Reassessing the proper motions of M31/M33 with Gaia DR3

Samuel Rusterucci, Nicolas F. Martin, Else Starkenburg, Rodrigo Ibata

A&A, 692, A30 (2024)
https://doi.org/10.1051/0004-6361/202452281

Astronomy & Astrophysics

Reassessing the proper motions of M31/M33 with *Gaia* DR3

Unravelling systematic uncertainties

Samuel Rusterucci[1,2,⋆], Nicolas F. Martin[1,3], Else Starkenburg[2], and Rodrigo Ibata[1]

[1] Université de Strasbourg, CNRS, Observatoire astronomique de Strasbourg, UMR 7550, 67000 Strasbourg, France
[2] Kapteyn Astronomical Institute, University of Groningen, Landleven 12, 9747 AD Groningen, The Netherlands
[3] Max-Planck-Institut für Astronomie, Königstuhl 17, 69117 Heidelberg, Germany

Received 17 September 2024 / Accepted 21 October 2024

ABSTRACT

We provide an updated inference of the proper motion of M31 using the *Gaia* DR3 proper motions of bright stars from the disc of M31. By refining the motion of the quasar reference frame, and statistically accounting for the variations in the inferred proper motions obtained across different regions of M31, we demonstrate that these inconsistencies most likely arise from systematic uncertainties. Our updated favoured values for the proper motion of M31 are $46.9 \pm 11.7(\text{stat}) \pm 50.6(\text{sys})$ μas yr^{-1} in the right ascension direction and $-29.1 \pm 9.4(\text{stat}) \pm 35.6(\text{sys})$ μas yr^{-1} in the declination direction, the systematics being determined at a 90% confidence level (the values for M33 are given in the paper). This clearly highlights that the systematics are the dominant source of uncertainty, their magnitudes being comparable to the proper motion of M31 itself. The analysis conducted using *Gaia* DR2 instead of DR3 revealed that a net reduction in these systematic uncertainties occurred between the two data releases. If similar progress is made with the upcoming DR4, the future *Gaia*-based estimates could match the uncertainty level of HST, and could be used to refine the dynamics and history of M31 and M33.

Key words. galaxies: kinematics and dynamics – Local Group – proper motion

1. Introduction

The proper motion of the Andromeda galaxy (M31) is an essential piece of information to constrain the past, present, and future of the Local Group. While it is often assumed that most of the motion of M31 is carried by its radial velocity component (e.g. the timing argument; Kahn & Woltjer 1959; Li & White 2008; Peñarrubia et al. 2016), useful measurements of its proper motion were only made possible over the last two decades from the use of indirect methods or significant advances in astrometric measurements.

Indirect measurements of the proper motion of M31[1] are primarily constrained by the radial velocities of its satellites, and were measured to be $(\mu_\alpha^{\text{M31}}, \mu_\delta^{\text{M31}}) = (21.5 \pm 11.1, -10.4 \pm 9.3)$ μas yr^{-1} (van der Marel & Guhathakurta 2008) and $(9.1 \pm 19.0, 5.6 \pm 16.3)$ μas yr^{-1} (Salomon et al. 2016). These results, which are compatible with each other, when translated into the galactocentric frame of reference, display a substantial tangential motion that is not negligible compared to its radial counterpart. Nonetheless, these estimates rely on the assumption that, dynamically, M31 satellites are pressure-supported and virialised, which may well not be the case given the observed presence of a plane of co-rotating satellites that includes about half of the known dwarf galaxies of M31 (Ibata et al. 2013).

The direct measurements make use of the individual proper motions of stars within the disc of M31. Sohn et al. (2012) and van der Marel et al. (2012b) were the first to make such an attempt with highly accurate Hubble Space Telescope (HST) data from three deep fields ($3.37' \times 3.37'$), observed at two different epochs. Accounting for the internal kinematics of M31, they derived a weighted average of $(\mu_\alpha^{\text{M31}}, \mu_\delta^{\text{M31}}) = (44.1 \pm 12.7, -31.8 \pm 12.2)$ μas yr^{-1} for the three fields, only marginally changing the results they obtained without considering the intrinsic motions of the stars within M31. This result is in much better agreement with a nearly radial orbit, but less so with the results from the indirect methods. They also derived a value by combining their direct measurement with indirect measurements to further constrain the motion of M31. Finally, the studies of van der Marel et al. (2019) and Salomon et al. (2021, hereafter S21), which respectively used the data from *Gaia* DR2 and EDR3, concluded that the proper motion of M31 is nearly radial for the latter study, but less so for the former. However, there are incompatibilities between the two samples analysed by S21. Their blue giant sample, which they consider to be more reliable due to its larger size and lower contamination, gives a nearly radial motion. In contrast, the red giant sample does not support this conclusion, even though it is inconceivable that two populations belonging to the disc of M31 move in different directions. Unfortunately, only hints could be offered in an attempt to explain these discrepancies, prompting us to revisit the study by S21, this time using data from *Gaia* DR3, with a particular focus on refining the reference frame around M31.

Our motivation stems from the possibility of some systematic uncertainties on the *Gaia* celestial reference frame (CRF) that are unaccounted for at small scales. As is shown in Figure 13 from Lindegren et al. (2021b) and pointed out quantitatively in their Table 7, the CRF presents local uncertainties of the order of ±17.2 μas yr^{-1} for angular scales larger than 0.5°. At

⋆ Corresponding author; samuel.rusterucci@astro.unistra.fr
[1] Throughout this paper, all results are presented in the heliocentric reference frame to eliminate any dependency on the assumed motion of the Sun.

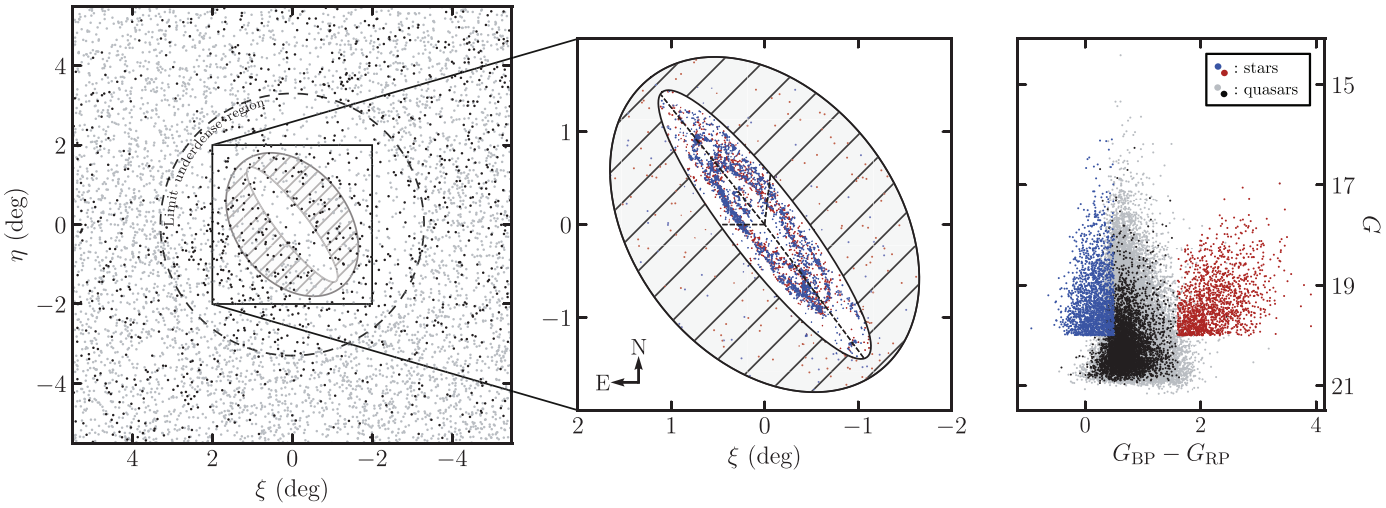

Fig. 1. Spatial and CMD locations of stars and quasars around M31. *Left-hand panel*: quasars from the *Gaia* CRF3 (in grey) and added by the latest *Gaia* catalogue of extra-galactic objects (in black) centred on M31. *Middle panel*: final blue and red samples of M31 stars superimposed on the central ellipse representing the geometrical cut made on M31. Each quadrant depicted by dashed lines contains the same number of blue stars and the hatched region is used to estimate the properties of the contamination. *Right-hand panel*: CMD of all the objects from the previous panels.

the distance of M31, this corresponds to $\pm 63.9\,\mathrm{km\,s^{-1}}$ (assuming $d_{\mathrm{M31}} = 785 \pm 25$ kpc, McConnachie et al. 2005), which is on the order of magnitude of the previously measured tangential velocities. Here, we aim to refine the CRF by complementing the catalogue of quasars used to produce the *Gaia* CRF, making use of the recently published extragalactic *Gaia* catalogue (Gaia Collaboration 2023a). Given the very small values of the proper motion of M31, we also aim to carefully assess the level of systematic uncertainties present in the *Gaia* data.

The structure of this paper is as follows. In Section 2, we describe our data and the statistical method used to infer the proper motion of M31. In Section 3, we derive an updated proper motion and estimate its systematic uncertainty. Finally, we discuss our results and conclude in Section 4.

2. Methods

2.1. Data selection

For the full analysis, we use the *Gaia* information provided in Gaia Collaboration (2023b).

Star samples: To isolate M31 disc stars in the *Gaia* catalogue, we have used the various spatial, colour-magnitude, proper motion[2], and *Gaia* quality cuts (they are the same as in S21). The purpose of these cuts is to remove the contamination from foreground Milky Way (MW) stars and also stars with poor astrometric solutions. This leads to two separate samples of blue (young giant candidates; B_{pm} sample) and red (older supergiant candidates; R_{pm} sample) stars containing 1867 and 1543 objects, respectively. These stars are displayed on the sky and in the colour-magnitude diagram (CMD) in Figure 1. Using the region surrounding the disc of M31 (hashed region in Figure 1), we estimate that the two samples suffer from only low levels of contamination, in agreement with the previous study (1.1% and 1.8%, respectively). The spatial distributions of the samples exhibit expected properties: stars from the blue sample are predominantly located in the ring of active star formation of the disc of M31 (Lewis et al. 2015), while the red stars are more sparsely distributed, but remain mainly confined to the region of the disc.

Quasar sample: Quasars used by the *Gaia* consortium to produce the CRF of the DR3 were selected from a large set of photometric and spectroscopic catalogues. These, however, may become heavily contaminated around nearby extended objects such as M31. In this case, only highly reliable quasars, confirmed through other means (e.g. VLBI observations), were kept to constrain the reference frame (Gaia Collaboration 2022). This leads to a drop in the density of quasars in the central ∼3° region around M31 (grey dots in the left-hand panel of Figure 1), which we can expect to locally lower the quality of the reference frame. To improve the latter, we complemented the set of quasars used to build the CRF with additional objects selected from the latest *Gaia* catalogue of extragalactic objects (Gaia Collaboration 2023a). In particular, we selected sources that are highly likely to be quasars[3], shown as black dots in Figure 1. These newly added sources increased the central density from $19\,\mathrm{deg^{-2}}$ to $34\,\mathrm{deg^{-2}}$, compared to an average of $42\,\mathrm{deg^{-2}}$ over the whole sky[4]. Our final sample consists of 26,741 quasar candidates.

2.2. Proper motion inference

We aim to infer the global proper motion of stars in the *Gaia* sample with a refined CRF. To do so, we assume that, for a given sample of objects (stars or quasars) of size n, we have a set of data points, $\mathcal{D} = \{\mathbf{d}_k\}_{1\leq k\leq n}$, with a datum, $\mathbf{d}_k$, defined by the *Gaia* position (α_k, δ_k) and proper motion $(\mu_{\alpha,k}, \mu_{\delta,k})$, along with their associated uncertainties[5] $(\delta\mu_{\alpha,k}, \delta\mu_{\delta,k})$ and correlation coefficient (ρ_k). That is: $\mathbf{d}_k = \{\alpha_k, \delta_k, \mu_{\alpha,k}, \mu_{\delta,k}, \delta\mu_{\alpha,k}, \delta\mu_{\delta,k}, \rho_k\}$. The likelihood of all n sources following a certain model specified by a

[2] We followed the convention used in the *Gaia* catalogue, noting $\mu^*_\alpha = \mu_\alpha \cos(\delta)$ as μ_α.

[3] Specifically, we select *Gaia* quasar candidates flagged as `astrometric_selection_flag == 1`, i.e. that have been identified with an estimated purity of 98% in Gaia Collaboration (2023a).

[4] We also checked the quasar catalogue of Hernitschek et al. (2016) based on the Pan-STARRS1 photometric variability of sources. Selecting likely quasars (`p_QSO>0.5`) leads to no additional sources compared to quasars already added from the *Gaia* extragalactic catalogue.

[5] Given the vanishingly small uncertainties on the position of stars, we only consider the proper motion uncertainties.

set of parameters, $\mathcal{P}$, is

$$P_{\rm tot}(\mathcal{D}|\mathcal{P}) = \prod_{k=1}^{n} P_k(\mathbf{d}_k|\mathcal{P}), \tag{1}$$

where $P_{\rm tot}(\mathcal{D}|\mathcal{P})$ is the sum of two probabilistic models, $P_{\rm pop}$ and $P_{\rm cont}$, which represent the main population of M31 stars and contaminants, respectively.

$$P_k(\mathbf{d}_k|\mathcal{P}) = (1-f_{\rm c})P_{\rm pop}(\mathbf{d}_k|\mathcal{P}^{\rm pop}) + f_{\rm c}P_{\rm cont}(\mathbf{d}_k|\mathcal{P}^{\rm cont}) \quad \text{with } \mathcal{P} = f_{\rm c} \cup \mathcal{P}^{\rm pop} \cup \mathcal{P}^{\rm cont}, \tag{2}$$

where the parameter $f_{\rm c}$ represents the fraction of contaminants.

We chose to model $P_{\rm pop}$ as a two-dimensional Gaussian in the (μ_α, μ_δ) space, centred on the mean motion given by parameters $\mathcal{P}^{\rm pop} = \{\mu_\alpha^{\rm pop}, \mu_\delta^{\rm pop}\}$. The width of the Gaussian in the model is assumed to be entirely driven by the proper motion uncertainties and we have foregone parameters that would represent the intrinsic dispersion of sources[6]. Therefore, we have

$$P_{\rm pop}(\mathbf{d}_k|\mathcal{P}^{\rm pop}) = \frac{1}{2\pi\delta\mu_{\alpha,k}\delta\mu_{\delta,k}\sqrt{(1-\rho_k^2)}} \times \exp\left[-\frac{1}{2(1-\rho_k^2)}\left(\frac{\Delta\mu_{\alpha,k}^2}{\delta\mu_{\alpha,k}^2} + \frac{\Delta\mu_{\delta,k}^2}{\delta\mu_{\delta,k}^2} - \frac{2\rho_k\Delta\mu_{\alpha,k}\Delta\mu_{\delta,k}}{\delta\mu_{\alpha,k}\delta\mu_{\delta,k}}\right)\right], \tag{3}$$

with $\Delta\mu_{\alpha,k} = \mu_{\alpha,k} - \mu_\alpha^{\rm pop}$ and $\Delta\mu_{\delta,k} = \mu_{\delta,k} - \mu_\delta^{\rm pop}$ being the offsets between a source's proper motion and the modelled motion of the considered population at this location. In the specific case of the stars of M31, a term accounting for the intrinsic motion of a star at position k was subtracted from $\Delta\mu_{\alpha/\delta,k}$. The value of this term was determined from the motion predicted by the disc model developed by Chemin et al. (2009) for a star located at (α_k, δ_k). In the case of the quasars, no additional term is needed as, on average, they are not expected to move.

We also chose $P_{\rm cont}$ to be modelled by a two-dimensional Gaussian. However, this time, the intrinsic dispersion of the contaminating stars cannot be assumed to be negligible anymore. The parameters are therefore the means and dispersions in the two-dimensional proper motion space; that is, $\mathcal{P}^{\rm cont} = \{\mu_\alpha^{\rm cont}, \mu_\delta^{\rm cont}, \sigma_\alpha^{\rm cont}, \sigma_\delta^{\rm cont}\}$. Therefore,

$$P_{\rm cont}(\mathbf{d}_k|\mathcal{P}^{\rm cont}) = \frac{1}{2\pi\sqrt{(\sigma_\alpha^{\rm cont})^2 + \delta\mu_{\alpha,k}^2}\sqrt{(\sigma_\delta^{\rm cont})^2 + \delta\mu_{\delta,k}^2}} \times \exp\left[-\frac{1}{2}\left(\frac{(\mu_{\alpha,k} - \mu_\alpha^{\rm cont})^2}{[(\sigma_\alpha^{\rm cont})^2 + \delta\mu_{\alpha,k}^2]} + \frac{(\mu_{\delta,k} - \mu_\delta^{\rm cont})^2}{[(\sigma_\delta^{\rm cont})^2 + \delta\mu_{\delta,k}^2]}\right)\right]. \tag{4}$$

Following Bayes' theorem, the posterior probability distribution function (PDF), $P_{\rm tot}(\mathcal{P}|\mathcal{D})$, is related to the likelihood, $P_{\rm tot}(\mathcal{D}|\mathcal{P})$, through priors, $P(\mathcal{P})$, such that

$$P_{\rm tot}(\mathcal{P}|\mathcal{D}) \propto P_{\rm tot}(\mathcal{D}|\mathcal{P})P(\mathcal{P}). \tag{5}$$

We assume uniform priors for all free parameters except for $f_{\rm c}, \sigma_\alpha^{\rm cont}$, and $\sigma_\delta^{\rm cont}$, for which a flat truncated prior is preferred. We chose $0 \leq f_{\rm c} \leq 0.1$ as we estimated that the contamination for both samples of stars does not exceed 2%. We chose $0 \leq \sigma_\alpha^{\rm cont}, \sigma_\delta^{\rm cont} \leq 2.5\,{\rm mas\,yr^{-1}}$, as we estimated it to be $\sim$1.5 mas yr^{-1}. We sampled the posterior PDFs using a Markov chain Monte Carlo (MCMC) method with our own implementation of the Metropolis-Hastings algorithm (Metropolis et al. 1953; Hastings 1970).

[6] At the distance of M31, a dispersion of 100 km s^{-1}, would correspond to a contribution to the proper motion dispersion of only $\sim$25 μas yr^{-1}, which is much smaller than the typical proper motion uncertainties at the considered magnitudes.

3. Results

3.1. Correcting for the motion of the reference

The *Gaia* data is of exceptional accuracy; however, at the distance of M31, motions of the order of a few tenths of a μas yr^{-1} cannot be neglected. The motion of the *Gaia* reference quasars is expected to be consistent with 0 μas yr^{-1}. Nonetheless, there are reasons to expect that small systematics of the order of 10 μas yr^{-1} remain on small spatial scales (see Lindegren et al. 2021b, Table 7). Here, we aim to quantify the level of these systematics so that they can be folded into the analysis of the movement of M31 stars.

We first created a spatial grid with bins of size 72″×72″, entirely covering M31 and parts of its outskirts. Using the method introduced in Section 2.2, we inferred, for each bin, the mean motion of all quasars within a radius of 2.5° from the centre of the bin. This choice of radius serves two purposes: it makes sure that our approach remains as local as possible, while also ensuring that each sample contains a minimum of 600 quasars. This number yielded reasonable uncertainties of $\sim$10 μas yr^{-1} on each inference. The resulting correction maps were then smoothed with a Gaussian kernel of a standard deviation of $\sigma = 5$ pixels to reduce the noise caused by small-scale variations. We are aware that this strategy (using overlapping samples and then smoothing the maps) introduces correlations between the corrections in different bins, but, in turn, the resulting smoothed correction maps are simpler, more interpretable, and far less dominated by noise, while maintaining high spatial resolution. The resulting quasar mean proper-motion maps are shown in the left-hand panels of Figure 2 in the right ascension and the declination directions. While the maps show a mean motion close to 0 μas yr^{-1}, they reveal peak-to-peak differences of up to $\sim$20 μas yr^{-1}.

As was expected from Lindegren et al. (2021b), the analysis using the DR3 proper motions shows significant improvement over the analysis using the DR2 proper motions. As an example, the right-hand panels of Figure 2 show the corresponding analysis based on the DR2 data[7]. Not only is the range of corrections broader, reaching peak-to-peak differences of the order of 40 μas yr^{-1}, but these corrections are systematically and significantly offset from 0 μas yr^{-1}, with typical values of the order of the expected proper motion of M31 (see Table 1).

The DR3 correction maps can finally be applied as zero-point offsets to the *Gaia* proper motions of stars in the B_{pm} and R_{pm} samples. This approach is more local than the correction applied by S21, who determined a mean offset over a region of 4 to 20° around M31. Applying the method described in Section 2.2 yields $(\mu_\alpha^{\rm M31}, \mu_\delta^{\rm M31}) = (44.6 \pm 14.9, -29.8 \pm 13.1)$ μas yr^{-1} for B_{pm} and $(41.3 \pm 14.3, -74.8 \pm 13.2)$ μas yr^{-1} for R_{pm}. The uncertainties[8] were here computed as the quadratic sum of the uncertainty resulting from the inference itself and the

[7] We find only 23,139 DR2 counterparts to our DR3-based quasar sample, with the difference arising from the intrinsically more complete DR3 dataset (see Gaia Collaboration 2021).

[8] All statistical uncertainties throughout this paper are determined as the central 68% confidence interval.

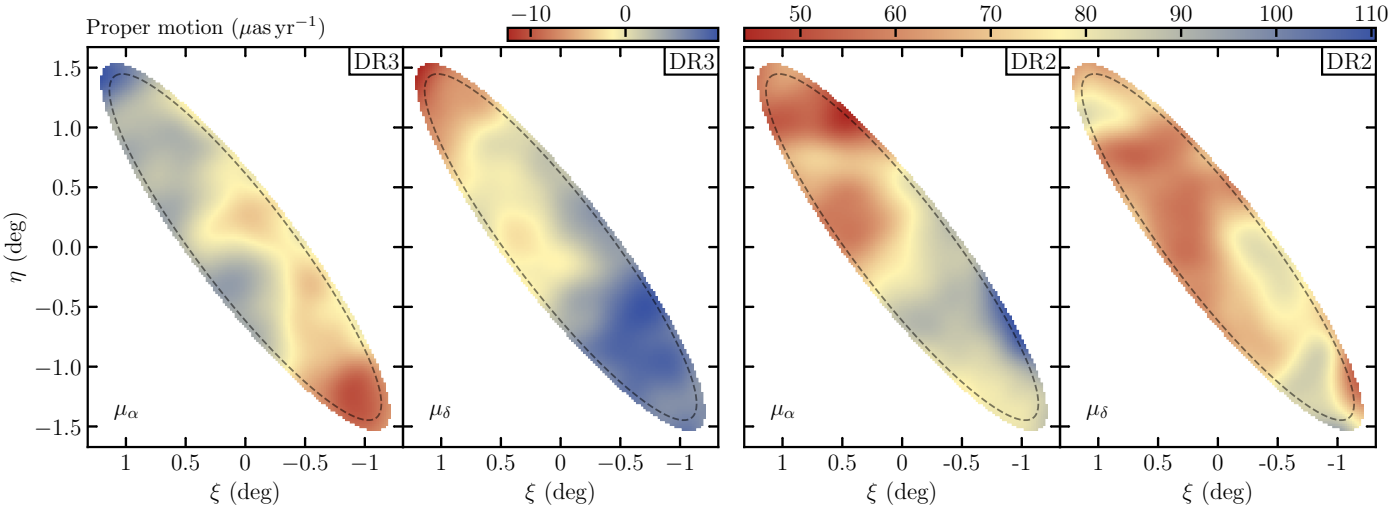

Fig. 2. Zero-point offsets for the *Gaia* stars' proper motions. *Left-hand panels*: smoothed maps of the quasar proper motions in the right ascension and the declination directions obtained using *Gaia* DR3. *Right-hand panels*: same but using *Gaia* DR2.

Table 1. Proper motion of M31 in the heliocentric reference frame.

Sample	Study	Dataset	$\mu_\alpha^{\rm M31}$(μas yr^{-1})	$\mu_\delta^{\rm M31}$(μas yr^{-1})
	vdM19	DR2	65.0 ± 24.0	−57.0 ± 21.9
	S21	EDR3	48.0 ± 10.2	−38.4 ± 7.8
B_{pm}	This work	DR2	−12.3 ± 23.1(stat) ± 98.2(sys)	−37.7 ± 27.1(stat) ± 122.3(sys)
	This work	DR3	**46.9** ± **11.7**(stat) ± **50.6**(sys)	**−29.1** ± **9.4**(stat) ± **35.6**(sys)
	S21	EDR3	41.3 ± 11.8	−87.4 ± 9.2
R_{pm}	This work	DR2	37.1 ± 23.1(stat) ± 93.6(sys)	−134.8 ± 20.7(stat) ± 79.1(sys)
	This work	DR3	**42.2** ± **11.3**(stat) ± **35.6**(sys)	**−75.0** ± **10.3**(stat) ± **32.8**(sys)

Notes. vdM19 corresponds to van der Marel et al. (2019). B_{pm} and R_{pm} are defined in S21 (Salomon et al. 2021) and in Section 2 of this work.

mean uncertainty from the inferences on the quasar corrections (9.8 μas yr^{-1} in the right ascension direction and 8.3 μas yr^{-1} in the declination direction). These results are similar to those obtained by S21 for the same samples, including the stark difference in the declination direction between the blue and red samples. Despite our more local correction of the mean motion of quasars, this measurement is still plagued by systematics that we now aim to constrain.

3.2. Quadrant analysis

To explore the reliability of the mean proper motion of M31 inferred above, we divided our blue and red samples into four quadrants centred on M31 (delimited by the dashed lines in the central panel of Figure 1), each containing the same number of stars, and then inferred the bulk motion of these smaller but independent samples. Consistent results in the different quadrants would mean that the systematics are folded into the statistical uncertainties, whereas incompatible results between quadrants would imply additional systematics that have not yet been taken into account.

The results for each quadrant are shown as the low-opacity symbols and associated error bars in Figure 3. For this particular quadrant configuration, the results for the blue stars are inconsistent, with one quadrant significantly offset from the three others. As we rotate the quadrants by small steps, we see rapid changes in the individual proper motions, leading to this inconsistency that is not specifically driven by our choice of quadrants. Potential contamination contributing to these inconsistencies is discussed in Section 4.3.

To assess the level of potential systematics in the proper motion inferences, we assume that the results from the four quadrants correspond to four independent measurements of the proper motion of M31[9], whose PDF on the mean ($\mu_{\alpha/\delta}$) and dispersion ($\sigma_{\mu_{\alpha/\delta}}$) informs us on the mean proper motion of the considered sample of stars and the unaccounted systematics. Examples of such PDFs are shown in the right-hand panel of Figure 3 for the right ascension proper motion, with the corresponding colours for the blue and red sample of stars.

An estimate of the mean and its statistical uncertainty were obtained from the marginalised PDFs of ($\mu_{\alpha/\delta}$), while the systematics were determined from the marginalised PDFs of ($\sigma_{\mu_{\alpha/\delta}}$). Conservatively, we chose to determine the systematics on the proper motion as the 90% confidence limit on this parameter. The corresponding values are shown as the dashed error bars in the left-hand panel of Figure 3. The final results, detailing the statistical and systematic uncertainties, are listed in Table 1 and are also displayed in the left-hand panel of Figure 4. In both cases, they are compared to literature values. The offset mentioned throughout this paper, between the blue and red stars, remains in our analysis. However, the systematics dominate the overall uncertainties, rendering the two star samples compatible

[9] As already mentioned above in Section 3.1, this is not entirely accurate due to the correlations introduced by our method of constructing the quasar correction maps and subsequently smoothed them.

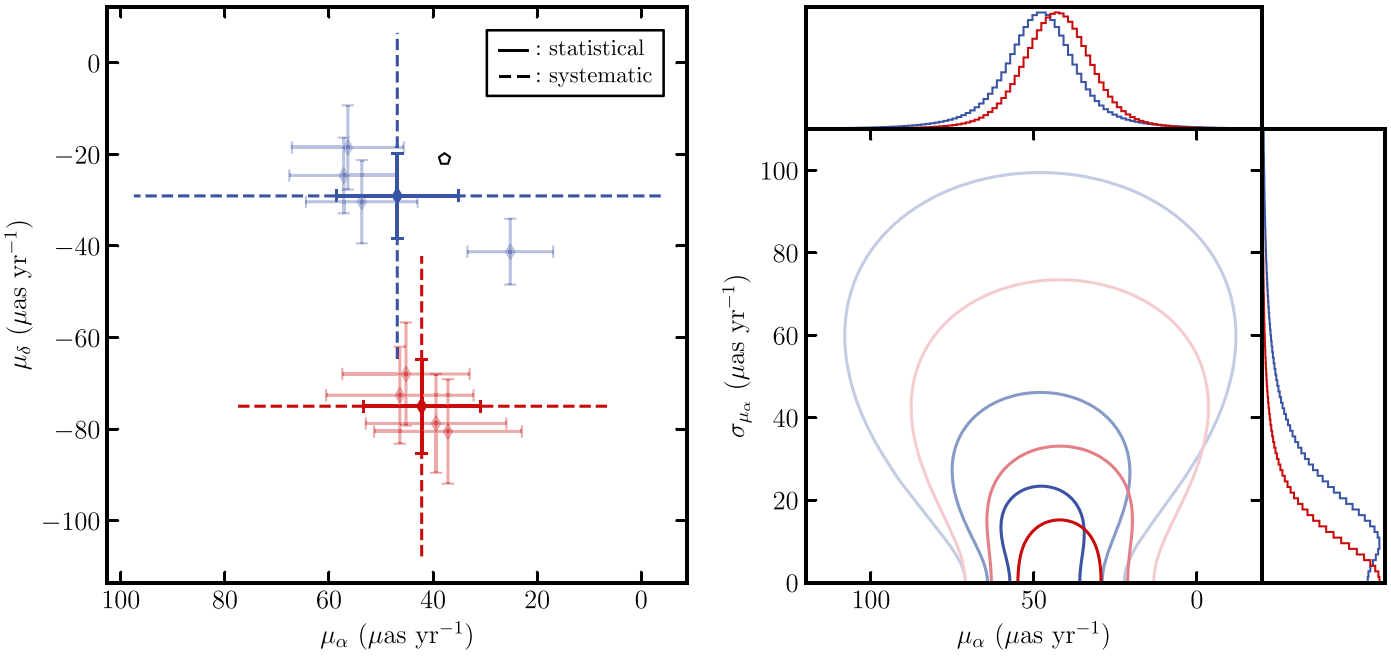

Fig. 3. Results of the quadrant analysis on M31. *Left-hand panel*: results from the quadrant analysis of M31 in the heliocentric frame. In low opacity are shown the inferences obtained for each individual quadrant, while shown clearly are the inferences resulting from the analysis of the 2D space from the right-hand panel. The solid and dashed error bars represent, respectively, the statistical and systematic uncertainties. The black pentagon represents a strictly radial motion in the galactic reference frame. *Right-hand panel*: example of the 2D PDFs (1, 2 and 3σ) obtained from the quadrant analysis in the direction of the right ascension and their marginalised 1D PDFs.

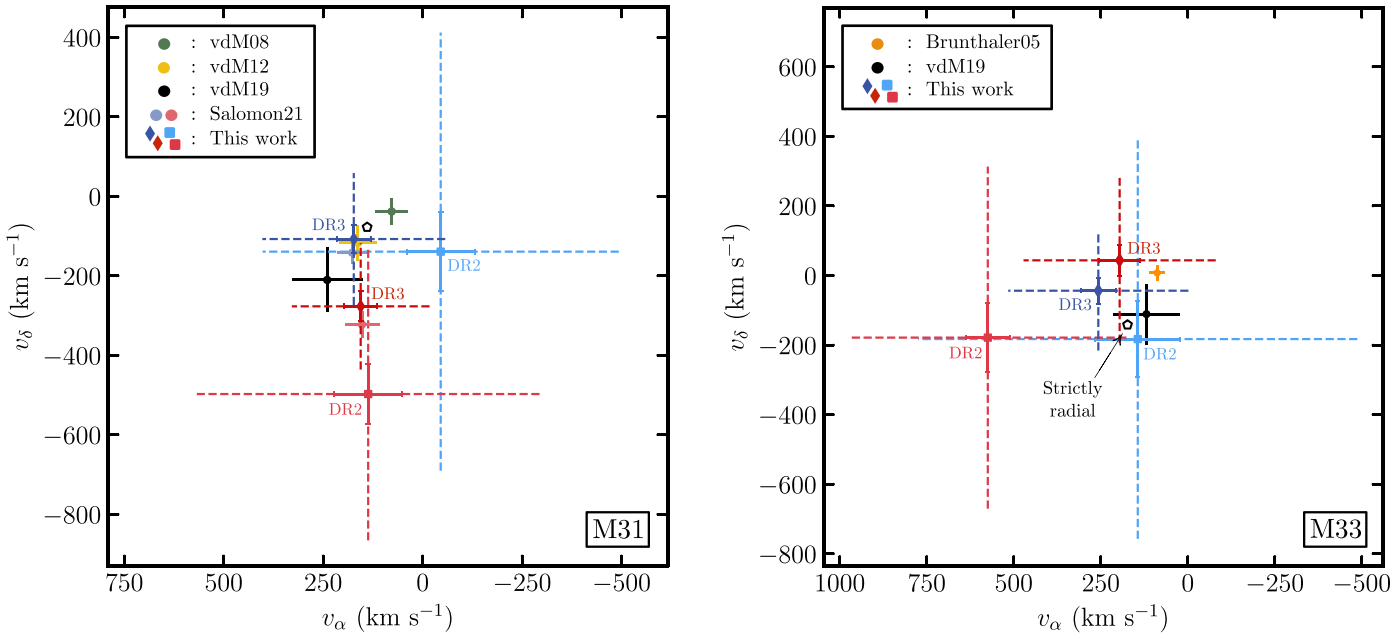

Fig. 4. Proper motions of M31 and M33 in the heliocentric reference frame. *Left-hand panel*: results for M31. The black pentagon represents a strictly radial motion in the galactic reference frame. The yellow dot is the weighted average of the three HST fields corrected for the internal kinematics of M31 (Sohn et al. 2012; van der Marel et al. 2012b). The black dot is the proper motion result based on the DR2 (van der Marel et al. 2019). The blue and red dots are the values derived by Salomon et al. (2021), while the diamonds and squares in different shades of blue and red are the results from this paper using DR3 and DR2, respectively. For our study, the solid error bars represent the statistical uncertainty and the dashed ones the systematic uncertainties. *Right-hand panel*: results for M33. The orange dot is the water maser weighted average of Brunthaler et al. (2005). The rest of the symbols are the same as in the left-hand panel.

Table 2. Proper motion of M33 in the heliocentric reference frame.

Sample	Study	Dataset	μ_α^{M33}(μas yr^{-1})	μ_δ^{M33}(μas yr^{-1})
	vdM19	DR2	31.0 ± 24.8	−29.0 ± 22.6
B_{pm}^{M33}	This work	DR2	37.5 ± 31.8(stat) ± 132.9(sys)	−47.7 ± 28.4(stat) ± 121.6(sys)
	This work	DR3	**67.1** ± **13.0**(stat) ± **54.6**(sys)	**−11.4** ± **9.7**(stat) ± **35.2**(sys)
R_{pm}^{M33}	This work	DR2	150.0 ± 16.5(stat) ± 87.5(sys)	−46.6 ± 26.1(stat) ± 102.3(sys)
	This work	DR3	**51.1** ± **15.3**(stat) ± **56.8**(sys)	**11.4** ± **11.9**(stat) ± **50.0**(sys)

Notes. vdM19 corresponds to van der Marel et al. (2019). B_{pm}^{M33} and R_{pm}^{M33} are defined in Section 3.3 of this work.

with one another, which was not the case before (the potential origins of those will be further discussed in Section 4). Moreover, these results are also compatible with the previous estimate of van der Marel et al. (2019), who used *Gaia* DR2.

3.3. The proper motion of M33

M33 is a companion of M31, located at roughly the same distance (d_{M33} = 794 ± 24 kpc, from McConnachie et al. 2004). It also shows signs of active star formation (Bastian et al. 2007; Peltonen et al. 2024), which, as in the case of M31, implies the presence of super-giants bright enough to be detected by *Gaia*. Before *Gaia*, estimates of the proper motion of M33 had already been made thanks to the presence of two water masers in its disc, IC 133 and M33/19, whose positions could be very accurately followed through interferometric VLBI measurements (Brunthaler et al. 2005). Although we can be highly confident in the precision of these measurements, much like the HST measurements of M31 (performed in small deep fields, Sohn et al. 2012), this estimate may be model-dependent and affected by local peculiar motions. As for the case of M31, an independent measurement, albeit a more uncertain one, was made possible through *Gaia* DR2 (van der Marel et al. 2019). The results from these previous works are shown in the right-hand panel of Figure 4 as orange and black dots, respectively.

We isolated M33 stars by applying the same cuts as were used for M31. The resulting blue (B_{pm}^{M33}) and red (R_{pm}^{M33}) samples are made of 1992 and 1093 stars, respectively. We corrected for the intrinsic motion of the stars with the flat disc model of Corbelli & Schneider (1997) and then followed the whole procedure described in the previous sections; that is, obtaining quasar correction maps and using them as zero-point offsets to the *Gaia* proper motions. These maps are very similar to the ones of M31, having mean motions compatible with 0 μas yr^{-1} (within the uncertainties of each bin inference) and peak-to-peak differences of ∼10 μas yr^{-1}.

The quadrant analysis reveals inconsistencies between the quadrant inferences, similar to those observed for M31, leading to similarly large systematic uncertainties. The final results are displayed in the right-hand panel of Figure 4 and are listed in Table 2. Interestingly, contrary to the case of M31, there is a better agreement between the B_{pm}^{M33} and R_{pm}^{M33} samples, which are nearly compatible within their statistical uncertainties. This might be the result of a less contaminated CRF around M33 than around M31, or the analysis being conducted on a smaller area (M33 being smaller). Our results also appear to be mostly consistent with the previous works, displaying an overall proper motion slightly less radial towards the MW than the result found by van der Marel et al. (2019).

4. Discussion

4.1. Summary

We reproduced the results from the study that was made by S21. In order to correct for the background motion of the quasars more locally, we created correction maps, which revealed local differences of ∼10 μas yr^{-1} (as expected from Lindegren et al. 2021b). Applying these correction maps as zero-point offsets for the stars yields results very similar to the ones previously found, and, despite the corrections, there remains an observed offset between the inferences of the blue and the red samples of stars. This offset is most likely due to systematics present in the *Gaia* DR3 data.

We show that we are able to assess the level of those systematics from the analysis of the overall inferred motion of different regions of the disc of M31. With a conservative 90% confidence level, these systematics are of the order of 30 μas yr^{-1} and commensurate with the proper motion of M31 itself. Such large systematics, explained by the spread of results from M31 star sub-samples, help to decipher the observed difference between the blue and red samples of stars, which was first pointed out by S21. These new results render our new estimates entirely compatible with the ones made by van der Marel et al. (2019) (using *Gaia* DR2) and Sohn et al. (2012) and van der Marel et al. (2012b) (using HST), all of which are displayed in the left-hand panel of Figure 4. No proper motion estimate of M33 had been made using *Gaia* DR3, but our updated results are consistent with previous works on the topic (see right-hand panel of Figure 4).

Overall, it is of course unsatisfactory to be unable to tightly constrain these systematics, which are large compared to the proper motions of either M31 or M33, but nevertheless very small and close to the accuracy expectations of *Gaia*. Below, we discuss potential sources that could explain these unaccounted systematics.

4.2. Impact of colour

The strong difference in the inferred proper motion of M31 from the blue and red samples of stars appear to imply that colour could (in part) play a role in the presence of systematics. Lindegren et al. (2021a), in their study of the dependence on colour and magnitude of the parallax systematics in *Gaia* DR3, point out that, since parallaxes and proper motions are jointly determined, such (small) systematics are likely also present for the *Gaia* proper motion values. Cantat-Gaudin & Brandt (2021) show that, at the bright end, these are indeed biased, with the biases being a function of the magnitude, the position, and the colour of an object. They estimate that the systematics could be

of the order of 10 μas yr^{-1} for objects that are bluer or redder than the mean colour of *Gaia* objects. The CMD in the right-hand panel of Figure 1 shows the colour distribution of stars from both samples along with that of the background quasars. It is clear that there is no strong overlap between the different populations, which may explain why, if there is indeed a colour-dependent systematic to the *Gaia* proper motions over the full magnitude range, the quasar correction did not remove this effect.

4.3. Potential contamination

With the accuracy that we are aiming for, biases may arise from differences between the assumed model and the true distribution of the data. In particular, it is easy to imagine that the presence of a single outlier with small proper motion uncertainties (e.g. a bright contaminating star) may shift the mean inferred proper motion. This shift may be small overall, but might still be comparable to the tiny mean proper motion of M31.

As has already been pointed out by S21, the PAndAS MW stream, located at a distance of ∼17 kpc (Martin et al. 2014) and crossing M31 in its most northern part, overlaps with the blue sample in the CMD for $G < 18$. However, re-doing the analysis only for stars with $G > 18$ does not change our conclusions.

We also investigated the intrinsically higher contamination of the red sample that was already mentioned in S21. Using the numerous photometric bands of the PHAT survey (Williams et al. 2014) and, in particular, the (F336W–F475W) versus (F110W–F160W) colour–colour space[10], we see a clear separation between stars removed by the proper motion cut and the R_{pm} sample. We interpret this as the separation between contaminating dwarfs and relevant giants. This separation, if truly discriminating the contaminants from true members, would imply that the 1.8% contamination (from Section 2.1) is underestimated at least by a factor of two, the newly found contamination fraction being of the order of 4% for the R_{pm} sample. Re-doing the full analysis without these possible contaminants highlighted in the PHAT data barely changes our previous results, the new inference remaining far from the blue inferences.

4.4. Improvements since Gaia DR2 and future prospects

Because we aim for proper motion precision that challenges *Gaia*'s current capabilities, it is worth checking whether the situation has improved between the DR2 and the DR3. To this end, when available, we extracted the DR2 proper motions for the quasar, the B_{pm} and R_{pm} samples built in Section 2.1. We then performed the full analysis on this new DR2-based dataset. The right-hand panels of Figure 2 show the corresponding quasar correction maps, which, as was already mentioned in Section 3.1, are significantly worse than for DR3: not only are the corrections no longer centred on ∼0 μas yr^{-1}, but the full amplitude of corrections spans ∼40 μas yr^{-1} (vs. ∼0 μas yr^{-1} and ∼20 μas yr^{-1}, respectively, for DR3). The quadrant inconsistencies are larger than for the DR3 analysis, implying larger final systematic uncertainties reaching ∼90 μas yr^{-1}, two to three times larger than for the DR3 analysis (see Table 1). In Figure 4, we see that these results have statistical uncertainties comparable to the ones from van der Marel et al. (2019). However, we see that the model's favoured values are very different and are only compatible because of the large systematic uncertainties. We observe the same trend for M33, where the results using DR2 are way more uncertain than the ones using DR3.

As is noted by Lindegren et al. (2021b), the proper motion uncertainties should decrease over time as $T^{-3/2}$; that is, by a factor of 0.51 between DR2 (22 months of data) and DR3 (34 months of data). It is reassuring to observe this refinement in our results, both for the statistical and systematic uncertainties. If a similar improvement is achieved between DR3 and DR4, the proper motion uncertainties should be reduced by nearly a factor of 3. We hope this will help resolve the discrepancies between the blue and red star samples. We also anticipate the possibility of reaching systematic uncertainties comparable to the ∼12 μas yr^{-1} reached by HST (in Sohn et al. 2012; van der Marel et al. 2012b), but over a sample of stars covering the entire extent of M31. Re-evaluating the proper motions of M31 and M33 with DR4 could further constrain their dynamics, allowing for studies of their orbital history to be compared with thorough past works (e.g. van der Marel et al. 2012a; Patel et al. 2017).

[10] The first combination is gravity-sensitive while the latter explores the infrared colours of the stars.

Acknowledgements. 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.

References

Bastian, N., Ercolano, B., Gieles, M., et al. 2007, MNRAS, 379, 1302
Brunthaler, A., Reid, M. J., Falcke, H., Greenhill, L. J., & Henkel, C. 2005, Science, 307, 1440
Cantat-Gaudin, T., & Brandt, T. D. 2021, A&A, 649, A124
Chemin, L., Carignan, C., & Foster, T. 2009, ApJ, 705, 1395
Corbelli, E., & Schneider, S. E. 1997, ApJ, 479, 244
Gaia Collaboration (Brown, A. G. A., et al.) 2021, A&A, 649, A1
Gaia Collaboration (Klioner, S. A., et al.) 2022, A&A, 667, A148
Gaia Collaboration (Bailer-Jones, C. A. L., et al.) 2023a, A&A, 674, A41
Gaia Collaboration (Vallenari, A., et al.) 2023b, A&A, 674, A1
Hastings, W. K. 1970, Biometrika, 57, 97
Hernitschek, N., Schlafly, E. F., Sesar, B., et al. 2016, ApJ, 817, 73
Ibata, R. A., Lewis, G. F., Conn, A. R., et al. 2013, Nature, 493, 62
Kahn, F. D., & Woltjer, L. 1959, ApJ, 130, 705
Lewis, A. R., Dolphin, A. E., Dalcanton, J. J., et al. 2015, ApJ, 805, 183
Li, Y.-S., & White, S. D. M. 2008, MNRAS, 384, 1459
Lindegren, L., Bastian, U., Biermann, M., et al. 2021a, A&A, 649, A4
Lindegren, L., Klioner, S. A., Hernández, J., et al. 2021b, A&A, 649, A2
Martin, N. F., Ibata, R. A., Rich, R. M., et al. 2014, ApJ, 787, 19
McConnachie, A. W., Irwin, M. J., Ferguson, A. M. N., et al. 2004, MNRAS, 350, 243
McConnachie, A. W., Irwin, M. J., Ferguson, A. M. N., et al. 2005, MNRAS, 356, 979
Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. 1953, J. Chem. Phys., 21, 1087
Patel, E., Besla, G., & Sohn, S. T. 2017, MNRAS, 464, 3825
Peltonen, J., Rosolowsky, E., Williams, T. G., et al. 2024, MNRAS, 527, 10668
Peñarrubia, J., Gómez, F. A., Besla, G., Erkal, D., & Ma, Y.-Z. 2016, MNRAS, 456, L54
Salomon, J. B., Ibata, R. A., Famaey, B., Martin, N. F., & Lewis, G. F. 2016, MNRAS, 456, 4432
Salomon, J. B., Ibata, R., Reylé, C., et al. 2021, MNRAS, 507, 2592
Sohn, S. T., Anderson, J., & van der Marel, R. P. 2012, ApJ, 753, 7
van der Marel, R. P., & Guhathakurta, P. 2008, ApJ, 678, 187
van der Marel, R. P., Besla, G., Cox, T. J., Sohn, S. T., & Anderson, J. 2012a, ApJ, 753, 9
van der Marel, R. P., Fardal, M., Besla, G., et al. 2012b, ApJ, 753, 8
van der Marel, R. P., Fardal, M. A., Sohn, S. T., et al. 2019, ApJ, 872, 24
Williams, B. F., Lang, D., Dalcanton, J. J., et al. 2014, ApJS, 215, 9