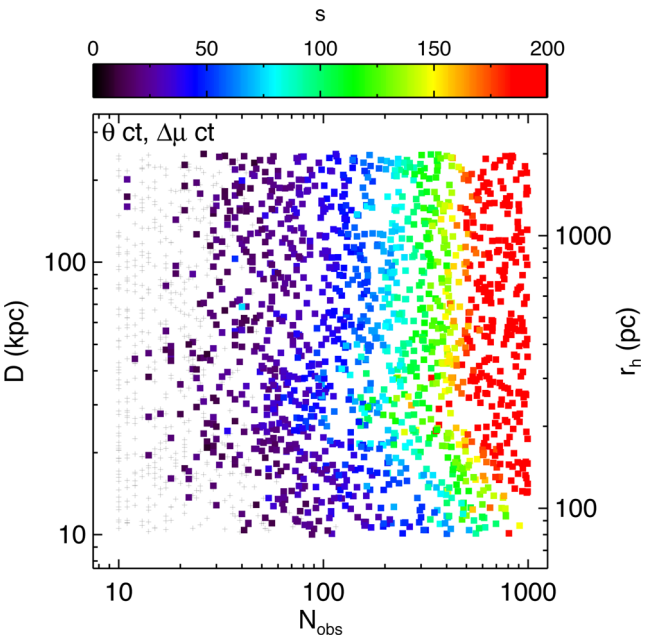

Figure 10. Detectability tests run with a library of synthetic UFDGs built with constant apparent size in the sky θ and proper motion plane σ_μ, as well as projected centre of mass velocity. Regions of constant effective parameters $N_{\rm obs}/\theta$ and $N_{\rm obs}/\sigma_\mu$ are vertical lines in this plot. The panel show the significance s as function of the physical parameters distance (left axis) and half-light radii (right axis) as a function of $N_{\rm obs}$. UFDGs with significance over 200 have been plotted with a colour saturated at this value. Black crosses indicate UFDGs that were not detected.

Fig. 10 shows the significance of the UFDGs of this experiment as function of $N_{\rm obs}$ and distance D (left axis). The right axis shows the half-light radii $r_{\rm h}$ which is related to D through equation (2) to produce the same θ. Although we do not include them in this plot, one could also draw other axes for $\sigma_{\rm V}$ and $V_{\rm gal}$, which are also related to D to produce the same $\Delta\mu$ and the same peak position in proper motion space. As θ and $\Delta\mu$ are constant for all these UFDGs, the effective parameters $N_{\rm obs}/\theta$ and $N_{\rm obs}/\sigma_\mu \sim N_{\rm obs}/\Delta\mu$ are constant along vertical lines in this plot. We see that equal significance contours are approximately vertical (but see discussion below), which illustrates that indeed, for constant effective parameters the significance does not depend on the physical parameters (D, $\sigma_{\rm V}$, $r_{\rm h}$) for these experiments (contrast this with the case of Fig. 9) but only on $N_{\rm obs}/\theta$ and $N_{\rm obs}/\sigma_\mu$.

However, the colours do not follow exactly vertical contours. This is due to the effects of observational errors in proper motion which make the apparent size of the UFDGs in the proper motion plane σ_μ oscillate with D as already explained in Section 2.5 (Fig. 6). This is the same effect as in Fig. 9 (bottom panel). Here, we generated these UFDGs with $\Delta\mu$ constant, but not σ_μ. In Fig. 10, we can see how UFDGs at around 30 kpc are better detected because at this distance the apparent size σ_μ decreases. Although there is a similar effect around 70 and 135 kpc, these are not so clearly seen here.

4.3 Limits of detection and completeness

Having seen that UFDGs with the same effective parameters have the same significance, we can explore the detection limits and completeness as function of them alone. In the following, we show the significance of a large synthetic library (15 000 UFDGs) that spans a large range in the effective parameters. All of the physical

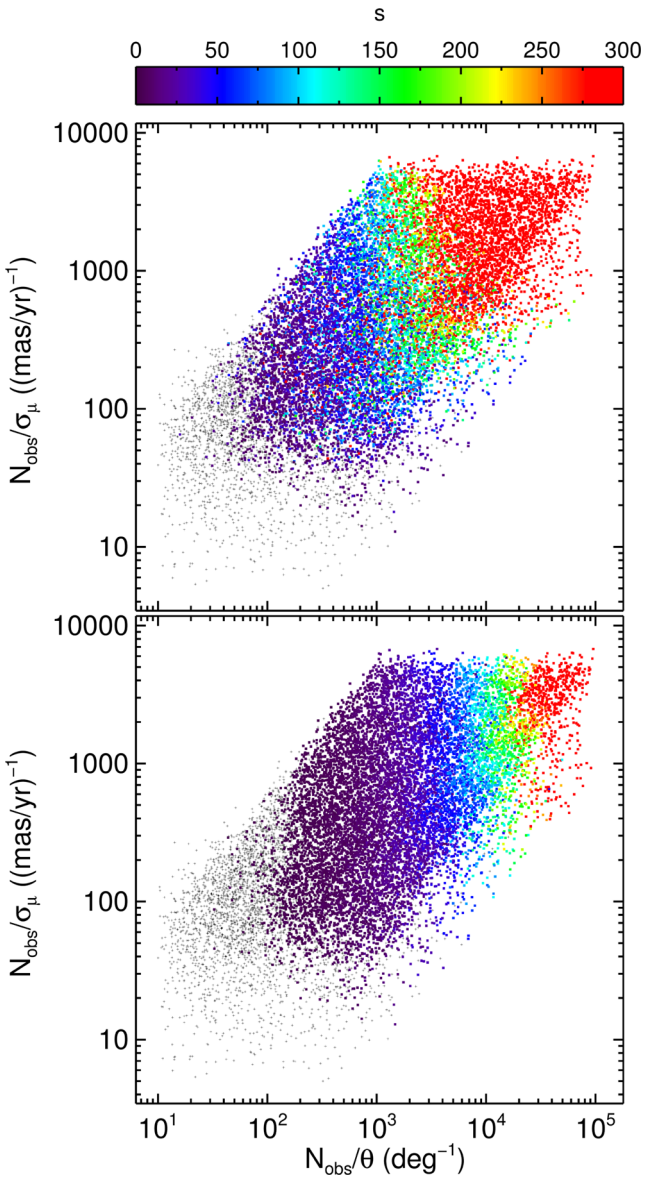

Figure 11. Top: detectability tests run with a large library of synthetic UFDGs as function of two effective parameters: $N_{\rm obs}/\theta$ and $N_{\rm obs}/\sigma_\mu$. The colour scale indicates the significance s. UFDGs with significance over 300 have been plotted with a colour saturated at this value. Black crosses indicate UFDGs that were not detected. Bottom: significance of the detections of the same library but being these only detections in the sky plane (see the text for details).

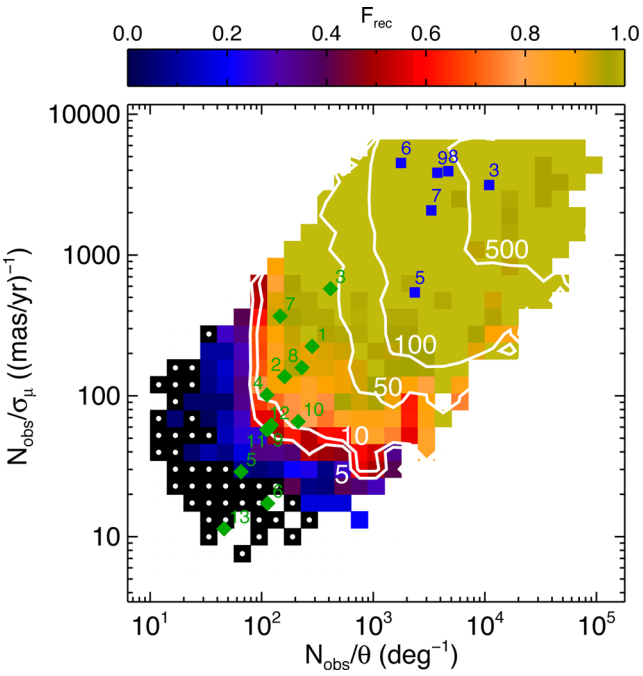

Figure 12. Fraction $F_{\rm rec}$ of detected UFDGs as function of effective parameters. Black squares with a central white dot are regions where the fraction recovered is exactly equal to 0 (this is to differentiate from regions with small recovered fraction). The blue squares and the green diamonds show the estimated positions of classical dSphs and UFDGs, respectively, with labels as in Fig. 2. White contours indicate the detection significance from Fig. 11 (top).

parameters of the UFDGs in this test are varied, except the LOS (l and b).

Fig. 11 (top panel) shows the results of this test as function of the two effective parameters $N_{\rm obs}/\theta$ and $N_{\rm obs}/\sigma_\mu$. The colours are well separated, i.e. not strongly mixed, in this plot, showing that despite the physical properties of the UFDGs being very different[14], the significance of the detections depends mainly of these two effective parameters. Though the third effective parameter (position of the peak in proper motion space) changes in the UFDGs of this test, the uniformity of the colours in a certain region of this plot indicates that its influence is not as relevant as that of the other parameters, as already shown.

More in detail, we can also see that for a higher fraction of the plot and specially the upper half, the colours follow a approximately vertical structure, i.e. the significance is mainly given by $N_{\rm obs}/\theta$. For the lower part, the contours are more curved. Finally, note also how the undetected objects lie in the regions of low $N_{\rm obs}/\theta$ and/or low $N_{\rm obs}/\sigma_\mu$, i.e. most diffuse objects.

We find that the minimum significance of our positive detections is $s \sim 5$. This is because of the threshold imposed to $\ln P$ in order to filter false detections (see Section 3.4).

In Fig. 12, we use the same test described above to estimate the fraction $F_{\rm rec}$ of detected UFDGs in each region of the effective parameter space. To do this, we have binned logarithmically this space and computed how many of the generated UFDGs in each bin are successfully detected. We only plot bins with at least four simulated UFDGs. The median number of UFDGs in each cell is 14. Note how for most of the space explored this fraction is close to 1 (ocher colours). There is a transition zone where fractions go from ∼0.5 to ∼0.2. To differentiate regions with small recovered fraction from regions with this fraction equal to 0 (all of them with dark colours), we have marked the latter with a white central dot. The region where our method is not able to detect objects is the low $N_{\rm obs}/\theta$ and/or low $N_{\rm obs}/\theta$, as expected.

In this plot, we superpose white contours indicating the significance s of the detections from Fig. 11 (top), computed as the median significance in the same grid used in this plot, including the cases that were not detected, that is with $s = 0$, in the computation. Note that these contours are just approximate, given that there is some dispersion in the significance. For instance, around the contours of $s = 5$ and 10, the dispersion in s is of 5–10. We see how the transition

[14] Remember that this is not a cross-section as in Fig. 9.

zone corresponds to the region of significance roughly around 5. For significance higher than 10, the recovered fractions is between 0.7 and 1.0.

We also plot in Fig. 12 the estimated values of effective parameters for the known MW satellites (classical dSph and UFDGs, blue squares and green diamonds, respectively). The values of σ_μ for these know systems have been estimated by interpolating in a plot similar to Fig. 6 (bottom) but for velocity dispersions σ_V between 5 and 10 km s^{-1}. As expected, all of the classical satellites of the MW lie in a region of effective parameters where our algorithm applied to the mock *Gaia* data successfully detects all simulated systems. Note that some of the classical satellites lie outside the higher limits of the plot. Out of the 13 UFDGs that would have observable stars by *Gaia* (Leo T would not be observable), 1 of them lies in a region where $F_{\rm rec}$ is 1.0 (Boo (3)) and 4 of them lie in regions with recovery fraction of 0.9 (CVnI (1), Her (2), UMaII (7), CmB (8)). Besides, WilI (10), UMa (4), SegII (12) and BooII (9) are recovered with fractions of 0.8, 0.7, 0.6, 0.5, respectively. SegI (11) and LeoIV (5) are in regions with $F_{\rm rec}$ of 0.3 and 0.1, respectively. Finally, CVnII (6) and LeoV (5) are outside the limits of detection. Note that in this plot we see how systems with effective parameters similar to the known UFDG would be detected by our algorithm. But some of the known UFDG have a small number of observables stars [e.g. LeoIV (5) and SegII (12) have $N_{\rm obs} \sim 5$ and $\sim$7, respectively] and our tests are done for a minimum of $N_{\rm obs} \sim 10$. Nevertheless, for these low-luminosity known cases, the estimated $N_{\rm obs}$ is very uncertain.

It is outstanding that our tests indicate that it is possible to detect with *Gaia* UFDGs similar to some of the ones detected by SDSS which is $\sim$2 mag deeper. This is because, whereas the UFDGs of SDSS were detected with photometry alone, our search is done also in the proper motion plane. The bottom panel of Fig. 11 is the same as the top panel, but plotting the significance that would correspond to these detections if the search had been made only in the sky plane, that is not including proper motion data in the detection algorithm. The significance s is now calculated through $s = (N_{\rm sky} - \langle N_{\rm sky}\rangle)/\sqrt{\langle N_{\rm sky}\rangle}$, where $N_{\rm sky}$ is the number of stars in the detected peak in the sky. We only plot the significance for the detections that had at least[15] $s = 3$. The remaining simulated UFDGs are plotted with black crosses. The colours follow now vertical contours as the vertical axis plays no role. Comparing this plot with the top panel, we see how much the significance increases when the proper motions are included in the search. Red colours (maximum significance) are only achieved for higher $N_{\rm obs}/\theta$, that is for more densely populated objects in the sky (right part). Also the limits of detection are now located at larger $N_{\rm obs}/\theta$.

Fig. 13 shows the surface brightness of the detected UFDGs as function of the effective parameters. Contours of similar surface brightness are approximately diagonal in this parameter space. We have marked with red the *detected* synthetic UFDGs with surface brightness dimmer than 30 mag arcsec^{-2} which is the global SDSS surface brightness limit as found in Koposov et al. (2008). Very interestingly, the red squares mark out an area in the parameter space of UFDGs less bright than the SDSS limit and that would be possible to explore with *Gaia*. Note, none the less, that this region has a recovery fraction $F_{\rm rec}$ smaller than 0.8.

In Fig. 13 (and in Fig. 12), all known satellites of the MW lie in an approximate diagonal line in the effective parameter space. We believe that this is a projection of the fundamental curve mass–

[15] Note that in the top panel the minimum s found was 5.

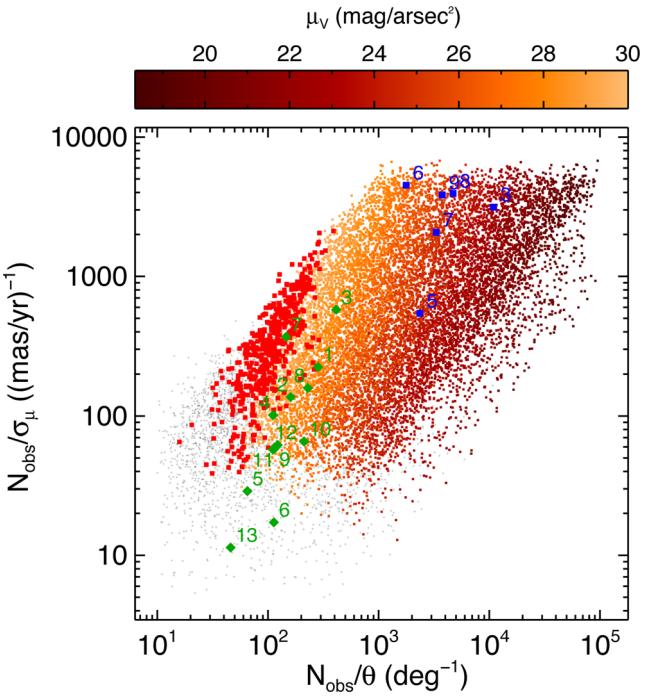

Figure 13. Surface brightness of the detected UFDGs as function of the effective parameters. Black dots indicate UFDGs that were not detected. Detected objects with surface brightness larger than 30 mag arcsec^{-2} are highlighted in red colours. The blue squares and the green diamonds show the estimated positions of classical dSphs and UFDGs, respectively, with labels as in Fig. 2.

radius–luminosity studied in e.g. Tollerud et al. (2011), or more in detail, a consequence of the Faber–Jackson and the $r_{\rm h}$-L scaling relations. The plot shows that our algorithm would be able to detect objects that are outside this diagonal. However, UFDGs that lie below the diagonal have surface brightness brighter than the SDSS limit and they would have already been detected, unless they are all located outside the SDSS footprint. On the other hand, part of the red region with surface brightness dimmer than the SDSS limit but that *Gaia* could probe lies outside the diagonal and, therefore, the detection of objects in it relies on the existence of them. Note, however, that the scatter across the diagonal is large.

Fig. 14 illustrates the recovered fraction (colour-scale) of UFDGs but now in terms of the physical parameters D, σ_V and $r_{\rm h}$ as a function of M_V. In these panels, the region with fractions between 0.9 and 1 (ocher colours) occupies a much smaller portion of the explored ranges. This is because, in these plots, in any given bin only two physical parameters are fixed while the remaining are varying in the entire explored range, which can result in a very different detection significance. This has the effect of lowering $F_{\rm rec}$ on average, while also resulting in more diffuse boundaries between areas with different $F_{\rm rec}$, as opposed to the sharp boundaries seen in the effective parameter plane of Fig. 12, which corroborates the fact that our detection scheme does depend mainly on the effective parameters. In this figure, we also show the positions of known UFDGs and classical dSph galaxies, though without the number labels. These are shown to illustrate the typical recovery fraction one would expect for a galaxy with, e.g. a given $r_{\rm h}$–M_V, if its velocity dispersion and other parameters are unknown but restricted to the range spanned by our library.

This implicit dependence on the other physical parameters means that the behaviour of $F_{\rm rec}$ in these plots will change depending on the

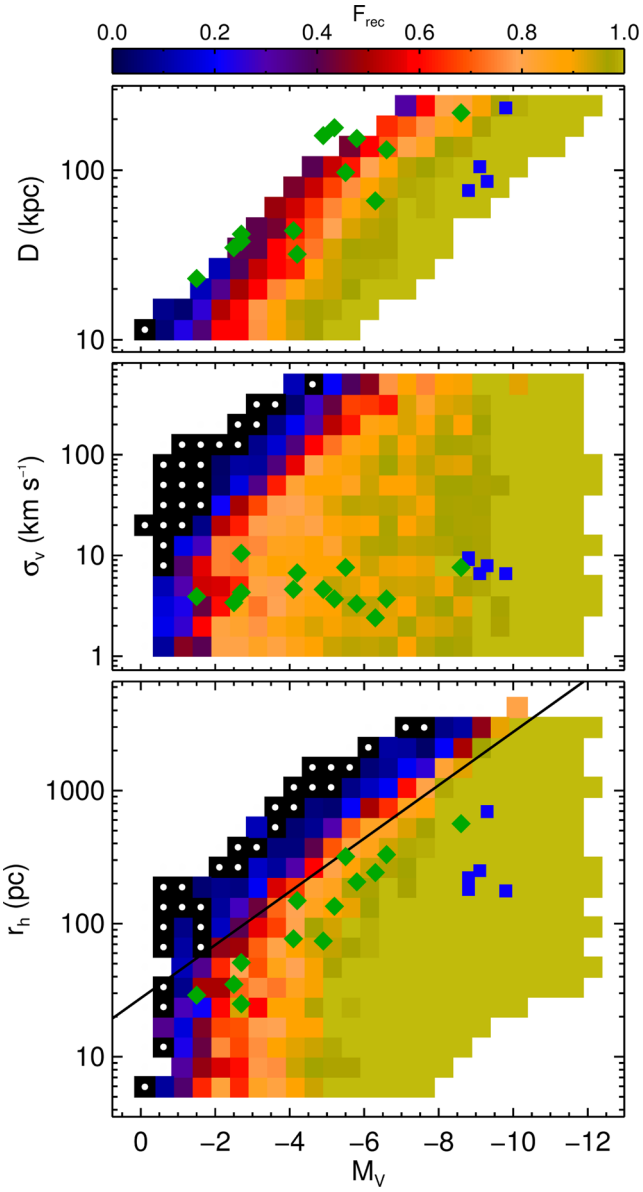

Figure 14. Fraction $F_{\rm rec}$ of detected UFDGs (colour-scale) as a function of D (top), $\sigma_{\rm V}$ (middle) and $r_{\rm h}$ (bottom) versus M_V. Black squares with a central white dot are regions where the fraction recovered is exactly equal to 0. The blue squares and the green diamonds show the estimated positions of classical dSphs and UFDGs, respectively. The black diagonal line in the bottom panel shows the SDSS surface brightness limit $\mu_V = 30$ mag arcsec^{-2}.

assumed distributions for the different parameters, and so, strictly speaking the reported $F_{\rm rec}$ is only valid under the assumed log-uniform distributions. For instance, if we consider only UFDGs with small velocity dispersion ($\sigma_{\rm V} < 10\,{\rm km\,s^{-1}}$; Fig. 15), the boundaries of detection improve significantly, i.e. the algorithm could detect larger UFDGs at the same given M_V. The simple distributions assumed for the physical parameters do allow us, however, to illustrate the limits of our method. Finally, it is worth noticing that the recovered fractions shown in Figs 12 and 14 can be interpreted in a probabilistic sense as the probability that any individual galaxy is detected by our method, given two of its physical or effective parameters.

The diagonal line in the bottom panel of Fig. 14 shows the SDSS surface brightness limit of μ_V =30 mag arcsec^{-2}. Lines of equal surface brightness are diagonal lines with slope of 5 in this plot. The detection limits of our algorithm (for example considering the line delineated by the red or blue coloured bins) have a similar slope at a slightly lower surface brightness with an additional vertical limit at $M_V \sim -1.5$ (but note that all these depends on the underlying distribution of physical parameters).

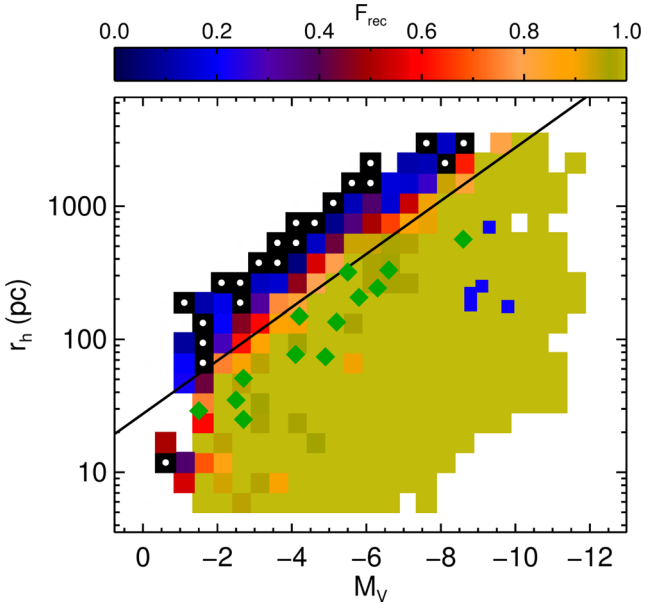

Figure 15. Same as low panel of Fig. 14 but only considering UFDGs with velocity dispersion $\sigma_{\rm V} < 10\,{\rm km\,s^{-1}}$.

The shape of the detection frontiers in the lower panel of Fig. 14 is similar to the ones of Koposov et al. (2008) in their figs 10 and 11. For brighter systems, the detection limits follows a diagonal line with the slope of a constant surface brightness line, followed by a vertical cut at certain absolute magnitude. In the case of Koposov et al. (2008), the surface brightness and absolute magnitude limits vary as function of distance. However, the detection limits of the two studies are not directly comparable because our method is based on different information, as it includes kinematics. Our effective parameter space, where the detection limits are defined, is essentially different (with more dimensions).

4.4 Changing the background

The results of the previous section correspond to our fiducial field $(l, b) = (90^\circ, 30^\circ)$. We now explore how these results change for different LOS's that have a different number of stars $N_{\rm BG}$ and a different distribution of proper motions.

Let s be the significance of a detection in the fiducial field, in which we have an UFDG with $N_{\rm obs}$ and a background with $N_{\rm BG}$ observed stars. Because $N_{\rm BG,\,in} \propto N_{\rm BG}$, from equation (7), it follows that $s \propto N_{\rm obs}/\sqrt{N_{\rm BG}}$. Then, given an UFDG with the same effective parameters but in an arbitrary LOS, for which the number of background stars $N'_{\rm BG}$ has changed in a ratio $r = N'_{\rm BG}/N_{\rm BG}$, its significance is

$$s' \sim \frac{s}{\sqrt{r}}. \tag{8}$$

This is a useful relation that allows us to establish the significance and detection limits in the effective parameter space for different LOS's without running additional experiments.

Table 4. Comparison of the significance of an ensemble of 1000 UFDGs located at different fields.

Field A		Field B		Expected	Median	AMD	P_{10}	P_{90}
l	b	l	b	s_B/s_A	s_B/s_A			
90	30	90	42	1.4	1.2	0.2	0.8	1.9
90	30	90	55	1.6	1.4	0.3	0.9	2.4
90	30	90	68	1.9	1.7	0.4	1.0	3.0
90	30	90	80	1.9	1.9	0.4	1.0	3.2
90	42	90	55	1.2	1.2	0.2	0.7	1.8
170	30	170	42	1.3	1.2	0.2	0.8	1.8
170	30	170	55	1.3	1.4	0.2	0.9	2.2
170	30	170	68	1.4	1.4	0.2	0.9	2.1
170	30	170	80	1.4	1.5	0.3	1.0	2.4
90	30	170	30	1.4	1.4	0.5	0.6	3.3
90	30	170	42	1.8	1.7	0.6	0.7	4.1
90	30	170	55	1.9	1.9	0.7	0.8	4.4
90	30	170	68	2.0	1.8	0.7	0.8	4.5
90	30	170	80	2.0	2.0	0.8	0.9	4.5

For instance, the number of background stars in the fiducial field is $N_{BG} = 1413$ and for two different LOS's at $(l, b) = (90^\circ, 55^\circ)$ and $(l, b) = (90^\circ, 80^\circ)$ this is $N_{BG} = 525$ and $N_{BG} = 377$, respectively. Therefore the background has decreased by factors $r = 0.37$ and 0.27, respectively, with respect to the fiducial case. Thus, we expect the significance of UFDGs with the same effective parameters to increase by $s' = 1.6s$ and $s' = 1.9s$, respectively.

In the following, we check that this relation is correct. We use a library of 1000 UFDGs (randomly extracted from the library of Section 4.3), and locate copies of it in different LOS's. We then compare the one-by-one significance for different pairs of LOS's. Note, however, that because we keep the proper motion of each UFDG constant, its relative position with respect to the centroid will change depending on the LOS (because the underlying distribution changes), thus changing one of the effective parameters. For this reason, and also because of the approximations used to derive equation (8) and that *Gaia* errors change with LOS, we expect a certain dispersion around the values predicted by equation (8).

Table 4 compares the expected value of s_B/s_A with the median observed values computed with all 1000 UFDGs at different pairs of LOS A and B. We also give the Absolute Median Deviation (AMD), and the 10 per cent (P_{10}) and 90 per cent (P_{90}) percentiles. The median ratios s_B/s_A differ at most 0.2 from the expected values. We see also that there is some expected dispersion with respect to this value. The cases where more dispersion is observed are when we compare fields at different longitudes (last rows). This is because in these cases the distribution of background proper motion changes the most. However, by looking at the percentiles we see that most of the dispersion comes from values that are higher than the expected value (thus improving the significance). The P_{10} is always around 1. This means that the significance of all the fields B is smaller than predicted, but larger than the significance of the fields A in ∼40 per cent of the cases. For ∼50 per cent of the cases, the significance of all the fields B is larger than expected.

In conclusion, the significance of the detections of our fiducial field at $l = 90^\circ$ and $b = 30^\circ$ are maintained or improved in ∼90 per cent of the simulated cases in the other fields where N_{BG} was smaller. The boundaries of the detection will also improve for these fields. But the scaling between significance and fraction of recovery is not straightforward and one would need to evaluate this in each particular field. Equation (8) offers, however, a fast approximate way of comparing the success of the detections in different LOS's.

5 CAVEATS

The method that we have introduced here has its limitations and assumptions, which we will review here.

First of all, we emphasize that our method does not aim to characterize and study UFDGs but it is a probabilistic method to identify possible candidates. When applied to real data, it will provide us with a list of candidates that will need to be studied in detail. The colour–magnitude diagrams of *Gaia* photometry can be used for this, as well as to derive morphological properties, kinematics, distances, etc, once a proper filter to select the UFDG population is designed, as has been done with SDSS (Willman et al. 2002). Also, a follow up using ground-based facilities will be required to obtain radial velocities and detailed chemical abundances.

Our procedure has been tested for Galactic latitudes above 30° in fields of $2^\circ \times 2^\circ$ and has been designed for UFDGs at distances larger than 10 kpc. If we want to apply it to search for nearer systems, different parallax and log g cuts would be needed. However, larger scales in the sky would have to be probed, increasing the background level. Another limitation of the method is that it is optimized for UFDGs with angular sizes smaller than the $2^\circ \times 2^\circ$ fields, and the detection of larger systems would require, again, to probe larger scales in the sky and perhaps a different strategy.

Likewise, it has been tested for UFDGs modelled as Plummer spheres with isotropic velocity distributions where light follows mass. A change in these assumptions that results on a variation in the footprints in the sky or proper motion planes, will change the effectiveness of our method, although the limits we have encountered should remain the same, when expressed in terms of the effective parameters. It is the mapping from structural to effective parameters that would need to be established for the new UFDG models. Similarly, the boundaries that define the limits of our detection method in N_{obs} will simply map into different boundaries in stellar luminosity, if the stellar population content is changed.

Our background clearly comes from smooth distributions without streams or clouds. A clumpy halo may affect the number of background stars compared to our estimations with GUMS. However, our algorithm will also detect other systems that are not necessarily UFDGs as long as they present some coherence in the four-dimensional space that we use. These detections, rather than being considered as additional false positives, will be interesting systems to be followed-up.

Previous studies to detect UFDGs with photometric surveys apply an isocrhone masking or a probabilistic modelling in the colour–magnitude or colour–colour diagram in order to filter out field stars (e.g. Koposov et al. 2008). Instead, here we do cuts using parallax and surface gravity. Some preliminary tests show that the addition of an isocrhone masking in the *Gaia* G versus G_{BP}–G_{RP} plane in our algorithm may be beneficial in particular cases. This merits a separate investigation that we aim to undertake in the future.

Although we have not included unresolved galaxies and quasars in our simulated background, we have checked that these will have a minor effect in our results. According to Bailer-Jones et al. (2013, their table 3) the fraction of misclassified galaxies and quasars is 2.5 per cent (2 per cent misclassified as stars and 0.5 per cent as binary systems) and 8.9 per cent (5.9 per cent as stars and 0.1 per cent as binary systems), respectively. From GUMS simulations, we have estimated that the number of galaxies and quasars in our fiducial field would be 4051 and 120, respectively. Therefore, there will be

∼110 objects (mainly galaxies) misclassified as stars, corresponding to an increase of 7 per cent of the back/foreground population in our fiducial field, and up to ∼30 per cent for other LOS ($(l, b) = (180^\circ, 80^\circ)$). However, this increase in the number of field stars will imply approximately a change in the significance of the detections only of a factor 0.96 and 0.88 (equation 8), respectively, in the two LOS described.

One important aspect of the *Gaia* astrometric data that we have not taken into account in this work is the issue of covariances in the estimated astrometric parameters. As explained in Lindegren et al. (2012), the statistical correlation between the different astrometric parameters will occur between the parameters of the same source and also between the parameters of different sources. The within-source error covariances can be similar for collections of sources in small areas of the sky, as can be seen for example in the statistical plots in Volume 1 to the Hipparcos Catalogue documentation (ESA 1997). In the proper motion plane for small areas on the sky (such as used in this study) this can lead to apparent structure in the proper motion distribution (caused by elongated and preferentially oriented error-ellipses).The between-source covariances will have a similar effect and are estimated in the case of *Gaia* to be most pronounced over areas of the order of $0.^\circ3$ radius on the sky (the value of the correlation half-length estimated in Holl, Hobbs & Lindegren 2010). This means that for a large fraction of UFDGs the between-source correlations will be important in addition to the within-source correlations. To first order the main effect will be that the interpretation of the WT maps will be more involved, where a distinction will have to be made between real and spurious structure in the proper motion plane.

The within-source covariance matrix of the astrometric parameters will be provided as part of the *Gaia* data releases. The covariance matrix of the astrometric parameters of different sources cannot be calculated for the full *Gaia* catalogue but it is feasible to do so for limited groups of sources as demonstrated in Holl & Lindegren (2012) and Holl, Lindegren & Hobbs (2012). Hence, we will be able to account for the error covariances but we defer to a future study the details of how to implement this in practice.

6 DISCUSSION AND CONCLUSIONS

We have introduced an automatic procedure to identify UFDG candidates in the future *Gaia* data base and charted its detection limits. The main advantages of using *Gaia* data on the search of UFDGs are, first, the inclusion of kinematics (proper motions) in the detection algorithm for the first time; and secondly, the *Gaia* full sky coverage, being the first unbiased homogeneous survey to be used for this purpose.

Our procedure identifies significant overdense peaks in the planes of the sky and of proper motions that share common stars. Then the probability of this occurring by chance is assessed and used to discard spurious detections. We have used a library of ∼30 000 synthetic UFDGs to probe the 9D space of intrinsic (L_V, r_h, σ_V) and extrinsic (l, b, D, V_gal, ϕ_V, θ_V) UFDG parameters, spanning ranges that extend well beyond those occupied by currently known systems.

We have identified the 'effective parameters' that our algorithm depends mainly on. The main two are the ratios of the number of observable stars by *Gaia* in the UFDGs to their apparent sizes in the sky (N_obs/θ) and proper motion planes ($N_\mathrm{obs}/\sigma_\mu$). The position of the peak in proper motion with respect to the background also influences the detection, but is not as relevant. These parameters reduce the dimensionality of our problem to 3, mainly 2, parameters.

We have charted the limits of detectability and completeness (recovery fraction) of our search in the effective parameter space (Fig. 12) for a LOS at $l = 90^\circ$ and $b = 30^\circ$. Detections can be made with high significance over most of the explored region, which includes the majority of the currently known UFDGs, with a recovered fraction that remains above 70 per cent over most of it. It is only in the corner of small effective parameters that the efficacy of our method decreases abruptly. On the other hand, the limits of our detection procedure cannot be described in terms of a limiting surface brightness alone (Fig. 13), because of the inclusion of kinematics in the search.

We have derived a relation that allows us to know the approximate detection significance of the synthetic UFDGs at LOSs with a different number of background stars. The translation from significance to recovery fraction is not straightforward and one would need a more thorough characterization per LOS. However, most of the results presented here are for a pessimistic case compared to higher latitudes, or to the outer galaxy ($l = 180^\circ$), where we expect less field contamination.

Furthermore, we have explored the extent to which current detectability limits can be pushed forward, opening the possibility of detecting real systems hitherto not found. We have found that there is a region in the effective parameter space where there are currently no observed systems. Part of this region corresponds to UFDGs with surface brightness brighter than the SDSS limit and, therefore, they would have already been detected, unless they are all located outside of the SDSS footprint. But more interestingly, we have seen that *Gaia* will be able to probe a region of the effective parameter space of surface brightness dimmer than the SDSS limit, if such objects exist, albeit with a recovery fraction smaller than 0.8. Note that the recent UFDG discoveries made with DECam have similar surface brightness to the ones detected by SDSS (see fig. 17 in Koposov et al. 2015). Also because of the different detection methodologies followed by SDSS and DECam compared to *Gaia*, the nature of the detection limits is completely different, thus offering the possibility to explore uncovered regions of the parameter space (both with respect other surveys in the north and in the south) and for all sky.

We can make a very rough estimation of the number of UFDGs that *Gaia* will detect from the recovery fractions that we have found for our synthetic search (Fig. 12), assuming isotropy on the distribution of satellites in the MW halo, and considering only SDSS UFDGs. There is 1 known object (Boo) that would be detected with a recovery fraction of 1.0, 4 objects (CVnI, Her, UMaII, CmB) with a fraction of 0.9, and four objects (WilI, UMa, SegII, BooII) with fractions of 0.8, 0.7, 0.6 and 0.5, respectively. We do not count objects with a recovery fraction below 0.5. This makes a total of 7.2 UFDGs in a sky area equivalent to SDSS (∼1/5 of the sky; Koposov et al. 2008). If we assume that *Gaia* will detect UFDGs only above $b = 30^\circ$, which corresponds to 1/2 of the sky, there should be of the order of ∼10 new UFDGs (i.e. currently not known) over the 1/2−1/5=3/10 of the sky that remains unexplored, that is subtracting the area already covered by SDSS. These calculations are based on the field at $l = 90^\circ$ and $b = 30^\circ$ but could be slightly better for higher latitudes.

However, by the arrival time of the *Gaia* catalogue (see below) other surveys such as ATLAS (Shanks et al. 2013), Pan-STARRS (Kaiser et al. 2010) and DES (Diehl et al. 2014) will have covered great fraction of this area. But a part of the South Galactic cap will still remain completely unexplored (at declinations below stripe SPT of DES). We estimate this to be a fraction of ∼0.0195 of the whole sphere (by taking the part of the spherical cap in equatorial coordinates below $\delta < -65^\circ$ that lies in the range $\alpha \sim [-60, 90]$).

Therefore, there should be of the order of ∼1 new UFDG in this unexplored area. However, we emphasize that our method uses information not used in other searches, namely proper motions, and thus, it could lead to new discoveries, made possible, not by the covered region in the sky or depth of probing, but by their motion in the sky. As such, our method complements present searches.

Moreover, the number of discovered new candidates could be higher because as discussed above, *Gaia* could also detect more UFDGs with lower surface brightness than the SDSS limit. Besides, under the assumption of anisotropy in the spatial distribution of satellites, this number could be larger if the *Gaia* footprint happens to cover preferential directions. In fact, the importance of having a full-sky catalogue in this type of search for the first time is that it will allow us to put constraints on the isotropic distribution of the satellites and, therefore, their origin.

The known UFDGs with high recovery fraction mentioned above could be seen as standard systems for future *Gaia* discoveries but only in terms of effective parameters. Thus, one cannot interpret this as if, for instance, all objects with the same half-light radius and the same distance as Boo will be detected, but rather as systems for which the combination of *all* physical parameters produce similar effective parameters will be detected with high probability. Note also that we have not considered in this calculation the influence of the third effective parameter, which we have shown to be less important.

Our proposed method can be applied fully to the third *Gaia* data release scheduled for[16] 2017/2018. This release will include the five-parameter astrometric solutions as well as the object classification (necessary to eliminate contaminant extra-Galactic objects) and astrophysical parameters such as $\log g$, necessary for filtering out foreground dwarfs. Preliminary searches could be conducted using earlier releases; e.g. with the first data release in summer 2016, using only on sky coordinates; or with the second data release in early 2017, using full sky and proper motion information, yet without the possibility of using the foreground filters as explained here, since astrophysical parameters will not yet be available.

Finally, there is the future possibility that the *Gaia* magnitude limit will be pushed down to $G = 20.7$. This will obviously be positive in terms of the number of observable stars in each UFDG, but will also increase the foreground/background contamination, so the effect in the detection probabilities will have to be assessed.

ACKNOWLEDGEMENTS

We thank V. Belokurov, the referee, for a thorough review and insightful comments on the original manuscript. TA is supported by an ESA Research Fellowship in Space Science. CM acknowledges support from a post-doctoral fellowship of DGAPA-UNAM, Mexico. CM and LA acknowledge support from DGAPA/UNAM grant IG100115 and the hospitality and support of the University of Barcelona during part of this investigation. FF acknowledge support from MINECO (Spanish Ministry of Economy) - FEDER through grant AYA2012-39551-C02-01 and ESP2013-48318-C2-1-R, and from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement GREAT-ITN FP7 264895. We thank the CEA and Nice Observatory for the MR software. We thank Merce Romero-Gomez and GaiaUB team for the code to simulate *Gaia* errors.

[16] http://www.cosmos.esa.int/web/gaia/release

REFERENCES

Anscombe F. J., 1948, Biometrika, 35, 246,
Antoja T., Figueras F., Fernández D., Torra J., 2008, A&A, 490, 135
Antoja T. et al., 2012, MNRAS, 426, L1
Arenou F., 2011, in Docobo J. A., Tamazian V. S., Balega Y. Y., eds, AIP Conf. Proc. Vol. 1346, International Workshop on Double and Multiple Stars: Dynamics, Physics, and Instrumentation. Am. Inst. Phys., New York, p. 107
Bailer-Jones C. A. L. et al., 2013, A&A, 559, A74
Bechtol K. et al., 2015, ApJ, 807, 50
Behroozi P. S., Wechsler R. H., Wu H.-Y., 2013, ApJ, 762, 109
Belokurov V., 2013, New Astron. Rev., 57, 100
Belokurov V. et al., 2006, ApJ, 647, L111
Belokurov V. et al., 2007, ApJ, 654, 897
Belokurov V. et al., 2008, ApJ, 686, L83
Belokurov V. et al., 2009, MNRAS, 397, 1748
Belokurov V. et al., 2010, ApJ, 712, L103
Brown A. G. A., Velázquez H. M., Aguilar L. A., 2005, MNRAS, 359, 1287
Brown T. M. et al., 2014, ApJ, 796, 91
Bullock J. S., 2010, preprint (arXiv:1009.4505)
Chabrier G., 2003, PASP, 115, 763
Close L. M., Siegler N., Freed M., Biller B., 2003, ApJ, 587, 407
de Bruijne J. H. J., 2012, Ap&SS, 341, 31
de Bruijne J. H. J., Allen M., Azaz S., Krone-Martins A., Prod'homme T., Hestroffer D., 2015a, A&A, 576, 74
de Bruijne J. H. J., Rygl K. L. J., Antoja T., 2015b, preprint (arXiv:1502.00791)
de Souza R. E., Krone-Martins A., dos Anjos S., Ducourant C., Teixeira R., 2014, A&A, 568, A124
Diehl H. T. et al., 2014, in Peck A. B., Benn C. R., Seaman R. L., eds, Proc. SPIE Conf. Ser. Vol. 9149, Observatory Operations: Strategies, Processes, and Systems V. SPIE, Bellingham, p. 0
Duquennoy A., Mayor M., 1991, A&A, 248, 485
Esa, ed., 1997, ESA SP-1200: The Hipparcos and Tycho catalogues. ESA, Noordwijk
Forbes D. A., Kroupa P., 2011, PASA, 28, 77
Hernández-Pérez F., Bruzual G., 2013, MNRAS, 431, 2612
Holl B., Lindegren L., 2012, A&A, 543, A14
Holl B., Hobbs D., Lindegren L., 2010, in Klioner S. A., Seidelmann P. K., Soffel M. H., eds, Proc. IAU Symp. 261, Relativity in Fundamental Astronomy: Dynamics, Reference Frames, and Data Analysis. Kluwer, Dordrecht, p. 320
Holl B., Lindegren L., Hobbs D., 2012, A&A, 543, A15
Hurley J. R., Tout C. A., Pols O. R., 2002, MNRAS, 329, 897
Jordi C. et al., 2010, A&A, 523, A48
Kaiser N. et al., 2010, in Stepp L. M., Gilmozzi R., Hall H. J., eds, Proc. SPIE Conf. Ser. Vol. 7733, Ground-Based and Airborne Telescopes III. SPIE, Bellingham, p. 0
Kirby E. N., Simon J. D., Geha M., Guhathakurta P., Frebel A., 2008, ApJ, 685, L43
Klypin A., Kravtsov A. V., Valenzuela O., Prada F., 1999, ApJ, 522, 82
Koposov S. et al., 2008, ApJ, 686, 279
Koposov S. E., Belokurov V., Torrealba G., Wyn Evans N., 2015, ApJ, 805, 130
Lada C. J., 2006, ApJ, 640, L63
Laevens B. P. M. et al., 2015, ApJ, 802, L18
Lindegren L., Lammers U., Hobbs D., O'Mullane W., Bastian U., Hernández J., 2012, A&A, 538, A78
McConnachie A. W., 2012, AJ, 144, 4
Mateu C., Bruzual G., Aguilar L., Brown A. G. A., Valenzuela O., Carigi L., Velázquez H., Hernández F., 2011, MNRAS, 415, 214
Mignard F. et al., 2008, in Jin W. J., Platais I., Perryman M. A. C., eds, Proc. IAU Symp. 248, A Giant Step: from Milli- to Micro-arcsecond Astrometry. Kluwer, Dordrecht, p. 224
Moore B., Ghigna S., Governato F., Lake G., Quinn T., Stadel J., Tozzi P., 1999, ApJ, 524, L19

Onions J. et al., 2012, MNRAS, 423, 1200
Perryman M. A. C. et al., 2001, A&A, 369, 339
Press W. H., Schechter P., 1974, ApJ, 187, 425
Robin A. C. et al., 2012, A&A, 543, A100
Romero-Gómez M., Figueras F., Antoja T., Abedi H., Aguilar L., 2015, MNRAS, 447, 218
Shanks T. et al., 2013, The Messenger, 154, 38
Simon J. D., Geha M., 2007, ApJ, 670, 313
Springel V., Frenk C. S., White S. D. M., 2006, Nature, 440, 1137
Starck J.-L., Murtagh F., eds, 2002, Astronomical Image and Data Analysis. Springer, Berlin
Tollerud E. J., Bullock J. S., Graves G. J., Wolf J., 2011, ApJ, 726, 108
White S. D. M., Rees M. J., 1978, MNRAS, 183, 341
Willman B., Dalcanton J., Ivezić Ž., Jackson T., Lupton R., Brinkmann J., Hennessy G., Hindsley R., 2002, AJ, 123, 848
Willman B. et al., 2005a, AJ, 129, 2692
Willman B. et al., 2005b, ApJ, 626, L85
Zucker D. B. et al., 2006a, ApJ, 643, L103
Zucker D. B. et al., 2006b, ApJ, 650, L41

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