

# The outer disc in shambles: Blind detection of Monoceros and the ACS with *Gaia*'s astrometric sample<sup>★</sup>

P. Ramos<sup>1,2</sup>, T. Antoja<sup>2</sup>, C. Mateu<sup>3</sup>, F. Anders<sup>2</sup>, C. F. P. Laporte<sup>4</sup>, J. A. Carballo-Bello<sup>5</sup>, B. Famaey<sup>1</sup>, and R. Ibata<sup>1</sup>

<sup>1</sup> Observatoire Astronomique de Strasbourg, Université de Strasbourg, CNRS, 11 rue de l'Université, 67000 Strasbourg, France  
e-mail: [p.ramos@unistra.fr](mailto:p.ramos@unistra.fr)

<sup>2</sup> Dept. FQA, Institut de Ciències del Cosmos (ICCUB), Universitat de Barcelona (IEEC-UB), Martí Franquès 1, 08028 Barcelona, Spain

<sup>3</sup> Departamento de Astronomía, Instituto de Física, Universidad de la República, Iguá 4225, CP 11400 Montevideo, Uruguay

<sup>4</sup> Kavli Institute for the Physics and Mathematics of the Universe (WPI), The University of Tokyo Institutes for Advanced Study (UTIAS), The University of Tokyo, Chiba 277-8583, Japan

<sup>5</sup> Instituto de Alta Investigación, Universidad de Tarapacá, Casilla 7D, Arica, Chile

Received 2 November 2020 / Accepted 2 December 2020

## ABSTRACT

**Context.** The *Gaia* astrometric sample allows us to study the outermost Galactic disc, the halo, and their interface. It is precisely at the very edge of the disc where the effects of external perturbations are expected to be the most noticeable.

**Aims.** Our goal is to detect the kinematic substructure present in the halo and at the edge of the Milky Way (MW) disc and provide observational constraints on their phase-space distribution.

**Methods.** We download, one HEALpix at a time, the proper motion histogram of distant stars, to which we apply a wavelet transformation to reveal the significant overdensities. We then analyse the large coherent structures that appear in the sky.

**Results.** We reveal a sharp yet complex anticentre dominated by Monoceros (MNC) and the Anticentre Stream (ACS) in the north – which we find have intensities comparable to the Magellanic Clouds and the Sagittarius stream – and by MNC South and TriAnd at negative latitudes. Our method allows us to perform a morphological analysis of MNC and the ACS, both of which span more than 100° in longitude, and to provide a high purity sample of giants with which we track MNC down to latitudes as low as ~5°. Their colour-magnitude diagram is consistent with extended structures at a distance of ~10–11 kpc that originated in the disc, with a very low ratio of RR Lyrae over M giants, and with kinematics compatible with the rotation curve at those distances or slightly slower.

**Conclusions.** We present a precise characterisation of MNC and the ACS, two previously known structures that our method reveals naturally, allowing us to detect them without limiting ourselves to a particular stellar type and, for the first time, using only kinematics. Our results will allow future studies to model their chemo-dynamics and evolution, thus constraining some of the most influential processes that shaped the MW.

**Key words.** Galaxy: kinematics and dynamics – Galaxy: formation – Galaxy: halo – astrometry

## 1. Introduction

Most of the studies that discovered new substructures within the second data release (DR2, [Gaia Collaboration 2018a](#)) of the *Gaia* mission ([Gaia Collaboration 2016](#)) used the full 6D phase-space sample, and their impact on our current understanding of the Milky Way (MW) and its history is undeniable. Examples of such studies are the works of [Belokurov et al. \(2018\)](#), [Gaia Collaboration \(2018b\)](#), [Haywood et al. \(2018\)](#), and [Helmi et al. \(2018\)](#), who identified a large group of stars that were accreted in the last major merger event of the MW, which took place ~10 Gyr ago. Another example is the advance in the study of the ‘moving groups’ and the possibility to now visualise the kinematic substructure directly in the plane of galactocentric radii against rotational velocity with the ridges ([Antoja et al. 2018](#); [Kawata et al. 2018](#); [Ramos et al. 2018](#); [Laporte et al. 2019a](#); [Fragkoudi et al. 2019](#); [Khanna et al. 2019](#)). Nevertheless, this sample is limited to approximately  $G < 13$  mag (above that, the

completeness drops significantly), restricting the exploitation of the kinematic data to a volume of ~3 kpc radius from the Sun. Despite some attempts to extend the kinematic maps to farther distances either by using statistical corrections to the parallax ([López-Corredoira & Sylos Labini 2019](#); [López-Corredoira et al. 2020](#)) or by adding photometric ([Anders et al. 2019](#)) or spectroscopic information ([Liu et al. 2017](#); [Wang et al. 2019](#)), these maps only have a significant amount of stars up to galactocentric radii of ~16 kpc.

In contrast, the 5D sample (only astrometry and no radial velocity) is more than two orders of magnitude larger than the 6D one. Its power is exemplified by, for instance, the work of [Castro-Ginard et al. \(2018, 2020\)](#), who used it to discover hundreds of new open clusters throughout the disc, the work of [Malhan et al. \(2018\)](#) and [Ibata et al. \(2019\)](#), who used it to reveal several new tidal streams in the halo, or the work of [Koppelman & Helmi \(2021\)](#), who used it to provide a large sample of halo stars selected using a combination of photometry and proper motions. Another good example is the detection of the Sagittarius (Sgr, [Ibata et al. 1994](#)) stream using mostly its kinematic signature ([Antoja et al. 2020](#); [Ibata et al. 2020](#); [Ramos et al. 2020](#)).

<sup>★</sup> Full Tables 1 and 2 are only available at the CDS via anonymous ftp to [cdsarc.u-strasbg.fr](ftp://cdsarc.u-strasbg.fr) (130.79.128.5) or via <http://cdsarc.u-strasbg.fr/viz-bin/cat/J/A+A/646/A99>

The astrometric sample reaches down to  $G \sim 21$  mag, significantly expanding the volume probed<sup>1</sup>, meaning that we can use it to trace kinematic structures well into the halo. Among the different stellar systems that we expect to find within the 5D sample are globular clusters (e.g., Baumgardt et al. 2019), streams (e.g., Belokurov et al. 2006), dwarf galaxies, and even ultra faint dwarf galaxies (e.g., Willman et al. 2005; Belokurov et al. 2007; Koposov et al. 2015). Moreover, this sample also covers the outermost regions of the MW disc, where Newberg et al. (2002) reported, almost two decades ago, the presence of a peculiar population above the mid-plane of the Galaxy, bluer than the thick disc and clearly appreciable as an overdensity of main sequence turn-off (MSTO) stars at a distance of  $\sim 10$  kpc from the Sun. Known as Monoceros (MNC, also referred to as the Galactic anticentre stellar structure, or GASS), this structure was observed to span more than a hundred degrees in longitude ( $100^\circ < l < 270^\circ$ , see, e.g., Rocha-Pinto et al. 2004; Morganson et al. 2016) both in the northern and southern hemispheres. Over the past two decades, there has been an intense debate over its origin, in part due to the difficulties of confronting the data with the different models available (Slater et al. 2014). Out of the many possible mechanisms proposed by Ibata et al. (2003), there have been two leading hypotheses: accretion and disc perturbation.

The idea that MNC is the tidal debris of an accreted satellite was based on its morphology (it looks like a stream) and on its metallicity and kinematics (e.g., Yanny et al. 2003; Crane et al. 2003; Wilhelm et al. 2005; Conn et al. 2005), and this idea is partially supported by the simulations of Helmi et al. (2003) and Peñarrubia et al. (2005). This led to the hunt for its progenitor and, after discarding the Canis Major (Martin et al. 2004) overdensity as a candidate (e.g., Momany et al. 2006; Rocha-Pinto et al. 2006; Carballo-Bello et al. 2021), none has yet been found despite attempts to detect the continuation of the hypothetical tidal stream at other Galactic latitudes (Conn et al. 2007, 2008). Nevertheless, upper limits for the total mass of the progenitor have been calculated (e.g., Guglielmo et al. 2018).

On the other hand, several works have shown that the close passage of a satellite can induce significant substructure in the outer parts of the disc (e.g., Younger et al. 2008; Purcell et al. 2011; Gómez et al. 2016). The interaction with a dwarf galaxy as massive as Sgr could cause some of the disc material to move to more inclined and eccentric orbits, and it could produce a stream of stars consistent with the observations (see also Kazantzidis et al. 2008, where, instead of a single satellite, the perturbers are six dark matter subhalos of masses  $\sim 10^{10} M_\odot$ ). More recently, the simulations by Laporte et al. (2018, 2019b) have shown that it is possible to create extended structures similar to MNC as well as feather-like structures during a satellite encounter while at the same time qualitatively reproducing part of the phase-space substructure observed in the solar neighbourhood. The detection of a vertical wave-like pattern in the disc (Widrow et al. 2012) that propagates almost radially (Xu et al. 2015; Schönrich & Dehnen 2018), in agreement with the simulations of Gómez et al. (2013), and the discovery of the phase-space spiral and consequent confirmation that our Galaxy is undergoing phase-mixing (Antoja et al. 2018) further support the perturbative scenario.

Other structures in the outer disc are the Anticentre Stream (ACS) and the Eastern Band Structure (EBS, Grillmair 2006), or the Triangulum-Andromeda (TriAnd1 and TriAnd2) overdensi-

ties (Majewski et al. 2004; Rocha-Pinto et al. 2004; Martin et al. 2007). The connection between all of these and MNC has also been subject to scrutiny for many years and is still not entirely clear (but see the models of Xia et al. 2015; Sheffield et al. 2018). For instance, the EBS, which was described as an independent structure by Grillmair (2011), has now been suggested to be part of the MNC ring by Deason et al. (2018), who used a combination of *Gaia* and the Sloan Digital Sky Survey (SDSS, York et al. 2000) data. de Boer et al. (2018) re-analysed the kinematics of MNC and ACS with SDSS astrometry calibrated with *Gaia* DR1 to provide accurate kinematic maps, showing that they have similar yet clearly distinct kinematic trends that can be used to establish the processes that form them. Very recently, Laporte et al. (2020) studied the  $[Mg/Fe]$ – $[Fe/H]$  distribution of the ACS and MNC with a combination of *Gaia* DR2 data and LAMOST-SEGUE-APOGEE to intercede in favour of a disc origin for the two structures, guided also by their previous simulations of an isolated MW interacting with Sgr. In their work, they use colour-magnitude diagrams (CMDs) to also show that both have a conspicuous red clump (RC), in contrast with previous studies that mainly focused on the main sequence (MS), the MSTO, or the Two-Micron All-Sky Survey (2MASS, Skrutskie et al. 2006) M-giants. Here we aim to provide an independent detection and characterisation of these structures that can help us clarify their true extent and 3D morphology, as well as their nature. A deeper understanding of the events that lead to the observed stellar distribution in the anticentre could be used to constrain the orbit and mass of Sgr, as well as its effect on the gas and stars of our Galaxy.

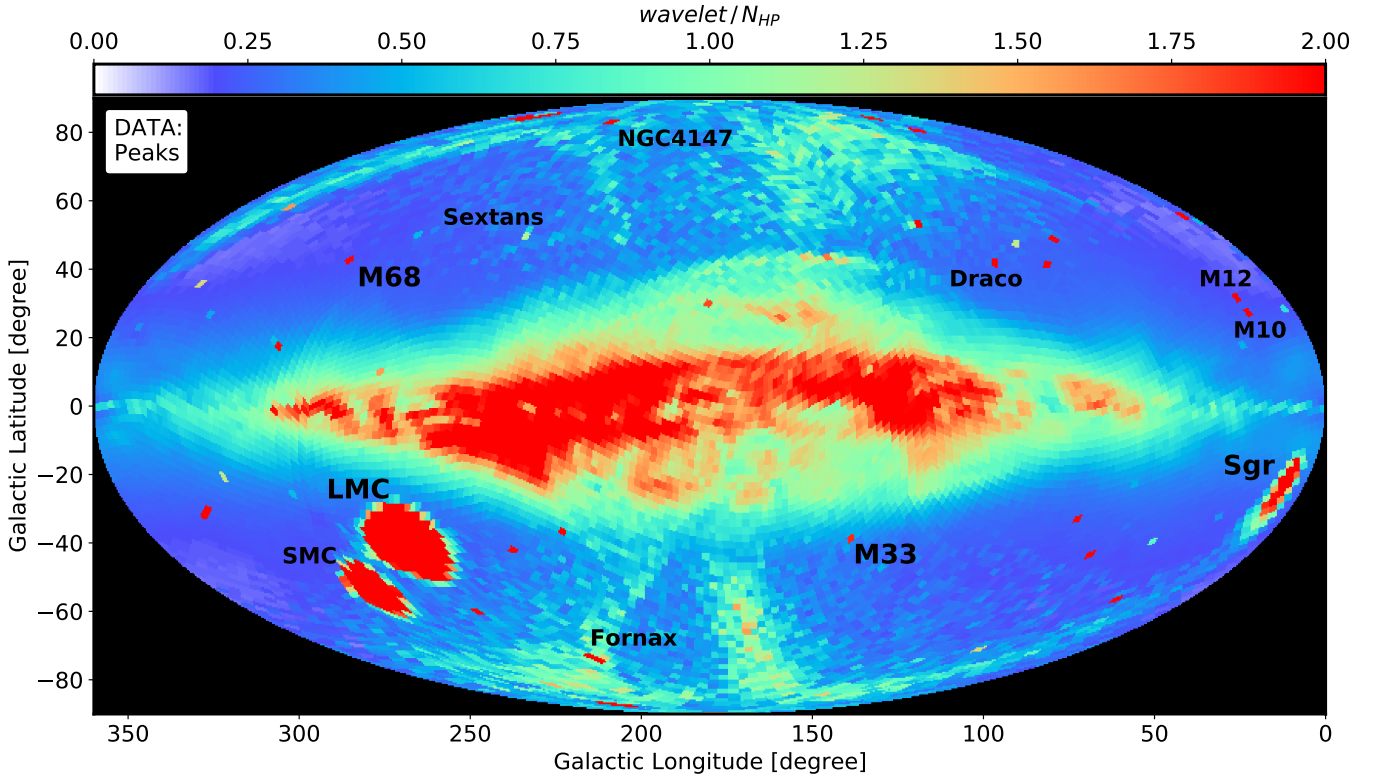
In this work, we search for substructure following the strategy devised by Antoja et al. (2015). The original goal of this method was to detect ultra faint dwarf galaxies in the halo using the fact that these galaxies should create, simultaneously, an overdensity in proper motion space and in the sky. Here we use the first half of the methodology – that is, its application in proper motion space only – to find the kinematic substructure at large heliocentric distances. This approach allows us to scan the whole celestial sphere systematically, homogeneously, and using a statically robust technique that can distinguish small but significant overdensities in proper motion space, as well as to track their changes as we move with Galactic longitude and latitude. As a result, our all-sky maps are dominated by three large structures: the Magellanic Clouds, the Sgr stream (as reported in Antoja et al. 2020), and MNC-ACS. Our goal is to map and study the morphology and kinematics of the structures in the Galactic anticentre using our methodology, which allows us to detect the structures and obtain a large set of members with (almost) no prior information.

This paper is organised as follows. In Sect. 2, we describe the strategy used to process the large amount of data available with *Gaia*. Section 3 then enumerates the different systems detected with our method, and, in Sect. 4, we focus on characterising the complex kinematics of the anticentre, especially in the north where we observe MNC and the ACS. We discuss the implications of our findings in Sect. 5, and present our conclusions in Sect. 6.

## 2. Data and methods

In this work, we used the same sample and methodology presented in Antoja et al. (2020, hereafter A20), which we reproduce here for convenience. We exploited the full *Gaia* DR2 (Gaia Collaboration 2018a), not restricting ourselves to any magnitude limit other than the one intrinsic to the instruments and only

<sup>1</sup> Using RC stars, and assuming no extinction, we can potentially reach up to distances of roughly 100 kpc from the Sun.



**Fig. 1.** Mollweide projection of the dominant peak in proper motion space detected at each HEALpix. The data are coloured by the relative intensity (Eq. (2)). By showing only the most significant kinematic overdensity and normalising to the number of stars in the HEALpix, a large number of structures become visible: the Sgr stream, tens of globular clusters, and an intricate anticentre. We have labelled some of most relevant ones.

applying the following two filters:

$$\varpi - \sigma_{\varpi} < 0.1 \text{ mas and } G_{BP} - G_{RP} > 0.2 \text{ mag,} \quad (1)$$

which are aimed at reducing the level of foreground contamination, that is, nearby stars that block our view of the halo and outer disc.

The resulting sample contains 700 412 152 sources. Properly processing and analysing such a large data set is obviously impractical with a regular desktop computer. It would typically require Big Data infrastructure. Nevertheless, given that we want to study changes in the velocity planes, our observables are the proper motion histograms themselves. Hence, we downloaded them in parallel directly from the *Gaia* Archive<sup>2</sup> with the query:

```
SELECT COUNT(*) as N, pmra_index*BINSIZE as pmra,
pmdec_index*BINSIZE as pmdec FROM (SELECT
source_id, FLOOR(pmra/BINSIZE) AS pmra_index,
FLOOR(pmdec/BINSIZE) AS pmdec_index FROM
gaiadr2.gaia_source WHERE source_id BETWEEN
HPNUM*2**35*4**(12-LVL) AND
(HPNUM+1)*2**35*4**(12-LVL) AND
parallax-parallax_error < 0.1 AND bp_rp >= 0.2
AND pmra IS NOT NULL AND pmdec IS NOT NULL) as
sub GROUP BY pmra_index, pmdec_index
```

where *LVL* is the level of the HEALpix grid (here, 5), *HPNUM* is the HEALpix to be processed, and *BINSIZE* is the size of the histogram binning (here, 0.24 mas yr<sup>-1</sup>).

Then we applied a wavelet transformation (WT, Starck & Murtagh 2002), followed by a peak detection algorithm, to each

of the 12 288 histograms we downloaded to detect the significant kinematic structures (assuming Poisson noise). The resulting wavelet coefficient of each peak is then, by construction, proportional to its density in proper motion space. To simplify the analysis, we only kept one structure at each HEALpix, the one with the highest relative intensity ( $WT/N_{hp}$ ):

$$\frac{WT}{N_{hp}} \times 1000, \quad (2)$$

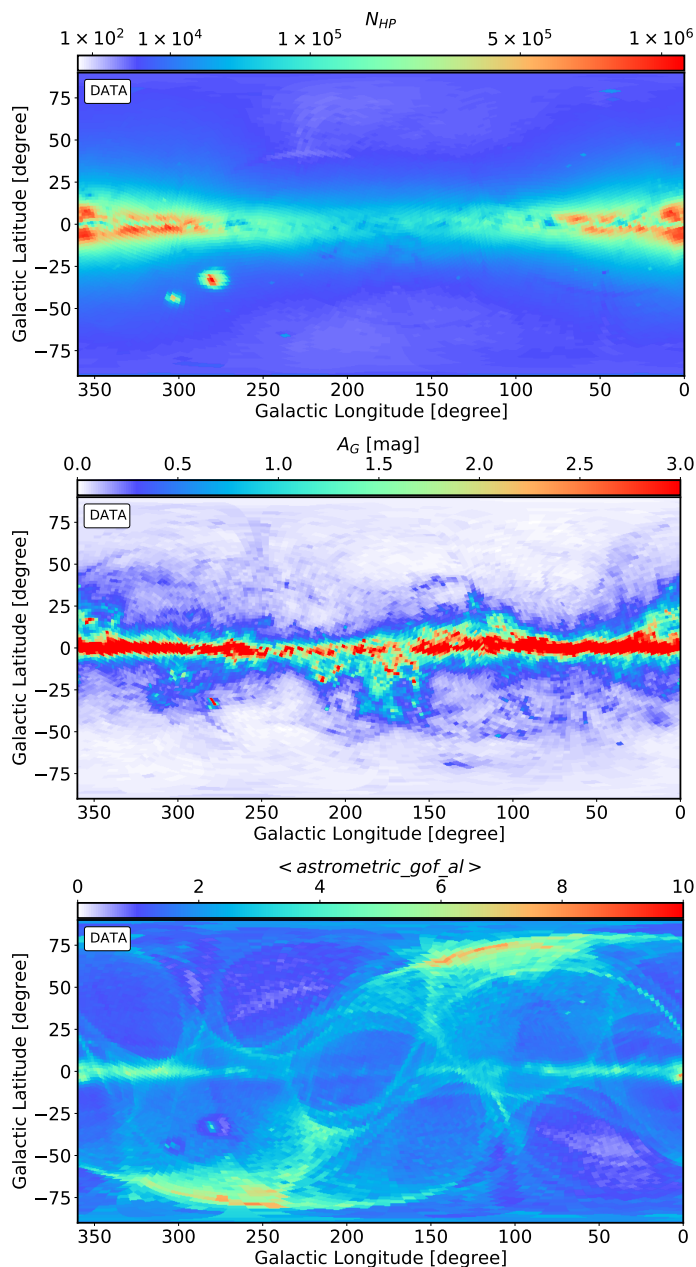
where  $N_{hp}$  is the total number of sources in the HEALpix and is used to normalise the wavelet coefficient (for more details on the method, see A20).

### 3. Global map of the substructures

Figure 1 shows the Mollweide projection of the sky in Galactic coordinates coloured by the relative intensity (Eq. (2)) of the highest peak in the proper motion histogram. By selecting only the overdensity with the largest intensity present at each proper motion histogram, we can focus on the dominant kinematic structure of the HEALpix<sup>3</sup>. The normalisation used in Eq. (2) compensates for the density gradient of the Galaxy and gives more contrast to the structures at higher latitudes. This figure reveals a wealth of substructure that cannot be seen with a simple density map of our sample. For instance, Fig. 2 contains the number of sources that pass our filters at each HEALpix (top panel) where we can only identify the Magellanic Clouds, some globular clusters, and the imprints of the

<sup>3</sup> In some cases, the dominant peak might not be the peak in the proper motion plane with the largest amount of stars inside it.

<sup>2</sup> Hosted at: <https://gea.esac.esa.int/archive/>



**Fig. 2.** Mean properties of the HEALpix. *Top*: number of stars in each HEALpix that fulfil the selection described in Sect. 2. *Middle*: absorption at infinity in the  $G$  band at each HEALpix, obtained from the Schlegel et al. (1998) maps with the re-calibration by Schlafly & Finkbeiner (2011) and using the mean  $G_{BP}-G_{RP}$  colour in the HEALpix together with the transformations described in Appendix A of Ramos et al. (2020). *Bottom*: average `astrometric_gof_al` at each HEALpix (see Appendix A).

extinction (which are shown in the middle panel for comparison, Schlegel et al. 1998) or the *Gaia* scanning law (bottom panel).

With Fig. 1 we have been able to detect the following. First, we have the Magellanic Clouds. Their angular size in the kinematic maps, in contrast with their apparent angular size in the star count map of Fig. 2 (top panel), is larger and shows the true extent of these systems as already noted in, for example, Gaia Collaboration (2018c). We also detect substructure within them (not shown here), and we are able to recover some of the glob-

ular clusters<sup>4</sup> that orbit the Large Magellanic Cloud, such as the recently detected *Gaia* 3 (Torrealba et al. 2019).

Second, we have the Sgr stream. The core of this dwarf galaxy and its stream are also clearly visible in our map. Interestingly enough, we do not observe it in the top panel of Fig. 2, which highlights the difficulty, even in the *Gaia* era, of detecting this structure with just stellar counts.

Third, almost vertically mirrored to the Sgr stream, we note a feature with a stream-like shape. The CMDs of the sources that produce these peaks do not present any coherent isochrone-like shape. Instead, they appear clumped around the faint and blue corner of the diagram. Added to the fact that their proper motions are always normally distributed around the origin, regardless of the position in the sky, this leads us to conclude that these sources are actually quasars. Although quasars are ubiquitous and should not produce a band in the sky, the scanning law of *Gaia* favours certain regions of the sky, as can be seen in the bottom panel of Fig. 2, with the `astrometric_gof_al`<sup>5</sup> that quantifies the quality of the astrometric solution.

The only source of dispersion in proper motion space for the quasars is the observational uncertainties. In contrast, the dispersion for the halo stars is driven by their kinetic temperature, which is usually larger than the uncertainties. Therefore, the overdensity caused by the quasars only becomes compact enough to be detected with an intensity larger than that of the halo population in the parts of the sky where the precision is sufficiently high.

Since we did not expect to obtain such a clear signal from the quasars, we did not remove them beforehand. Nevertheless, our results are not affected by their presence. With the list of extended objects that will be published in *Gaia* DR3, we will be able to remove these objects upfront within the queries.

Fifth, we have nearby galaxies. Apart from the Magellanic Clouds, we also detect M 31 and M 33, the latter clearly visible in Fig. 1 at  $(l, b) \sim (133^\circ, -31^\circ)$ .

Sixth, we have dwarf spheroidal galaxies. Our method is able to detect several of them, such as Fornax, Sextans, and Sculptor, as well as fainter ones such as Draco.

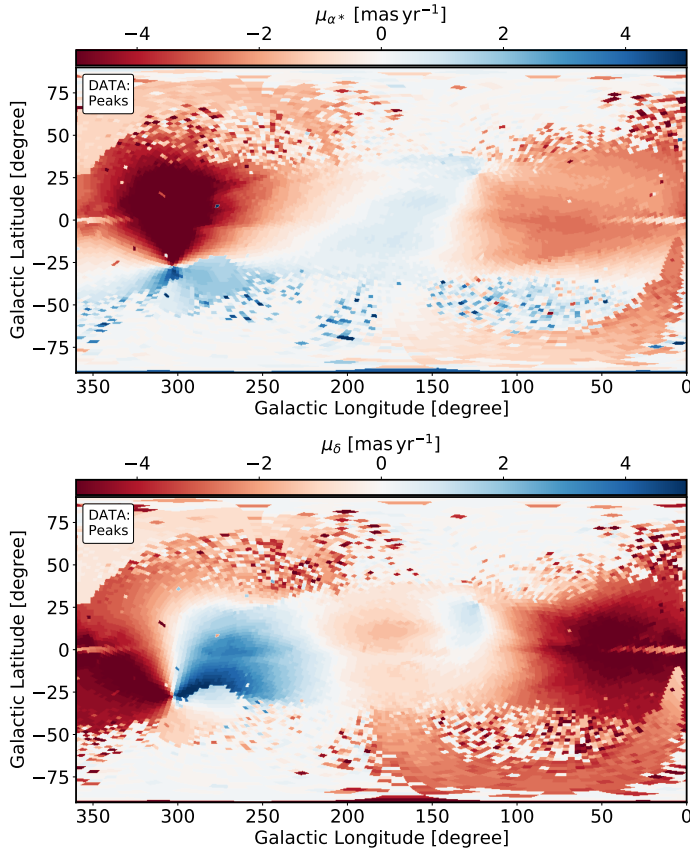
Seventh, we have globular clusters. We recover 51 globular clusters from the Bica et al. (2019) catalogue. From their 200 objects classified as globular clusters, more than half are in the bulge where we do not detect any, either because the contrast with the foreground is too low or due to the cut in parallax that is applied.

Eighth, we have ultra faint dwarf galaxies. We do not recover any of the known ultra faint dwarfs based solely on their kinematic signature. Their proper motion uncertainties are too large and there are too few members (see Massari & Helmi 2018) to produce a significant overdensity. Nevertheless, we note that when we apply the full methodology described in Antoja et al. (2015), which includes the search of peaks in the sky and not only in proper motion as done here, we can effectively recover most of them.

Finally, we have the anticentre. Apart from all the substructure we find in the halo, our methodology reveals complex kinematic substructures towards the anticentre of the MW, dominated by two arch-like features in the northern Galactic hemisphere. After comparing them with the extinction map shown in the

<sup>4</sup> More specifically, we have detected Magellanic Halos's Clusters as they are described in Bica et al. (2019).

<sup>5</sup> This 'Gaussianised chi-square' is an indicator of the quality of the astrometric solution. Values above +3 therefore indicate a bad fit to the data. Other indicators are `astrometric_excess_noise` or `astrometric_n_gof_obs_al`.



**Fig. 3.** Proper motion coordinates of the dominant peak at each HEALpix. *Top*: proper motion in right ascension. *Bottom*: proper motion in declination. Two conspicuous stream-like patterns crossing the entire Celestial sphere can be seen, one of which is Sgr (bottom right to top left) and the other corresponds to the quasars (bottom left to top right).

middle panel of Fig. 2, we confirm that these features are not aligned with regions of high absorption. Also, we checked how the `astrometric_gof_al` map (bottom panel of Fig. 2) superposes to the WT intensity map, from which we conclude that the shape of the bottom arch is artificially enhanced by the scanning law. The cavity at  $l \sim 180^\circ$ ,  $b \sim 20^\circ$  coincides with a region poorly sampled by *Gaia*, and therefore the intensity (proportional to stellar counts) is lower.

We note that some of these structures also appear in Fig. 3, where we colour the sky according to the proper motion<sup>6</sup> of the highest peak (the peak used to colour Fig. 1). The clearest one is the Sgr stream, which we analysed in detail in A20. Furthermore, the structure at latitude  $b \sim 35^\circ$  ( $140^\circ < l < 200^\circ$ ) appears as a conspicuous arch in the proper motion map. We devote the following sections to the analysis and characterisation of the kinematic substructure present at the outer disc, focusing mostly in the north where we observe these two conspicuous aforementioned arches.

<sup>6</sup> We represent the proper motions in equatorial coordinates since this is the space where we detect the peaks. The sky coordinates of different peaks do not have to be the same even if they are located at the same position in proper motion space, thus they will spread in velocity after being transformed into Galactic coordinates. Trying to convert to Galactic coordinates artificially increases the dispersion and introduces an undesirable correlation between the two components of velocity.

## 4. Kinematic features in the anticentre

### 4.1. MNC and the ACS

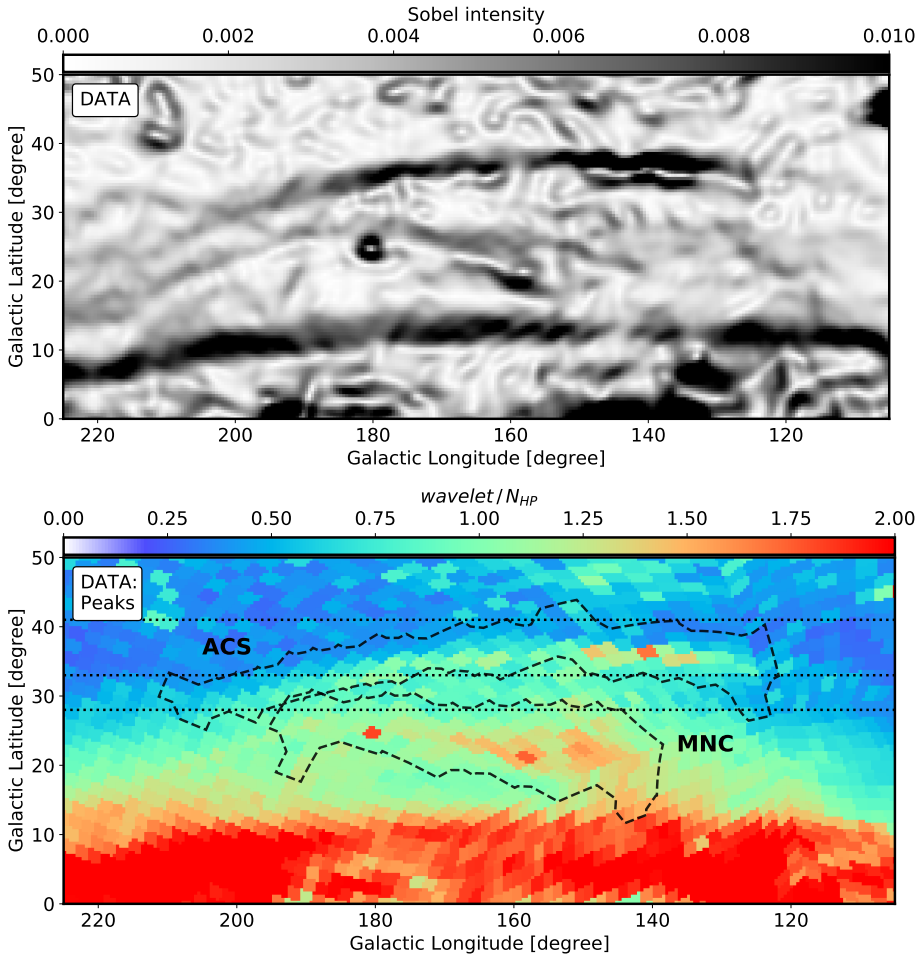
In Fig. 4 we present a zoom-in of Fig. 1 towards the anticentre and show our selection of the two structures that appear after colouring the sky according to the relative intensity of the dominant structure in proper motion. To build this selection in an objective manner, we first applied a Gaussian softening (two sigmas) of the 2D image to erase the HEALpix limits and then applied a bi-directional *Sobel* filter<sup>7</sup> to reveal edges. By doing so, the two arches are cleanly separated at all longitudes. The final step is to select only the HEALpix whose *Sobel* intensity is above a certain threshold (0.0035 for the bottom arch, 0.0040 for the top one). However, if we applied this selection blindly, we would obtain a long list of HEALpix that comprises several structures. Instead, we first drew a rectangle around each arch in an appropriate coordinate system. This coordinate system, different for the bottom and top arches, was obtained by rotating the celestial sphere with respect to the Galactic reference frame until the structure lies roughly flat at zero latitude in the new reference frame. The resulting final selections are the contours shown in the bottom plot of Fig. 4. The structures can be seen to continue beyond the defined contours, especially for the feature at lower latitudes, but we focus on the regions where they are the most intense. We added three horizontal lines that represent an approximated latitude limit of each structure at  $l \sim 180^\circ$ .

By comparing the shape of these structures and their location in the sky, we note that they match with the MNC (bottom) and the ACS (top) (e.g., Newberg et al. 2002; Grillmair 2006; Slater et al. 2014; Morganson et al. 2016). The patches we obtain are also in good agreement with the regions delineated by Laporte et al. (2020) but are much more concise. However, in contrast to previous works, their morphology appears sharper and well defined since we did not rely on counts but instead detected these structures in relative intensity (Eq. (2)). For instance, we observed an MNC structure that has a clear arch-like shape<sup>8</sup> extending from  $\sim 120^\circ$  to  $\sim 230^\circ$  in longitude, where it meets the disc at a latitude of  $\sim 10^\circ$ . Nevertheless, we stress again that in the case of MNC, although the structure is physical, the Sobel filter enhances the edge caused by the scanning law of *Gaia*. In contrast, the ACS is thinner, stays above MNC for the ranges of longitudes where we detect it, and has its strongest signal at  $l \sim 140^\circ$ , where MNC has already almost merged with the disc. This is the most precise picture of the anticentre available to date and is possible thanks to the introduction of kinematic information in the detection of these structures.

With the MNC and ACS regions now identified, we explore their kinematics and CMDs in more detail in Fig. 5. Each row contains the results for different regions: The first row is MNC, the second corresponds to the list of HEALpix that fall between MNC and the ACS (hereafter, the bridge), the third is the ACS, and, finally, the fourth is a region above the ACS. The last row is an example of what we would expect from a galaxy with no substructure obtained from a mock catalogue (Appendix B). With this exercise we can evaluate the continuity of these structures, compare their characteristics with nearby regions in the sky where we do not see an enhancement in relative intensity (cf. Fig. 4), and contrast them with the predictions of an MW

<sup>7</sup> Included in the *Python* package `Scikit-image` (van der Walt et al. 2014).

<sup>8</sup> The strong red HEALpix at  $(l, b) \sim (180^\circ, 25^\circ)$  is the globular cluster NGC 2419, which is far beyond MNC at a distance of  $\sim 83$  kpc (Forbes et al. 2008).



**Fig. 4.** Zoom-in on the anticentre region and definition of the patches. *Top:* result of applying a Gaussian smoothing plus Sobel filter to Fig. 1 in the region:  $110^\circ < l < 220^\circ$ ,  $-50^\circ < b < 50^\circ$ . We use this map to isolate the ACS and MNC. *Bottom:*  $wavelet/N_{HP}$  (relative peak intensity) of the same region. The black contours delineate the regions that we isolated according to the *upper panel*, while the dotted horizontal lines ( $b = [28^\circ, 33^\circ, 41^\circ]$ ) give an approximate limit for the structures in latitude at  $l \sim 170^\circ$ .

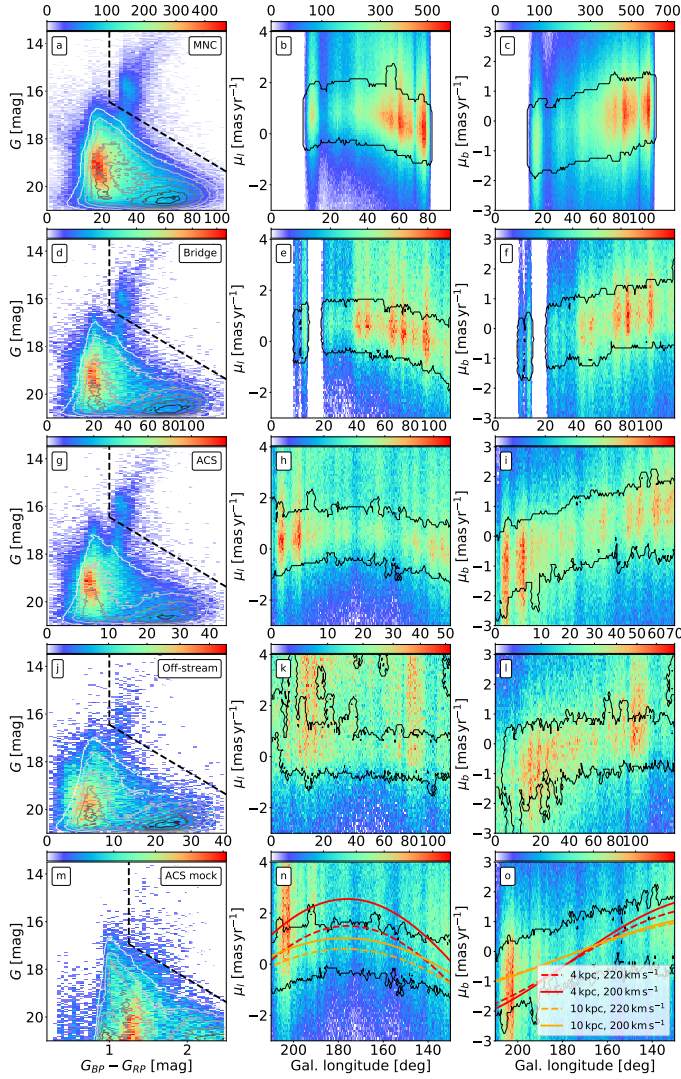
model. In the first column of this plot, we have aggregated all the stars that, in their respective HEALpix, fall within the highest intensity proper motion peak. We refer to such stars as ‘peak stars’. To show what we would see if we had not done this kinematic selection, the grey contours on top represent the CMD of all the stars of the region. The second and third columns contain, respectively, the trends of  $\mu_l$  and  $\mu_b$  with Galactic longitude for all the stars within the region. Here the black line encircles the kinematically selected stars, that is, the outer contour of the volume that the peak stars occupy in this space. To provide some contrast with a fiduciary galaxy, we repeated the same process in the bottom row for the particles in the mock catalogue that fall within the ACS footprint. In this case, the peak stars were selected according to the position and size of the peaks detected in the data, which is why the contours of panels h (i) and n (o) are so similar.

The first thing that we note is the presence of a giant branch all the way from MNC to the ACS, which disappears once we explore latitudes larger than  $\sim 40^\circ$ . The fact that we see a well-defined RC means that these stars share a similar distance that, based on their magnitudes ( $\sim 15.5$  mag) and Galactic latitudes ( $b > 15^\circ$ ), puts them well above the mid-plane of the Galaxy ( $z > 2$  kpc) at a height larger than the scale height of even the thick disc. If we compare the observed CMDs with the CMD of the mock catalogue, we note that we do not expect many stars in this region of the diagram since the nearby giants have already been removed with the cut in parallax (they are bright enough to have a reliable parallax) and the farther ones are not in the model as there are not many stars at such heights and galactocentric distances.

We also note that a dense clump of stars accompanying the giant branch appears in panels a, d, and g; this clump is bluer than the MS of the disc seen in the mock catalogue. Newberg et al. (2002) reported that the MS turnoff of MNC was bluer than the thick disc, and we detected the same behaviour for the stars in the peaks found within MNC and the ACS. This group of stars is consistent with being the MS of an isochrone containing the RC discussed above. The rest of the stars that fall outside said isochrone seem to follow the contours of the CMD obtained with all the stars (no kinematic selection, grey contours in those panels) and are most likely nearby dwarf field stars that overlap with these structures in the proper motion plane. In contrast to previous works (e.g., Newberg et al. 2002; Ivezić et al. 2008; Xu et al. 2015; Thomas et al. 2019) that detected an overdensity in counts for a given population (MS, MSTO, blue stragglers, or M-giants), here we instead unveiled the whole sequence in the CMD by performing a blind kinematic selection of the stars.

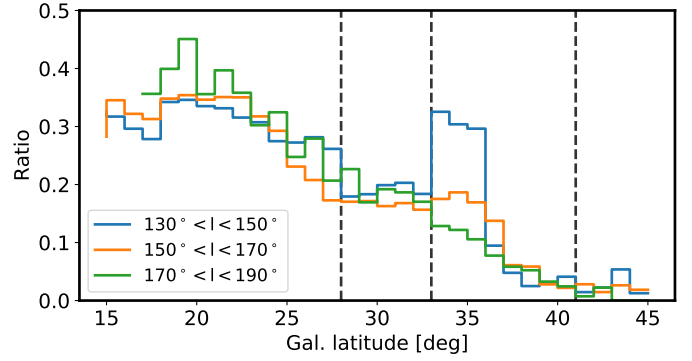
In the bottom row of Fig. 5, we have included a few curves that represent the proper motion that *Gaia* would measure for a star at a given distance if it only had azimuthal velocity<sup>9</sup>. In orange (red), this distance is 10 kpc (4 kpc), and for the solid

<sup>9</sup> Here we simply used the analytical expressions that transform the velocity of a star at a given position ( $l$ ,  $b$ , and distance) and which has only rotational velocity to proper motions in  $\mu_l$  and  $\mu_b$ . For the position and velocity of the Sun with respect to the Galactic centre, we have used  $R_\odot = 8.178$  kpc (GRAVITY Collaboration 2019) and  $V_\odot = [11.1, 248.5, 7.25]$  km s<sup>-1</sup> (Schönrich et al. 2010; Reid & Brunthaler 2020). In all the cases we keep the latitude constant to  $30^\circ$  since we compare the lines with the tracks obtained in the ACS region, but we obtain similar results for the MNC region.



**Fig. 5.** CMD (*left*) and proper motion trends with Galactic longitude in  $\mu_l$  (*middle*) and  $\mu_b$  (*right*) of each structure. The CMDs contain the histogram of the stars within the highest intensity proper motion peaks (peak stars) with grey contours on top that represent the CMDs of all the stars in the region. The dashed black lines represent our selection of giant stars. The proper motion maps contain all the stars in the region, and the black line is the zero-contour of the peak stars, that is, the stars that fall within the highest intensity proper motion peak of their HEALpix. *First row:* MNC. *Second row:* bridge between MNC and the ACS. *Third row:* ACS. *Fourth row:* above the ACS. *Fifth row:* same region as the ACS but for the mock particles, selecting the stars for the CMD according to the contours of *panels h* and *i*. In the *bottom panel* we also include the proper motions expected from a structure  $30^\circ$  above the plane at a given distance – 4 (red) or 10 (orange) kpc – rotating at a given velocity – 200 (solid) or 220 (dashed)  $\text{km s}^{-1}$  – but with no radial or vertical velocity.

(dashed) line the rotation velocity is  $200 \text{ km s}^{-1}$  ( $220 \text{ km s}^{-1}$ ). A structure that is too near – such as the brown curve – does not match the contour delineated by the peak stars (black contours), while a structure that does not rotate – such as, potentially, the halo – would have to be at a distance larger than 50 kpc on average to fall within the black lines. Even then, its shape would not be compatible with the data. While we note that other combinations of distances and velocities could produce a similar shape (even if the result is not physically supported), the dashed red line shows a good agreement with our observations



**Fig. 6.** Ratio of giants in the peak compared to all the giants as a function of Galactic latitude for three different ranges in longitude:  $130^\circ < l < 150^\circ$  (blue),  $150^\circ < l < 170^\circ$  (orange), and  $170^\circ < l < 190^\circ$  (green). The vertical lines give an orientation of the end of each structure with Galactic latitude (cf. Fig. 4). A sudden increase in the ratio can be seen in the part where the ACS is the more intense.

and corresponds to a structure at  $\sim 10$  kpc rotating slightly more slowly than the disc. In other words, the peak stars in the ACS have proper motions that change with Galactic longitude in a way that is compatible with a structure at a distance of  $\sim 10$  kpc rotating at a speed similar to the disc or slower, in agreement with the analysis by [de Boer et al. \(2018\)](#); the same also applies to MNC.

By comparing the CMDs inside and outside the patches defined in Fig. 4, it is clear that we can gain contrast with the MW foreground by focusing only on the giants, even though the MS of these structures is the dominant fraction. Therefore, we introduced another tag, in addition to the one we have already been using to separate stars inside and outside the proper motion peaks. The stars will be called giants whenever they are redder than  $G_{BP} - G_{RP} > 1$  mag and their apparent magnitude is smaller (brighter) than the line:

$$G < 1.95 (G_{BP} - G_{RP}) + 14.50. \quad (3)$$

Here we used the slope calculated in [Romero-Gómez et al. \(2019\)](#) to follow the extinction vector, and the zero-point was adjusted by eye to reduce the contamination from the disc while preserving the RC as much as possible. We note, however, that we still selected some faint red dwarfs at all latitudes, the great majority of which are not classified as peak stars (for sources redder than 2 mag in  $G_{BP} - G_{RP}$  and  $G > 14$  mag, only  $\sim 100$  out of  $\sim 8000$  in MNC and  $\sim 30$  out of  $\sim 4000$  in the ACS). This means that the giants also tagged as peak stars are more likely to be true giants, whereas field stars tagged as giants have a larger probability of being nearby red dwarfs.

Figure 6 shows the fraction of giants inside the peaks (i.e. stars tagged as giants and peak stars simultaneously) with respect to all the stars tagged as giants as a function of latitude. Given that our classification is rather rough, we should treat these as simple estimates and focus on the trends instead. What we observe is that the parts where MNC and the ACS have the strongest signal in relative intensity (cf. Fig. 4) coincide with the regions where this ratio is the highest. We have already seen that the relative intensity of the ACS decreases with longitude, and here we note that the ratio of giants also diminishes moving from one curve to the other. We also observe that the bridge keeps a constant ratio, showing that it is just the region where the tail of the two structures overlap. Finally, we note that our patch around the ACS is too broad as the ratio drops abruptly at

**Table 1.** MNC stars classified both as peak and giants (top two rows).

Source_id	RA [°]	Dec [°]	$\mu_{\alpha^*}$ [mas yr <sup>-1</sup> ]	$\sigma_{\mu_{\alpha^*}}$ [mas yr <sup>-1</sup> ]	$\mu_{\delta^*}$ [mas yr <sup>-1</sup> ]	$\sigma_{\mu_{\delta^*}}$ [mas yr <sup>-1</sup> ]	$G$ [mag]	$G_{BP}-G_{RP}$ [mag]	$A_G$ [mag]
926626863262740096	116.29	43.76	0.0070	0.1237	-0.9832	0.0973	16.40	1.20	0.10
926656550076613376	116.15	44.00	-0.6643	0.0817	-0.6905	0.0564	15.66	1.11	0.10

**Notes.** Column 1 contains the source\_id followed, in Cols. 2 and 3, by the right ascension and declination of the star. Columns 4–7 contain the proper motions (equatorial coordinates) and the corresponding uncertainties. Columns 8 and 9 are the apparent magnitude in the  $G$  band and the  $Gaia$  colour  $G_{BP}-G_{RP}$ , respectively. The Galactic coordinates,  $l$  and  $b$ , are given in Cols. 12 and 13. Finally, in the last column, we give the absorption in the  $G$  band (see text). The full table is available at the CDS.

**Table 2.** Same as Table 1 but for the ACS (top two rows).

Source_id	RA [°]	Dec [°]	$\mu_{\alpha^*}$ [mas yr <sup>-1</sup> ]	$\sigma_{\mu_{\alpha^*}}$ [mas yr <sup>-1</sup> ]	$\mu_{\delta^*}$ [mas yr <sup>-1</sup> ]	$\sigma_{\mu_{\delta^*}}$ [mas yr <sup>-1</sup> ]	$G$ [mag]	$G_{BP}-G_{RP}$ [mag]	$A_G$ [mag]
703742318576330496	128.16	26.59	-0.2006	0.0740	-0.7579	0.0510	15.12	1.23	0.09
703750633632989056	128.28	26.72	-0.1967	0.1102	-0.3101	0.0762	16.26	1.18	0.11

**Notes.** The full table is available at the CDS.

$b \sim 37^\circ$ , coinciding with the place where we observe a discontinuity in the kinematics<sup>10</sup>.

Tables 1 and 2, available at the CDS, contain the list of sources classified simultaneously as peak stars and giants for MNC (10 079 sources) and the ACS (2104 sources). These tables include, apart from sky positions and proper motions, the apparent magnitude, colour and absorption.

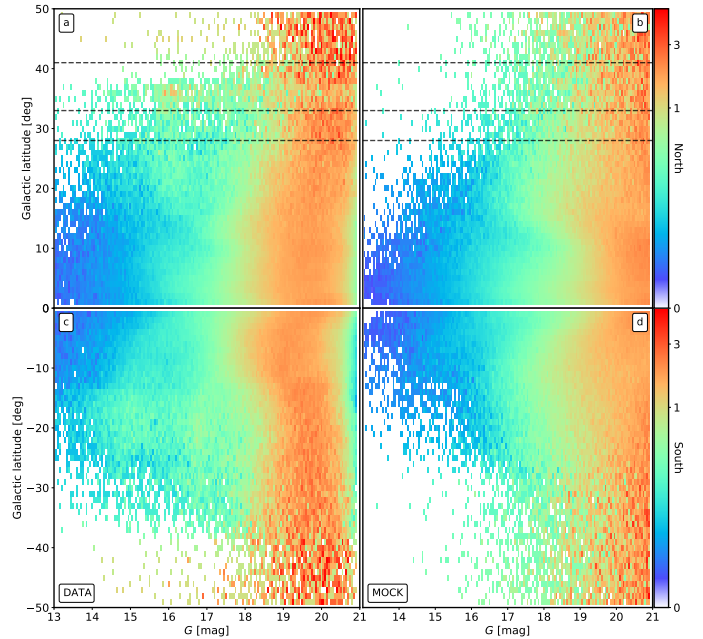
#### 4.2. Anticentre region: North versus south

Above, we have proven that our methodology can detect kinematic substructures and isolate them effectively from the rest of the disc. We have also proven that the giants inside the proper motion peaks are good tracers of MNC and the ACS since the contamination in that region of the CMD is expected to be very low once we have filtered by parallax and kinematics. Therefore, we can use the location of the RC to trace the structures in physical space. To do so, we now explore the changes in apparent magnitude of the stars that we tagged as peak stars for different ranges of longitude, both in the northern and southern hemispheres.

Figure 7 shows the distribution of apparent magnitudes of peak stars with respect to Galactic latitude using a histogram normalised by bins of  $b$  for the range  $130^\circ < l < 150^\circ$ . This figure compares the data (top) with the expectations from the mock catalogue (bottom). In this case, the particles selected in the mock catalogue correspond to the peaks that are detected in the mock itself (in contrast to panel m of Fig. 5, where we used the peaks detected in the data). We find an overdensity of stars in the north at a magnitude  $\sim 16$ , which corresponds to the RC seen in Fig. 5. The overdensity is most intense above  $b > 30^\circ$  and corresponds to the ACS. We see it extending rather continuously down to  $b \sim 10^\circ$  where it merges with the disc, following an arch that is compatible with the increase in extinction.

In the south, we observe an excess of bright stars ( $G < 17$  mag) at latitudes between  $15^\circ$  and  $25^\circ$  with respect to the mock. Interestingly enough, the intensity maps (Fig. 1) do not show an enhancement, as is the case for the north, not even

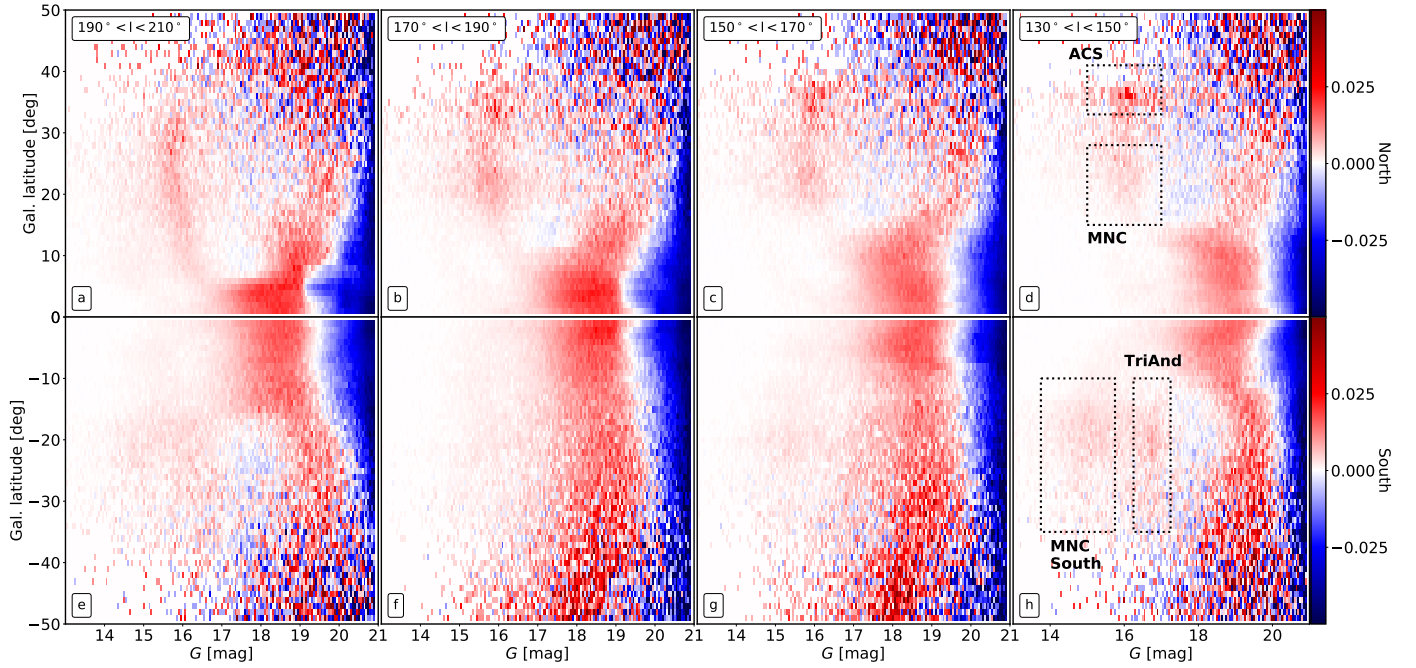
<sup>10</sup> We have verified that the trends in proper motion as a function of latitude do also indeed suffer a sudden change at around  $\sim 37^\circ$ , as can be seen in Fig. 3.



**Fig. 7.** Apparent magnitude as a function of Galactic latitude for the stars in the peaks with  $130^\circ < l < 150^\circ$ . *Left:* data. *Right:* mock. The histograms are normalised such that the sum of all pixels in a given bin of  $b$  adds up to unity. A conspicuous overdensity of stars can be clearly seen appearing at  $G \sim 16$  mag and extending from  $15^\circ < b < 40^\circ$ . The horizontal lines are taken from Fig. 4 and represent the latitude limits of MNC and the ACS as seen in the sky.

compared to the mock map (Fig. B.1). Based on their apparent magnitudes and location in the sky, it is very likely that these stars form the diffuse stellar population detected by [Ibata et al. \(2003\)](#) and which is sometimes called MNC South. This is also the region where the TriAnd overdensities have been reported ([Majewski et al. 2004](#); [Rocha-Pinto et al. 2004](#); [Martin et al. 2007](#)), and, as we show below, we do detect it with our method. A detailed study of these structures and a comparison with previous studies (e.g., Fig. 4 from [Perottoni et al. 2018](#)) is beyond the scope of this work, but we will revisit it once the EDR3 *Gaia*





**Fig. 8.** Differences between the data and the mock in the plane of apparent magnitude against Galactic latitude. *Panels a and h* are obtained by subtracting the counts (normalised by latitude) from the top and bottom panels of Fig. 7. The same is done for the rest of panels but at different ranges of longitude. *From right to left:*  $130^\circ < l < 150^\circ$ ,  $150^\circ < l < 170^\circ$ ,  $170^\circ < l < 190^\circ$ , and  $190^\circ < l < 210^\circ$ . *Panels d and h:* we highlight the regions of interest. The boxes around the ACS and MNC are based on the latitude ranges defined in Fig. 4.

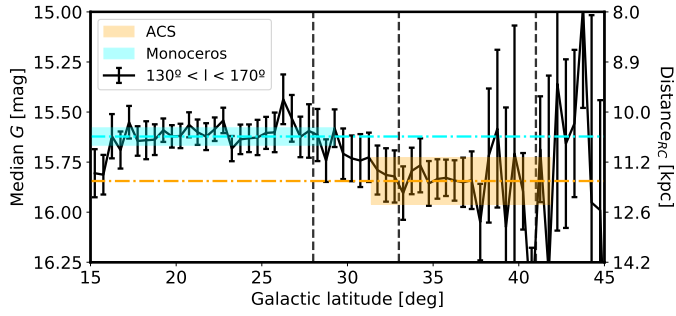
data release (Brown 2019) is made public and we have more and better astrometric data.

We now present, in Fig. 8, the difference between the data and the mock in the same plane of apparent magnitude against Galactic latitude for different bins in longitude. From right to left, they are:  $130^\circ < l < 150^\circ$ ,  $150^\circ < l < 170^\circ$ ,  $170^\circ < l < 190^\circ$ , and  $190^\circ < l < 210^\circ$ . In panels d and h, we see the subtraction of the left-hand panels of Fig. 7 from the right-hand panels (i.e. data minus mock after properly normalising the histograms). As already mentioned, we see two distinct overdensities in the south, a bright ( $14 < G < 15.5$  mag) one corresponding to MNC South and a fainter one ( $G \sim 17$  mag) corresponding to TriAnd, at a magnitude consistent with the most recent determinations of its heliocentric distance (e.g., Bergemann et al. 2018). These features cannot be seen as clearly at other longitudes with the exception of in panel e, where the diffuse overdensity appears again (latitudes between  $\sim 20^\circ$  and  $\sim 30^\circ$ , brighter than  $\sim 17$  mag). We cannot rule out the southern structures being discontinuous; however, taking the corrugations in the disc reported by Xu et al. (2015) into account, the most likely scenario is that we can only detect them with our method where the extinction is sufficiently low. We will be able to better assess their continuity with the next *Gaia* releases.

In the north (top panels), we note that the ACS decreases in intensity and shifts to lower latitudes when we move towards the third quadrant of the Galaxy, as we also see in Fig. 4. In the middle panels (b and c), we observe two concentrations at different latitudes, but similar apparent magnitudes, corresponding to MNC and the ACS, whose tails overlap to form the aforementioned bridge. More importantly, we see MNC extending more and more towards lower latitudes, keeping roughly the same apparent magnitude throughout. This is interesting as MNC is usually hard to trace so deep into the disc due to the foreground stars. And yet, in panel a, its RC can be traced down to a latitude of  $\sim 5^\circ$  using our kinematic selection.

By measuring the median  $G$  for the giant stars only, selected according to Eq. (3), we can investigate the relative distance of these structures. To do so, however, we focus only on the range  $130^\circ < l < 170^\circ$ , where they remain rather flat and the bridge is quite wide (see Fig. 4), since the latitudes probed by MNC and the ACS change with longitude. The median of the  $G$  magnitude and the associated one-sigma interval of uncertainty at different bins in latitude are shown in Fig. 9 for the giant stars inside the proper motion peaks. To compensate for the effects of extinction, we first corrected the apparent magnitude using the  $G_{BP} - G_{RP}$  colour of each source individually and the prescription detailed in Appendix A of Ramos et al. (2020). What we observe is that, below  $b \sim 28^\circ$  where we identify MNC, the RC is brighter than above  $b \sim 31^\circ$  where the ACS begins. Since we used the integrated extinction up to infinity (Schlegel et al. 1998), the separation that we observe is an upper limit: If we assume that the extinction applied to the ACS is correct, since it is at a higher latitude, then MNC could actually be less extinguished than assumed and therefore be intrinsically fainter than the value we are recovering. Nevertheless, since both structures are quite far, this effect should be small and we can safely conclude that the ACS is farther away than MNC.

To be more quantitative, we then measured the median  $G$ , corrected for extinction, for all the peak giant stars in this longitude range. The result is the shaded areas shown in Fig. 9, where we can clearly see that the ACS is, once we convert the difference in magnitude to distance, roughly 1 kpc farther away than MNC, with a discrepancy of more than  $3\sigma$ . Furthermore, we have estimated the median distance to each of the two structures. In doing so, we assumed that the median apparent magnitude measured (horizontal lines in Fig. 9) corresponds to the magnitude of the RC. By imposing that the absolute magnitude of an RC star is  $M_G = 0.495$  mag (Ruiz-Dern et al. 2018), we obtained the following median distances and their



**Fig. 9.** Apparent magnitude of giant peak stars in the anticentre region ( $130^\circ < l < 170^\circ$ ) as a function of Galactic latitude. The magnitudes shown on the left y-axis have been corrected for extinction (see text). The error bars denote the  $1\sigma$  uncertainty on the median computed as  $\sigma \sqrt{\frac{\pi}{2N}}$ , where  $\sigma$  is the standard deviation of the apparent magnitude in the bin. Vertical lines represent the approximate limits of each structure in that range of Galactic longitudes (see Fig. 4), and the right axis represents the distance to an RC star, with the apparent magnitude shown in the left axis. The horizontal lines correspond to the median  $G$  magnitude for the giant peak stars within MNC (cyan) and the ACS (orange). The shaded areas contain the  $\pm 3\sigma$  interval of uncertainty on the median, and they extend from the minimum to the maximum latitude of the peak stars within each patch (the vertical dashed lines serve only as an orientation). As can be seen, the ACS is fainter than MNC, and this translates into a difference in distance of  $\sim 1$  kpc.

statistical uncertainties<sup>11</sup>:  $D_{\text{MNC}} \sim 10.6 \pm 0.1$  kpc and  $D_{\text{ACS}} \sim 11.7 \pm 0.2$  kpc.

Nevertheless, without a precise calibration of each individual star and its extinction, we cannot investigate the changes in distance with longitude and latitude, which is key for revealing the 3D shape of these structures. We first made an attempt to study the distribution of the structures we detected along the line of sight by cross-matching the peak stars with StarHorse (Anders et al. 2019), a catalogue of Bayesian-derived astrophysical parameters obtained from the photometry of *Gaia*, Pan-STARRS1, 2MASS, and AllWISE combined. We downloaded all the stars in the anticentre ( $100^\circ < l < 260^\circ$  and  $-60^\circ < b < 60^\circ$ ) with SH\_OUTFLAG equal to ‘00000’, as recommended in Anders et al. (2019), and with a distance (50th percentile) less than 20 kpc. From the 13 098 038 peak stars in the north, we found 429 565 in StarHorse. In the south, the cross-match returned 514 167 stars out of the 13 669 647 peak stars. Most of them, however, are faint dwarfs found at low latitudes, closer than 10 kpc, whose parallax quality is not good enough to discard them with the filter presented in Sect. 2. If instead we restrict ourselves to the giant peak stars, we find 378 955 (out of 1 449 250) in the north and 458 403 (out of 1 286 132) in the south.

Figure 10 shows the distribution of StarHorse distances as a function of Galactic latitude for the four ranges of longitude explored above. The first thing we note is the effect of our selection function since the nearby giants are missing and a wall of stars at a distance of  $\sim 6$  kpc is formed. The tails extend up to  $\sim 15$  kpc, beyond which point StarHorse distance uncertainties become too large. Compared with the corresponding figure for the mock (Fig. B.3), where we see the disc extending much

farther away, we note a clear excess of stars in the data at latitudes larger than  $20^\circ$  and at a distance of  $>7$  kpc. We associate these with MNC and the ACS. In the right-hand column, in the range of longitudes where the ACS is more intense, we see it clearly separated from MNC and slightly farther away. As we shift our view towards the third Galactic quadrant, these structures recede, becoming less prominent and shifting to lower latitudes (as shown above). MNC covers a large range of latitudes and connects smoothly with the disc, but the lack of stars and the uncertainties prevent us from determining if there is a distance gradient with latitude or not.

The south does not show the same structures as the north, but we note that, at least for panels f and g, the extinction is higher than in the north, which could block our line of sight. We note that we do not recover the structures detected in panel h of Fig. 7 as clearly, probably due to low statistics. However, we do observe an increase in stars in panel e in the form of a diffuse distribution of distant stars at latitudes between  $-20^\circ$  and  $-30^\circ$ , which coincides with the location of the structure S200-24-19.8 reported in Newberg et al. (2002) as well as the detection by Xu et al. (2015) that they associated with the aforementioned TriAnd.

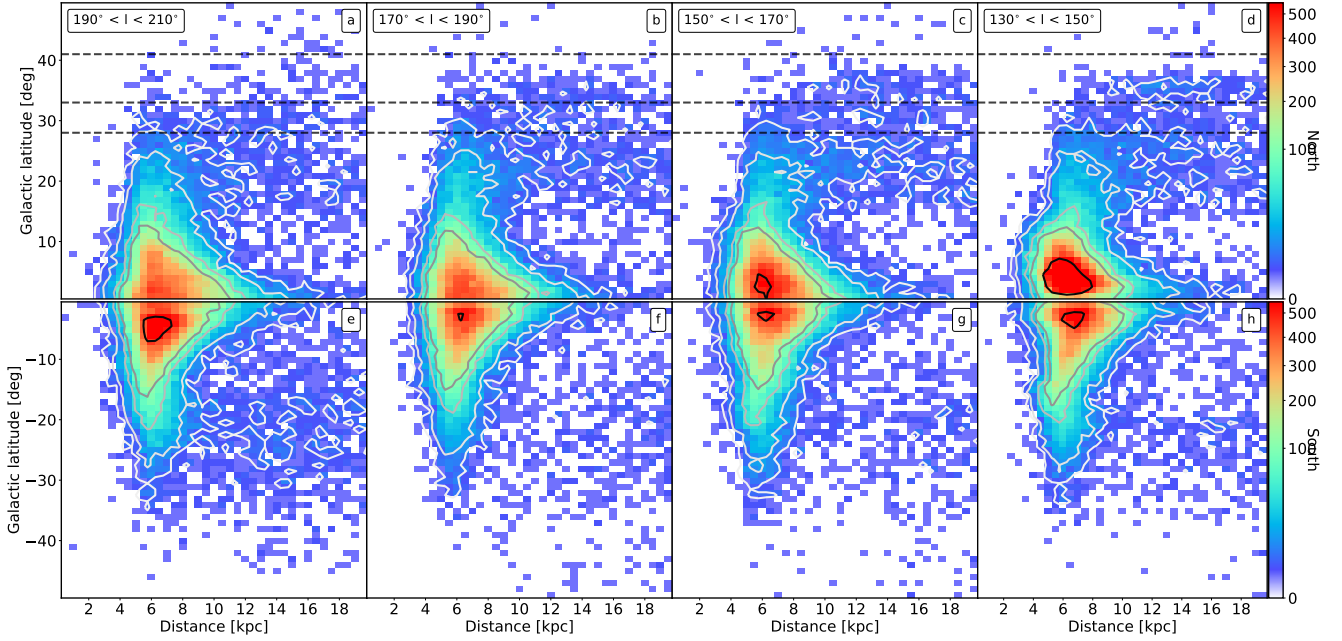
#### 4.3. RR Lyrae to M giant ratio

With a kinematically selected sample of stars for both MNC and the ACS, we can now check if their population is consistent with having been born in an extragalactic system or not. All known MW dwarf spheroidal galaxies have a large fraction of RR Lyrae stars (Vivas & Zinn 2006) and a low fraction of M giants (Price-Whelan et al. 2015), whereas the opposite occurs with the Galactic discs. Hence, we followed Price-Whelan et al. (2015), who used the ratio between the number of RR Lyrae and M giant stars in TriAnd (Rocha-Pinto et al. 2004; Martin et al. 2007) to argue that these structures were probably disc stars kicked out by an external perturbation. For that, we needed to estimate the number of RR Lyrae and M giants within MNC and the ACS. Based on previous studies, we expect a low number of RR Lyrae in these structures (Kinman et al. 2004) but a large number of M giants (e.g., Rocha-Pinto et al. 2004). Now the question is how low the ratio between the two populations is.

Starting with the RR Lyrae, we used the catalogue described in Mateu et al. (2020), which combines the VariClassifier and Specific Objects Studies (SOS) catalogues from *Gaia* DR2 (Holl et al. 2018; Clementini et al. 2019) with the ASAS-SN-II catalogue Jayasinghe et al. (2019), providing optimal completeness at the bright end ( $G < 15$ ). We selected only those sources that fall within the sky patches defined in Fig. 4, for a total of 900 (800) RR Lyrae in MNC (ACS), 253 of which fall within the one standard deviation range around the median distance in the case of MNC (6.7 to 16.7 kpc) and another 253 for the ACS (7.8 to 17.7 kpc). Of these, only 12 (six) are also consistent with the kinematic signature we detected in the space of  $l$ - $b$ - $pmra$ - $pmdec$ <sup>6</sup>. Based on these results, and after correcting for the completeness of the RR Lyrae catalogue (estimated at 80% at these magnitudes in Mateu et al. 2020), we conclude that MNC has, at most, 15 RR Lyrae and that the ACS has no more than eight. In parallel, we cross-matched the list of RR Lyrae with our sample to see the effect that the cut in parallax<sup>12</sup> has, and we observed that, from the 253 (253) stars that we had, only 123 (123) remain in MNC (ACS). Finally, if we now keep only stars classified as peak stars (i.e. probable MNC or ACS members), we find only one and two RR Lyrae for,

<sup>11</sup> Here, the dominant source of uncertainty is the systematic errors, which are not included in the given error bars. The two more important ones are (i) the assumption that the median magnitude is the magnitude of the RC and (ii) not using 3D extinction maps but instead correcting with the integrated extinction to infinity. Other sources of systematic uncertainty are: the error on the absolute magnitude of the RC, contamination from stars that are not giants, and errors in the extinction map.

<sup>12</sup> The effect of the cut in colour is negligible.



**Fig. 10.** Distance from StarHorse as a function of Galactic latitude for different slices in Galactic longitude. We show only the giant stars within the peak. *From right to left:*  $130^\circ < l < 150^\circ$ ,  $150^\circ < l < 170^\circ$ ,  $170^\circ < l < 190^\circ$ , and  $190^\circ < l < 210^\circ$ . The horizontal lines are at the same latitudes as in Fig. 4.

respectively, MNC and the ACS. These figures are lower and, even if we correct by the 48.6% reduction in completeness caused by the cut in parallax, the maximum amount of RR Lyrae in MNC (ACS) would be three (five).

On the other hand, the M giants are much more numerous. We used the official *Gaia* cross-match with 2MASS and the selection proposed by Majewski et al. (2003) to obtain the corresponding  $G_{BP}-G_{RP}$  colour cut necessary for MNC and the ACS, finding that we can confidently select M giants in our sample of candidates with the following limits:  $G < 15.5$  mag and  $G_{BP}-G_{RP} > 1.5$  mag. These cuts result in a total of 959 M giants for MNC and 155 for the ACS. Of course, these values are not corrected for the completeness of the sample, as opposed to those corresponding to the RR Lyrae. Therefore, if we take the maximum number of RR Lyrae that these structures can have, the ratio  $f_{RR:MG}$  that we provide becomes a strict upper limit:  $f_{RR:MG} < 1.5\%$  for MNC and  $f_{RR:MG} < 5.2\%$  for the ACS.

The obtained fractions are consistent with previous independent estimates (Sheffield et al. 2018), are totally compatible with a stellar population of the MW alpha-poor disc, and are much lower than the expected values for extragalactic systems like Sgr or the Large Magellanic Cloud, for which we expect a fraction of  $\sim 50\%$  (Price-Whelan et al. 2015). In fact, it is hard to reconcile these values with the hypothesis that MNC and the ACS are tidal tails of an accreted satellite. This would require either a young system that managed to reach high metallicities (perhaps formed from already metal-enriched gas) or the effect of a dynamical mechanism that could segregate RR Lyrae from giants within the stream, something that we do not observe in other streams such as the Sgr stream (e.g., Antoja et al. 2020; Ibata et al. 2020; Ramos et al. 2020).

## 5. Discussion

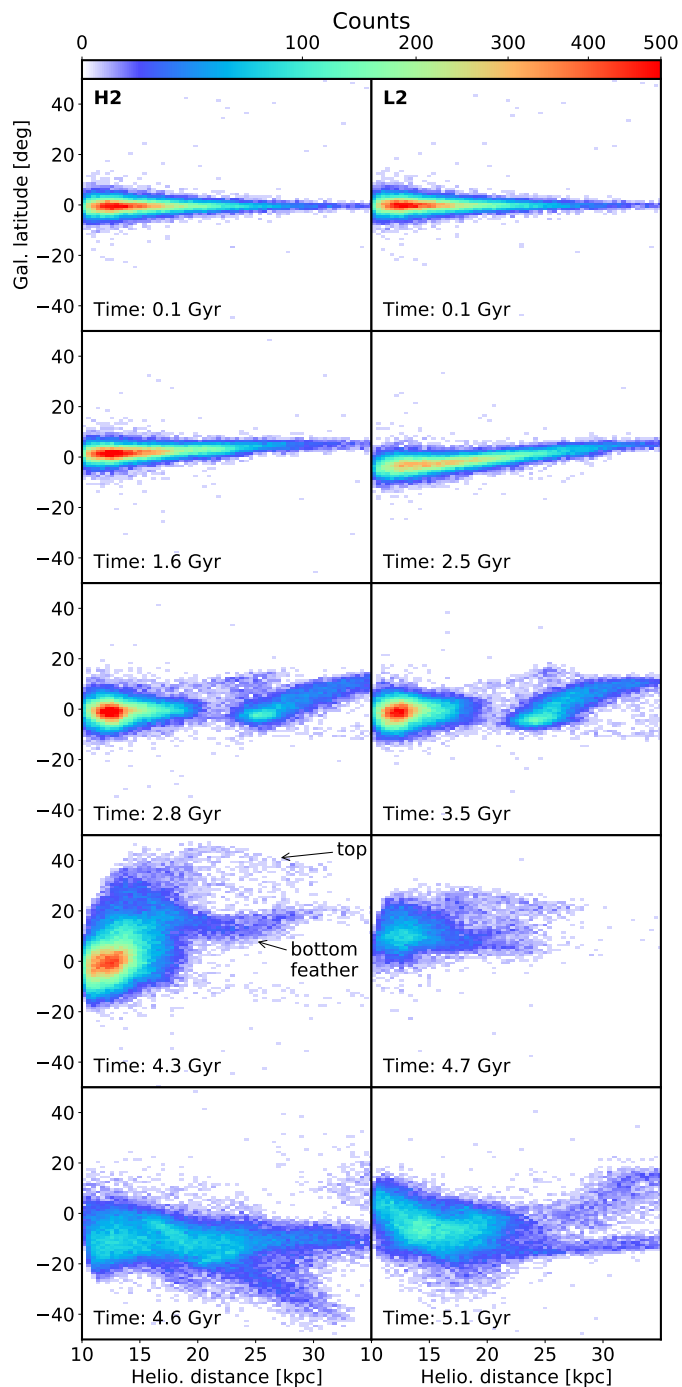
Although recent work favours a disc origin for these structures, the debate is still ongoing, with the alternative being that these

structures are the tidal debris of an accreted MW satellite. Based solely on the morphology that our method allows us to observe, both MNC and the ACS could very well be different wraps of the same tidal stream. If that were the case, then we should be able to see at least a hint of continuity in the south, unless the tails only emerge from behind the disc at Galactic longitudes where the disc is already too dense for our method to detect them. Peñarrubia et al. (2005) presented an  $N$ -body model fitted to the observations of MNC available at the time, which was later used by Slater et al. (2014) to show that there is a broad agreement with the PanSTARRS-1 (Chambers et al. 2016) data. We find that the arch described by the debris generated with their model is too wide to explain MNC as we detected it<sup>13</sup>, but we note the presence of a tail (top-right panel of Fig. 5 from Slater et al. 2014) that resembles the ACS. Comparing the morphology now with the  $N$ -body simulation of Laporte et al. (2019b)<sup>14</sup>, where these ‘feathers’ appear as a result of the interaction with Sgr, we note that the overall agreement is good (top-left panel of their Fig. 1): for instance, the difference in latitude between the two structures and the fact that the one on top stops abruptly at a given longitude. Nonetheless, the feather corresponding to the ACS obtained with their  $N$ -body simulations is thicker than the observed one and does not present a higher density of stars close to the turning point of the vertical oscillation (the highest point in latitude) as we see in the data. However, this could simply be due to resolution limitations and the fact that these simulations were not meant to be an exact match to the MW.

To explore the simulations of Laporte et al. (2018) a little more deeply, we show in Fig. 11 the plane of Galactic latitude against the heliocentric distance of the particles that are

<sup>13</sup> However, the shape that we recover could be affected by the *Gaia* scanning law, as we mentioned in Sect. 4.1.

<sup>14</sup> We tried to also compare with the simulations of Kazantzidis et al. (2008, Fig. 5) and Gómez et al. (2016, Fig. 4), but there are too few particles to make a good assessment given the level of detail that the data provide.



**Fig. 11.** Heliocentric distance against Galactic latitude for different snapshots of the simulations by Laporte et al. (2018). On the *left* is the H2 simulation, and on the *right* is the L2. In both cases, the Sun has been placed at  $(x = -8 \text{ kpc}, y = 0 \text{ kpc})$ . The stars shown in each panel correspond to those in slices of Galactic longitude  $120^\circ < l < 240^\circ$ . The snapshot at 4.3 Gyr for H2 corresponds to the one where ACS- and MNC-like structures (top and bottom feathers, respectively) were reported in Laporte et al. (2019b).

at galactocentric radii  $>18 \text{ kpc}$ . To do that, we located the Sun at  $(x, y) = (-8, 0) \text{ kpc}$  and looked at the particles with  $120^\circ < l < 240^\circ$ . The snapshots were chosen such that we can see the outer disc at the beginning of the simulation (first row), right after the first pericentric passage of Sgr (second), at the time of first apocentre (third), right after the second pericentre (fourth), and, finally, at the second apocentre (fifth). The first thing we

note is that – while the first pericentre passage seems to have a similar effect in both simulations, causing some material to be ejected from the system (third row) – the difference becomes more noticeable with the second passage. In particular, we can see how in the H2 simulation a couple of feathers appear in the fourth row, with roughly constant declination ( $\sim 20^\circ$  and  $\sim 40^\circ$ ) and a large extension in distance. They correspond to the MNC- and ACS-like structures reported in Laporte et al. (2019b). The shape that we observe for MNC in the *Gaia* data seems more concentrated in distance while being more extended in latitude (see Fig. 7). Nevertheless, a more detailed analysis of the distances in our data, as mentioned in previous sections, is still necessary to produce a more precise 3D characterisation of MNC and the ACS.

Secondly, we note that, even at timescales of  $\sim 100 \text{ Myr}$ , the large-scale distribution of stars in the whole outer disc of the simulated MW changes significantly with time in both the radial and vertical directions. In turn, this would suggest that the phase-space configuration of the outer disc and of its feathers at the present time can in principle pose strong constraints on the time evolution of the perturbation (assuming that they were caused by a single perturber). The data that we have obtained in this work, being precise, continuous, and covering a large range of longitudes and latitudes, can be used for describing MNC and the ACS with analytical or semi-analytical models, such as the analytic method presented in Weinberg (1998). This efficient way of exploring the parameter space could be used to obtain a quantitative measure of the goodness of fit for each model, beyond the small set of simulations analysed here. Such an exercise is crucial as it would allow us to predict where the continuation of these structures should appear, both in Galactic coordinates as well as in distance, providing a way to actively search for them. More importantly, we would be able to quantify the mass of the perturber and the history of its orbit. We could even, as mentioned in Laporte et al. (2019b), use MNC and the ACS to constrain the Galactic potential: the rotation curve at that distance, its slope, the shape of the dark matter halo, etc.

Another option for objectively comparing models is to add particles generated with  $N$ -body models able to reproduce the observed MNC and the ACS, to a mock catalogue of the Galaxy. In this work we used the Rybizki et al. (2018) catalogue as an example of a Galaxy without substructure, but we now know that it was not very representative of the MW and, also, that it underestimated the observational errors (Rybizki et al. 2020). The recent Rybizki et al. (2020) catalogue has fixed much of these issues and provides a mock to use with the upcoming *Gaia* EDR3 (Brown 2019). With this approach, and taking the nuances of our methodology into account, we can attempt a quantitative comparison with the models. One of the key parameters for generating the  $N$ -body models would be the stellar mass contained inside these structures, which is currently poorly measured (Morganson et al. 2016). In this work, we have attempted to quantify the fraction of the disc that is within MNC and the ACS in Fig. 6, but it is just a rough estimate based solely on the number of giants.

We also studied the kinematic information obtained with the proper motion of the peaks, and we see that MNC and the ACS rotate slightly more slowly than the disc at the solar position, in agreement with de Boer et al. (2018). Nevertheless, this alone does not prove that these structures were once part of the disc. Since the two mechanisms proposed (extra-galactic or internal) have distinct formation timescales and chemical properties, we can more easily distinguish between them by exploring their CMDs. In this sense, the ratio of RR Lyrae to M giants that we

find (<5%) is unlikely for structures composed of tidal debris from an accreted satellite. This adds to the recent studies of Bergemann et al. (2018) and Laporte et al. (2020), who show that the abundances, the distribution in the [Mg/Fe] versus [Fe/H] plane, and the mean metallicity of MNC and the ACS are inconsistent with the extragalactic scenario.

If indeed MNC and the ACS were once part of the disc, then we can use them as chemical fossils, an idea already put forward in Laporte et al. (2020). After they were kicked out of the disc, the stellar formation of these structures most likely came to a halt as any gas that initially accompanied the stars must have quickly settled back onto the disc thanks to its efficient energy dissipation mechanisms. As a result, their current population is a frozen relic of the outskirts of the MW at the time when the perturbation occurred. With enough spectroscopic abundances, we could learn about the gas that dwelled at the edge of our Galaxy some gigayears ago and use that information to constrain the chemo-dynamical models of the MW.

## 6. Conclusions

The application of the WT to the proper motion space has proven extremely useful for revealing the kinematic substructure of the halo and outer disc. By removing most of the foreground with a simple yet effective cut in parallax, our method is able to efficiently detect kinematic substructure in the halo and even in external galaxies such as M33 or the Magellanic Clouds, several dwarf spheroidal galaxies, and dozens of globular clusters, as well as the Sgr stream. It has also revealed the sharpest picture of the anticentre, with MNC and the ACS appearing as the third most prominent structures in the distant sky (only after the Magellanic Clouds and Sgr).

We have been able to blindly detect the whole MNC North as well as the ACS from  $l \sim 120^\circ$  to  $l \sim 230^\circ$ . Our findings are in good agreement with previous studies such as Laporte et al. (2020), who also used DR2 data to investigate these structures. Nevertheless, we have been able to characterise their morphology with great detail, which is crucial for obtaining the orbital parameters of these groups of stars. We observe MNC with an arch-like shape, broader at small longitudes and becoming thinner towards larger longitudes. Nevertheless, the RC stars that we selected can be seen to span a wider range of latitudes, and therefore a detailed study of the selection function of *Gaia* and the extinction is needed to confirm how much of this shape is caused by the scanning law. The ACS can be seen at larger latitudes than MNC throughout the whole longitude range where we detect them; the ACS has a maximum of relative intensity when it reaches the highest latitude at  $l \sim 140^\circ$  (consistent with a pileup of stars at the maximum height in the orbit) and stops abruptly at a longitude of  $\sim 110^\circ$ . This behaviour, added to the fact that we do not observe a clear continuity in the south, favours the perturbative scenario proposed by Ibata et al. (2003) and supported by the simulations of many other authors (e.g., Gómez et al. 2016; Laporte et al. 2019b). Moreover, the kinematics of these features, which differ from the bulk motion of the disc stars that lie in front, are compatible with a low eccentricity orbit at  $\sim 10$  kpc that rotates similarly to the disc.

By analysing the apparent magnitude of the RC stars selected by proper motion, we have been able to trace MNC down to a latitude of  $\sim 5^\circ$ , closer to the disc than ever before. Also, by measuring the median apparent magnitude of the RC stars of each structure and converting to heliocentric distance, we have determined that the ACS ( $\sim 11.7$  kpc) is roughly 1 kpc farther away from the Sun than MNC ( $\sim 10.6$  kpc). This actually means that

both structures are at roughly the same galactocentric radius (but at heights above the disc of, respectively,  $\sim 6.5$  kpc and  $\sim 4.5$  kpc). We have also shown that MNC and the ACS, despite being different structures, are extended in distance and in the sky, and their tails overlap both in the 3D physical space and in kinematic space.

In the south, we found a diffuse population of giants at  $130^\circ < l < 150^\circ$  and  $190^\circ < l < 210^\circ$ , coinciding with the regions of low extinction, which we do not observe in the mock catalogue nor in the north. Their apparent magnitudes span the range  $14 < G < 15.5$  mag, which implies distances for an RC star not affected by extinction between 5 and 10 kpc. These could be related to the vertical wave described in Xu et al. (2015) and are most likely the so-called MNC South that was first reported by Ibata et al. (2003). On the other hand, in the longitude range  $130^\circ < l < 150^\circ$ , we observed a faint trace of an RC at  $G \sim 17$  mag (heliocentric distance  $\sim 16$  kpc) that most likely corresponds to the TriAnd overdensity. Nevertheless, due to the contamination of nearby stars close to the disc and the large distance uncertainties, we have not been able to explore the morphological connection between these structures and MNC North.

Studies like this will benefit the most from the next *Gaia* release (EDR3), which is expected to contain proper motions twice as precise (on average) and an increased number of stars at the faintest magnitudes (Brown 2019). As a result, the structures that we detect will become more concentrated in the proper motion space and in turn produce stronger signals in our maps of relative intensity. Furthermore, the effects of the scanning law should diminish as a result of the additional year of observations. With it, we should be able to remove the foreground contamination more efficiently and detect the anticentre structures continuously at all latitudes, providing a direct observation of their 3D morphology. Other important surveys regarding the anticentre will be the all-sky, multi-epoch spectroscopic SDSS-V survey (Kollmeier et al. 2017) and the Subaru Prime Focus Spectrograph (PFS, Takada et al. 2014), which targets MSTOs in the anticentre region (operation phase planned for 2023). Moreover, the WEAVE spectroscopic survey (Dalton et al. 2014), with the first light expected for early 2021, will also observe this region; as such, we could potentially obtain radial velocities and abundances for a large fraction of the giants in our sample. This means that we will be able to trace MNC and the ACS, and probably even TriAnd, more clearly and deeply and obtain a less contaminated sample of members.

The challenge now is to find a way to use our data to form a coherent and unified picture of the outer disc, constraining the properties of the different agents involved (the MW disc, the dark matter halo, and Sgr). Any method that we build, though, must also be ready to work with the future samples that will soon become available and which will bring more and better data of the Galactic anticentre.

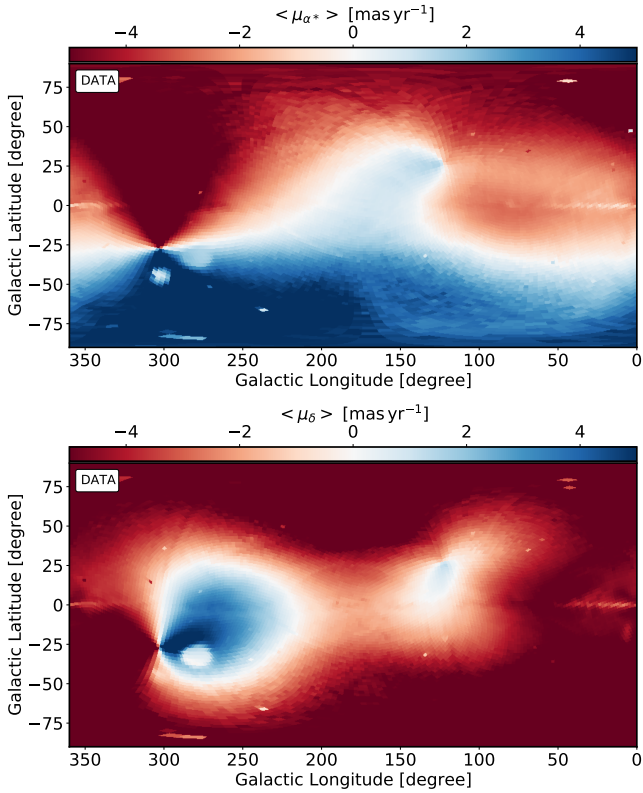
*Acknowledgements.* We thank the referee for her/his comments. Also to the language editor, for her/his thorough revision of the manuscript. This work has made use of data from the European Space Agency (ESA) mission *Gaia* (<https://www.cosmos.esa.int/gaia>), processed by the *Gaia* Data Processing and Analysis Consortium (DPAC, <https://www.cosmos.esa.int/web/gaia/dpac/consortium>). Funding for the DPAC has been provided by national institutions, in particular the institutions participating in the *Gaia* Multilateral Agreement. This work has been supported by the Agence Nationale de la Recherche (ANR project SEGAL ANR-19-CE31-0017). It has also received funding from the project ANR-18-CE31-0006 and from the European Research Council (ERC grant agreement No. 834148 and No. 852839). This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 745617 and No. 800502. This work used the Extreme Science and

Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number OCI-1053575. This work was supported in part by World Premier International Research Center Initiative (WPI Initiative), MEXT, Japan. This work was supported by the MINECO (Spanish Ministry of Economy) through grant ESP2016-80079-C2-1-R and RTI2018-095076-B-C21 (MINECO/FEDER, UE), and MDM-2014-0369 of ICCUB (Unidad de Excelencia ‘María de Maeztu’). This project has received support from the DGAPA/UNAM PAPIIT program grant IG100319. CM is grateful for the hospitality of the ICCUB, during visits in which part of this research was carried out.

## References

- Anders, F., Khalatyan, A., Chiappini, C., et al. 2019, *A&A*, **628**, A94
- Antoja, T., Monari, G., Helmi, A., et al. 2015, *ApJ*, **800**, L32
- Antoja, T., Helmi, A., Romero-Gómez, M., et al. 2018, *Nature*, **561**, 360
- Antoja, T., Ramos, P., Mateu, C., et al. 2020, *A&A*, **635**, L3
- Baumgardt, H., Hilker, M., Sollima, A., & Bellini, A. 2019, *MNRAS*, **482**, 5138
- Belokurov, V., Zucker, D. B., Evans, N. W., et al. 2006, *ApJ*, **642**, L137
- Belokurov, V., Zucker, D. B., Evans, N. W., et al. 2007, *ApJ*, **654**, 897
- Belokurov, V., Erkal, D., Evans, N. W., Koposov, S. E., & Deason, A. J. 2018, *MNRAS*, **478**, 611
- Bergemann, M., Sesar, B., Cohen, J. G., et al. 2018, *Nature*, **555**, 334
- Bica, E., Pavani, D. B., Bonatto, C. J., & Lima, E. F. 2019, *AJ*, **157**, 12
- Brown, A. G. A. 2019, <https://doi.org/10.5281/zenodo.2637972>
- Carballo-Bello, J. A., Martínez-Delgado, D., Corral-Santana, J. M., et al. 2021, *MNRAS*, **501**, 1690
- Castro-Ginard, A., Jordi, C., Luri, X., et al. 2018, *A&A*, **618**, A59
- Castro-Ginard, A., Jordi, C., Luri, X., et al. 2020, *A&A*, **635**, A45
- Chambers, K. C., Magnier, E. A., Metcalfe, N., et al. 2016, ArXiv e-prints [arXiv:1612.05560]
- Clementini, G., Ripepi, V., Molinaro, R., et al. 2019, *A&A*, **622**, A60
- Conn, B. C., Martin, N. F., Lewis, G. F., et al. 2005, *MNRAS*, **364**, L13
- Conn, B. C., Lane, R. R., Lewis, G. F., et al. 2007, *MNRAS*, **376**, 939
- Conn, B. C., Lane, R. R., Lewis, G. F., et al. 2008, *MNRAS*, **390**, 1388
- Crane, J. D., Majewski, S. R., Rocha-Pinto, H. J., et al. 2003, *ApJ*, **594**, L119
- Dalton, G., Trager, S., Abrams, D. C., et al. 2014, in *Groundbased and Airborne Instrumentation for Astronomy V*, eds. S. K. Ramsay, I. S. McLean, & H. Takami, *SPIE Conf. Ser.*, **9147**, 91470L
- de Boer, T. J. L., Belokurov, V., & Koposov, S. E. 2018, *MNRAS*, **473**, 647
- Deason, A. J., Belokurov, V., Koposov, S. E., & Lancaster, L. 2018, *ApJ*, **862**, L1
- Forbes, D. A., Lasky, P., Graham, A. W., & Spitler, L. 2008, *MNRAS*, **389**, 1924
- Fragkoudi, F., Katz, D., Trick, W., et al. 2019, *MNRAS*, **488**, 3324
- Gaia Collaboration (Prusti, T., et al.) 2016, *A&A*, **595**, A1
- Gaia Collaboration (Brown, A. G. A., et al.) 2018a, *A&A*, **616**, A1
- Gaia Collaboration (Babusiaux, C., et al.) 2018b, *A&A*, **616**, A10
- Gaia Collaboration (Helmi, A., et al.) 2018c, *A&A*, **616**, A12
- Gómez, F. A., Minchev, I., O’Shea, B. W., et al. 2013, *MNRAS*, **429**, 159
- Gómez, F. A., White, S. D. M., Marinacci, F., et al. 2016, *MNRAS*, **456**, 2779
- GRAVITY Collaboration (Abuter, R., et al.) 2019, *A&A*, **625**, L10
- Grillmair, C. J. 2006, *ApJ*, **651**, L29
- Grillmair, C. J. 2011, *ApJ*, **738**, 98
- Guglielmo, M., Lane, R. R., Conn, B. C., et al. 2018, *MNRAS*, **474**, 4584
- Haywood, M., Di Matteo, P., Lehnert, M. D., et al. 2018, *ApJ*, **863**, 113
- Helmi, A., Navarro, J. F., Meza, A., Steinmetz, M., & Eke, V. R. 2003, *ApJ*, **592**, L25
- Helmi, A., Babusiaux, C., Koppelman, H. H., et al. 2018, *Nature*, **563**, 85
- Holl, B., Audard, M., Nienartowicz, K., et al. 2018, *A&A*, **618**, A30
- Ibata, R. A., Gilmore, G., & Irwin, M. J. 1994, *Nature*, **370**, 194
- Ibata, R. A., Irwin, M. J., Lewis, G. F., Ferguson, A. M. N., & Tanvir, N. 2003, *MNRAS*, **340**, L21
- Ibata, R. A., Malhan, K., & Martin, N. F. 2019, *ApJ*, **872**, 152
- Ibata, R., Bellazzini, M., Thomas, G., et al. 2020, *ApJ*, **891**, L19
- Ivezic, Ž., Sesar, B., Jurić, M., et al. 2008, *ApJ*, **684**, 287
- Jayasinghe, T., Stanek, K. Z., Kochanek, C. S., et al. 2019, *MNRAS*, **486**, 1907
- Kawata, D., Baba, J., Ciuca, I., et al. 2018, *MNRAS*, **479**, L108
- Kazantzidis, S., Bullock, J. S., Zentner, A. R., Kravtsov, A. V., & Moustakas, L. A. 2008, *ApJ*, **688**, 254
- Khanna, S., Sharma, S., Tepper-García, T., et al. 2019, *MNRAS*, **489**, 4962
- Kinman, T. D., Saha, A., & Pier, J. R. 2004, *ApJ*, **605**, L25
- Kollmeier, J. A., Zasowski, G., Rix, H. W., et al. 2017, ArXiv e-prints [arXiv:1711.03234]
- Koposov, S. E., Belokurov, V., Torrealba, G., & Wyn Evans, N. 2015, *ApJ*, **805**, 130
- Koppelman, H. H., & Helmi, A. 2021, *A&A*, **645**, A69
- Laporte, C. F. P., Johnston, K. V., Gómez, F. A., Garavito-Camargo, N., & Besla, G. 2018, *MNRAS*, **481**, 286
- Laporte, C. F. P., Minchev, I., Johnston, K. V., & Gómez, F. A. 2019a, *MNRAS*, **485**, 3134
- Laporte, C. F. P., Johnston, K. V., & Tzanidakis, A. 2019b, *MNRAS*, **483**, 1427
- Laporte, C. F. P., Belokurov, V., Koposov, S. E., Smith, M. C., & Hill, V. 2020, *MNRAS*, **492**, L61
- Liu, C., Xu, Y., Wan, J.-C., et al. 2017, *Res. Astron. Astrophys.*, **17**, 096
- López-Corrodoira, M., & Sylos Labini, F. 2019, *A&A*, **621**, A48
- López-Corrodoira, M., Garzón, F., Wang, H. F., et al. 2020, *A&A*, **634**, A66
- Mafz Apellániz, J., & Weiler, M. 2018, *A&A*, **619**, A180
- Majewski, S. R., Skrutskie, M. F., Weinberg, M. D., & Ostheimer, J. C. 2003, *ApJ*, **599**, 1082
- Majewski, S. R., Ostheimer, J. C., Rocha-Pinto, H. J., et al. 2004, *ApJ*, **615**, 738
- Malhan, K., Ibata, R. A., & Martin, N. F. 2018, *MNRAS*, **481**, 3442
- Martin, N. F., Ibata, R. A., Bellazzini, M., et al. 2004, *MNRAS*, **348**, 12
- Martin, N. F., Ibata, R. A., & Irwin, M. 2007, *ApJ*, **668**, L123
- Massari, D., & Helmi, A. 2018, *A&A*, **620**, A155
- Mateu, C., Holl, B., De Ridder, J., & Rimoldini, L. 2020, *MNRAS*, **496**, 3291
- Momany, Y., Zaggia, S., Gilmore, G., et al. 2006, *A&A*, **451**, 515
- Morganson, E., Conn, B., Rix, H.-W., et al. 2016, *ApJ*, **825**, 140
- Newberg, H. J., Yanny, B., Rockosi, C., et al. 2002, *ApJ*, **569**, 245
- Peñarrubia, J., Martínez-Delgado, D., Rix, H. W., et al. 2005, *ApJ*, **626**, 128
- Perotoni, H. D., Rocha-Pinto, H. J., Girardi, L., et al. 2018, *MNRAS*, **473**, 1461
- Price-Whelan, A. M., Johnston, K. V., Sheffield, A. A., Laporte, C. F. P., & Sesar, B. 2015, *MNRAS*, **452**, 676
- Purcell, C. W., Bullock, J. S., Tollerud, E. J., Rocha, M., & Chakrabarti, S. 2011, *Nature*, **477**, 301
- Ramos, P., Antoja, T., & Figueras, F. 2018, *A&A*, **619**, A72
- Ramos, P., Mateu, C., Antoja, T., et al. 2020, *A&A*, **638**, A104
- Reid, M. J., & Brunthaler, A. 2020, *ApJ*, **892**, 39
- Rocha-Pinto, H. J., Majewski, S. R., Skrutskie, M. F., Crane, J. D., & Patterson, R. J. 2004, *ApJ*, **615**, 732
- Rocha-Pinto, H. J., Majewski, S. R., Skrutskie, M. F., et al. 2006, *ApJ*, **640**, L147
- Romero-Gómez, M., Mateu, C., Aguilar, L., Figueras, F., & Castro-Ginard, A. 2019, *A&A*, **627**, A150
- Ruiz-Dern, L., Babusiaux, C., Arenou, F., Turon, C., & Lallement, R. 2018, *A&A*, **609**, A116
- Rybizki, J., Demleitner, M., Fouesneau, M., et al. 2018, *PASP*, **130**, 074101
- Rybizki, J., Demleitner, M., Bailer-Jones, C., et al. 2020, *PASP*, **132**, 074501
- Schlafly, E. F., & Finkbeiner, D. P. 2011, *ApJ*, **737**, 103
- Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, *ApJ*, **500**, 525
- Schönrich, R., & Dehnen, W. 2018, *MNRAS*, **478**, 3809
- Schönrich, R., Binney, J., & Dehnen, W. 2010, *MNRAS*, **403**, 1829
- Sheffield, A. A., Price-Whelan, A. M., Tzanidakis, A., et al. 2018, *ApJ*, **854**, 47
- Skrutskie, M. F., Cutri, R. M., Stiening, R., et al. 2006, *AJ*, **131**, 1163
- Slater, C. T., Bell, E. F., Schlafly, E. F., et al. 2014, *ApJ*, **791**, 9
- Starck, J.-L., & Murtagh, F. 2002, *Astronomical Image and Data Analysis* (Berlin: Springer)
- Takada, M., Ellis, R. S., Chiba, M., et al. 2014, *PASJ*, **66**, R1
- Thomas, G. F., Laporte, C. F. P., McConnachie, A. W., et al. 2019, *MNRAS*, **483**, 3119
- Torrealba, G., Belokurov, V., & Koposov, S. E. 2019, *MNRAS*, **484**, 2181
- van der Walt, S., Schönberger, J. L., Nunez-Iglesias, J., et al. 2014, *PeerJ*, **2**, e453
- Vivas, A. K., & Zinn, R. 2006, *AJ*, **132**, 714
- Wang, H.-F., Carlin, J. L., Huang, Y., et al. 2019, *ApJ*, **884**, 135
- Weinberg, M. D. 1998, *MNRAS*, **299**, 499
- Widrow, L. M., Gardner, S., Yanny, B., Dodelson, S., & Chen, H.-Y. 2012, *ApJ*, **750**, L41
- Wilhelm, R., Beers, T. C., Allende Prieto, C., Newberg, H. J., & Yanny, B. 2005, in *Cosmic Abundances as Records of Stellar Evolution and Nucleosynthesis*, eds. I. Barnes, G. Thomas, & F. N. Bash, *ASP Conf. Ser.*, **336**, 371
- Willman, B., Dalcanton, J. J., Martínez-Delgado, D., et al. 2005, *ApJ*, **626**, L85
- Xia, Q., Liu, C., Xu, Y., et al. 2015, *MNRAS*, **447**, 2367
- Xu, Y., Newberg, H. J., Carlin, J. L., et al. 2015, *ApJ*, **801**, A105
- Yanny, B., Newberg, H. J., Grebel, E. K., et al. 2003, *ApJ*, **588**, 824
- York, D. G., Adelman, J., Anderson, J. E. J., et al. 2000, *AJ*, **120**, 1579
- Younger, J. D., Besla, G., Cox, T. J., et al. 2008, *ApJ*, **676**, L21

## Appendix A: Queries to the *Gaia* archive



**Fig. A.1.** Average proper motion of the sample in the sky. Only the stars selected according to the criteria set in Sect. 2 are used. *Top*: proper motion in right ascension. *Bottom*: proper motion in declination. Only a few globular clusters, the Large Magellanic cloud, and the most inner parts of the Sgr core are noticeable as the rest of the field is dominated by the solar reflex (peculiar motion of the Sun and rotation of the local standard of rest).

To obtain the mean quantities used in this work all around the sky and in a single file, we used the following query to the *Gaia* Archive:

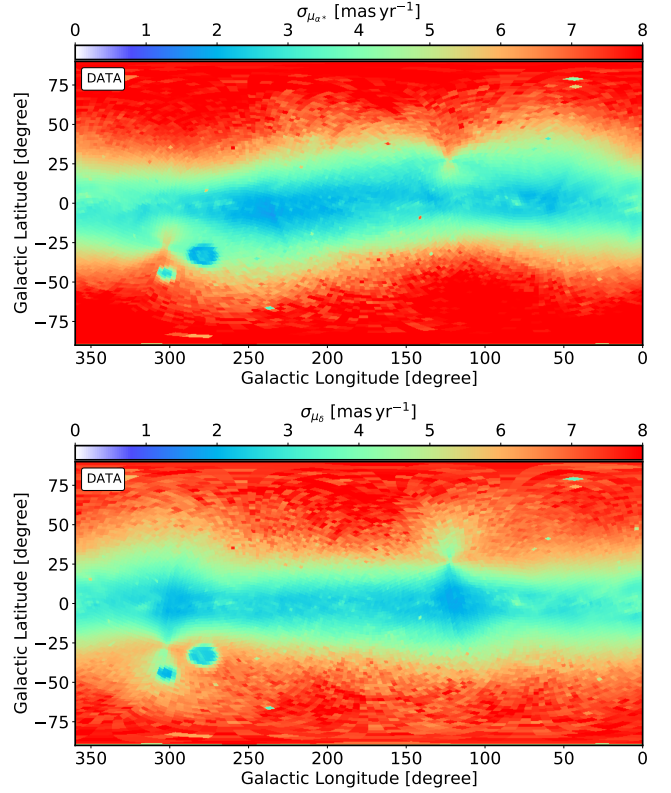
```
SELECT gaia_healpix_index(5, source_id) AS healpix_5,
count(*) as N, avg(astrometric_n_good_obs_al) AS
avg_n_good_al, avg(astrometric_gof_al) AS
avg_gof_al, avg(astrometric_excess_noise) as,
avg_excess_noise, avg(bp_rp) as avg_bprp,
avg(phot_g_mean_mag) as avg_g, avg(pmra) as
avg_pmra, avg(pmdec) as avg_pmdec,
avg(pmra_error) as avg_pmra_error,
avg(pmdec_error) as avg_pmdec_error
FROM gaiadr2.gaia_source WHERE
parallax-parallax_error < 0.1 AND BP_RP>0.2
GROUP BY healpix_5
```

Computing the standard deviations is not so straightforward, though, as there is no implemented function to do so in ADQL. Instead, we repeated the same query but this time using the average squared of the quantities. A simple subtraction of both tables returns the standard deviation:

$$\sigma = \sqrt{E[X^2] - (E[X])^2}, \quad (\text{A.1})$$

where  $E[X]$  is the mean of a vector  $X$ .

Figures A.1 and A.2 contain, respectively, the average proper motion and its dispersion for both components, right ascension



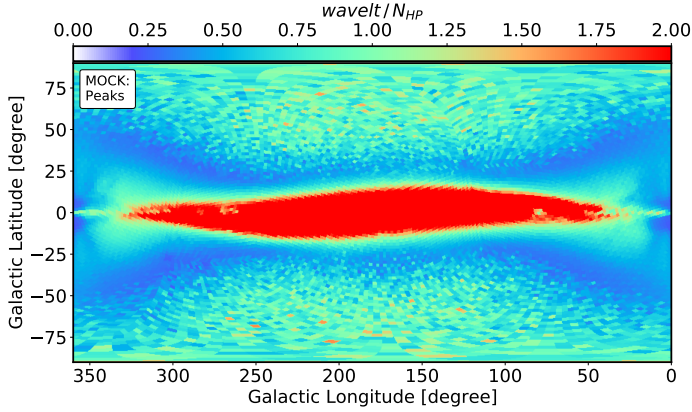
**Fig. A.2.** Dispersion in proper motion of the sample in the sky. *Top*: proper motion in right ascension. *Bottom*: proper motion in declination.

and declination. With the mean proper motions, we basically see a kinematic field dominated by the combination of the solar peculiar motion and the rotation of the Galaxy. Given the large amount of nearby stars that our filter (Eq. (1)) cannot remove, only the very dense structures can be seen here: the Magellanic Clouds, some globular clusters, and, although very faint, the Sgr stream. To better understand the extent of the disc in our sample, we used the relation between the kinematic dispersion and the kinetic temperature of each population (disc vs. halo). We note a transition at  $\sim 30^\circ$  in latitude from a rather cold population (disc) to a hotter one (halo).

## Appendix B: *Gaia* DR2 mock catalogue

To test the signal that we would expect to see in a galaxy without substructure, we ran our whole method on the proper motion histograms obtained from a mock catalogue. For that, we queried the Rybizki et al. (2018) catalogue and downloaded, for each HEALpix, up to 2 000 000 stars. Among the quantities available for download, we selected all the astrometry and photometry, as well as the age of the stars, which we then used to separate the stars into thin disc, thick disc, and halo categories. We then drew for each star one realisation from a normal distribution centred on the true values and with a dispersion equal to the provided observational uncertainties to produce the mock particles. In the case of the colours, we applied the errors to the fluxes and then converted the observed fluxes to magnitudes using the equations published on the *Gaia* web page<sup>15</sup> (Maíz Apellániz & Weiler 2018). We note that these are meant to be used with synthetic fluxes derived using the same passbands, which is not the case

<sup>15</sup> <https://www.cosmos.esa.int/web/gaia/dr2-known-issues>

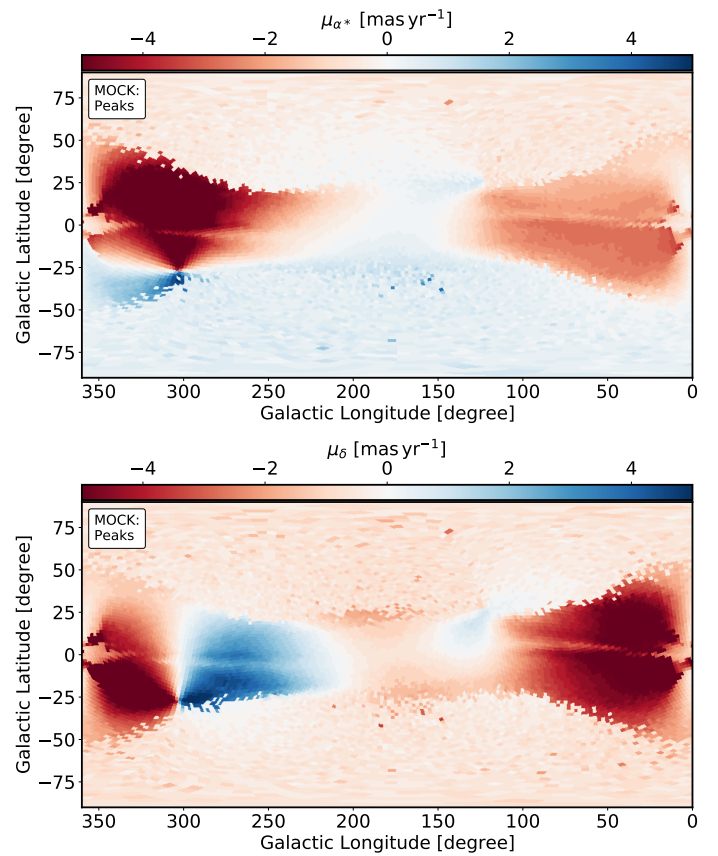


**Fig. B.1.** Relative intensity of the dominant structure in the proper motion plane of each HEALpix for the mock catalogue.

here; however, since we only used the photometry for reference (the cut in colour has little impact), we do not need a perfect match with reality.

Once we had the mock particles, we applied the cuts in parallax and colour described in Sect. 2, generated the proper motion histograms, and analysed them with the WT in the same manner as we did for the data. The value of the WT coefficient is sensitive to the absolute number of counts and, as a consequence, we scaled the histograms such that the sum of all the bins equals the number of stars observed in that same HEALpix with *Gaia*. No substructure was added in this way since a scaling of the histogram does not bias the centroid, and, whenever we show the coefficients, we do it after normalising by the number of stars in the HEALpix.

Figure B.1 shows the result of applying our methodology to the mock catalogue. The only structure present is the geometrical warp introduced in the underlying model of the galaxy. Apart from that, we note a sharp transition between the disc and the halo, noticeable as a drastic change in the relative intensity. There is also a change between the thin disc that dominates the anticentre and the thick disc that dominates the central parts of the mock MW. If we analyse the proper motions of the obtained peaks (Fig. B.2), we recover the reflex of the solar motion, with the location of the poles of the equatorial sphere clearly visible as singular points, as well as the perspective effect caused by the rotation, which introduces a gradient in the proper motions

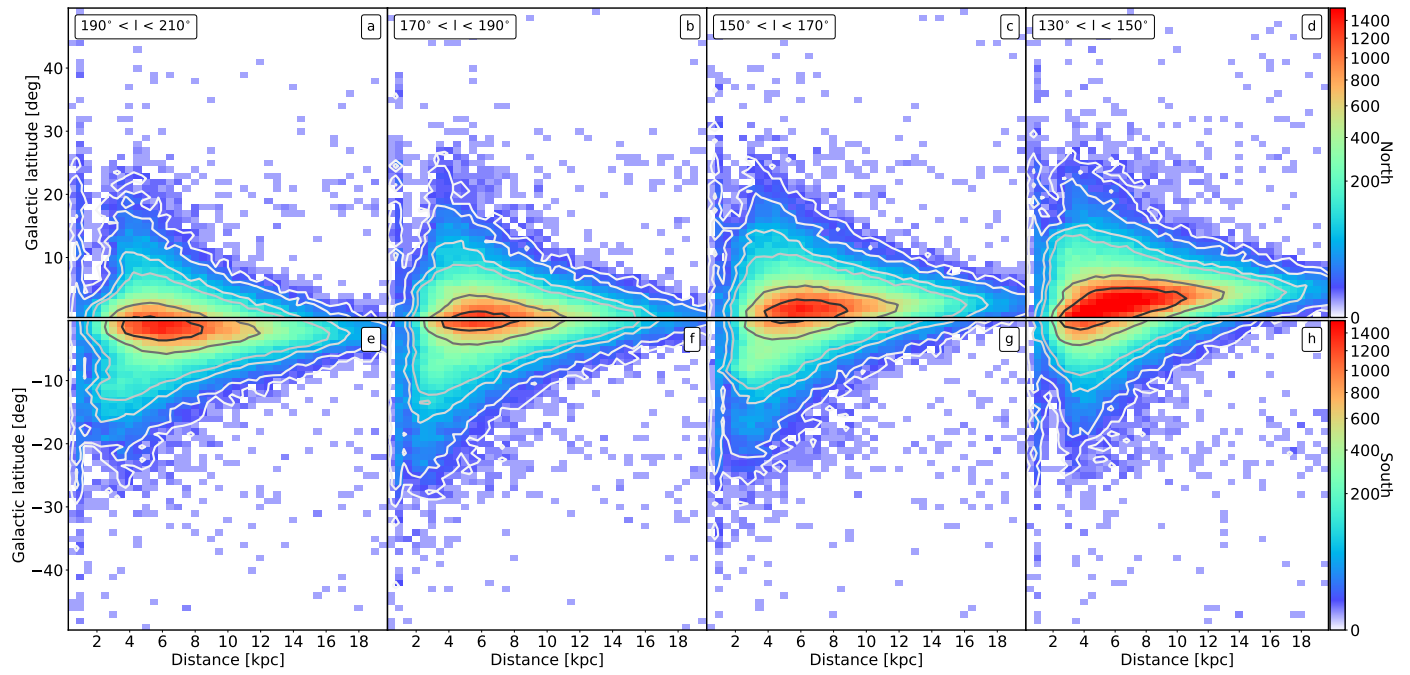


**Fig. B.2.** Same as Fig. 3 but for the mock catalogue.

with Galactic longitude and latitude. This is the reason why the transition to the halo is so sharp: Once the dominant structure in the proper motion plane is the halo, we observe the reflex of a non-rotating stellar system that only has four lobes instead of the eight present in the case of the disc.

In some of the figures, for example Figs. 5 or B.3, we did not select the particles according to the peak obtained with the mock but with the coordinates of the peaks detected in the data. In doing so, we can check what is the distribution in the CMD and what is the distribution in the distance of the particles that have the observed kinematics.





**Fig. B.3.** Same as Fig. 10 but for the mock catalogue. We show only the particles in the mock catalogue that fall inside the peaks detected in the data.