Gaia Data Release 2. Catalogue validation

F. Arenou, X. Luri, C. Babusiaux, C. Fabricius, A. Helmi, T. Muraveva, A. C. Robin, F. Spoto, A. Vallenari, T. Antoja, et al.

A&A 616, A17 (2018)
https://doi.org/10.1051/0004-6361/201833234

Astronomy & Astrophysics

Gaia Data Release 2

Special issue

Gaia Data Release 2

Catalogue validation

F. Arenou[1], X. Luri[2], C. Babusiaux[18,1], C. Fabricius[2], A. Helmi[3], T. Muraveva[16], A. C. Robin[4], F. Spoto[11,19], A. Vallenari[5], T. Antoja[2], T. Cantat-Gaudin[5,2], C. Jordi[2], N. Leclerc[1], C. Reylé[4], M. Romero-Gómez[2], I.-C. Shih[1], S. Soria[2], C. Barache[8], D. Bossini[5], A. Bragaglia[16], M. A. Breddels[3], M. Fabrizio[12,13], S. Lambert[8], P. M. Marrese[12,13], D. Massari[3], A. Moitinho[9], N. Robichon[1], L. Ruiz-Dern[1], R. Sordo[5], J. Veljanoski[3], L. Eyer[6], G. Jasniewicz[20], E. Pancino[15], C. Soubiran[14], A. Spagna[17], P. Tanga[11], C. Turon[1], and C. Zurbach[20]

[1] GEPI, Observatoire de Paris, Université PSL, CNRS, 5 Place Jules Janssen, 92190 Meudon, France
e-mail: Frederic.arenou@obspm.fr
[2] Dept. FQA, Institut de Ciències del Cosmos, Universitat de Barcelona (IEEC-UB), Martí Franquès 1, 08028 Barcelona, Spain
[3] Kapteyn Astronomical Institute, University of Groningen, Landleven 12, 9747 AD Groningen, The Netherlands
[4] Institut UTINAM, CNRS, OSU THETA Franche-Comté Bourgogne, University Bourgogne Franche-Comté, 25000 Besançon, France
[5] INAF, Osservatorio Astronomico di Padova, Vicolo Osservatorio, Padova 35131, Italy
[6] Observatoire de Genève, Université de Genève, 1290 Versoix, Switzerland
[7] Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB30HA, UK
[8] SYRTE, Observatoire de Paris, Université PSL, CNRS, Sorbonne Université, LNE, 61 avenue de l'Observatoire, 75014 Paris, France
[9] CENTRA, Universidade de Lisboa, FCUL, Campo Grande, Edif. C8, 1749-016 Lisboa, Portugal
[10] Leiden Observatory, Leiden University, Niels Bohrweg 2, 2333 CA Leiden, The Netherlands
[11] Université Côte d'Azur, Observatoire de la Côte d'Azur, CNRS, Laboratoire Lagrange, Bd de l'Observatoire, CS 34229, 06304 Nice Cedex 4, France
[12] INAF – Osservatorio Astronomico di Roma, Via di Frascati 33, 00078 Monte Porzio Catone (Roma), Italy
[13] ASI Science Data Center, Via del Politecnico, Roma, Italy
[14] Laboratoire d'astrophysique de Bordeaux, Université de Bordeaux, CNRS, B18N, allée Geoffroy Saint-Hilaire, 33615 Pessac, France
[15] INAF – Osservatorio Astrofisico di Arcetri, Largo Enrico Fermi 5, 50125 Firenze, Italy
[16] INAF – Osservatorio di Astrofisica e Scienza dello Spazio di Bologna, via Piero Gobetti 93/3, 40129 Bologna, Italy
[17] INAF – Osservatorio Astronomico di Torino, via osservatorio 20, Pino Torinese, Torino, Italy
[18] Université Grenoble Alpes, CNRS, IPAG, 38000 Grenoble, France
[19] IMCCE, Observatoire de Paris, PSL Research University, CNRS, Sorbonne Université, UPMC Univ. Paris 06, Univ. Lille, 77 av. Denfert-Rochereau, 75014 Paris, France
[20] Laboratoire Univers et Particules de Montpellier, Université Montpellier, CNRS, Place Eugène Bataillon, CC72, 34095 Montpellier Cedex 05, France

Received 13 April 2018 / Accepted 4 June 2018

ABSTRACT

Context. The second *Gaia* data release (DR2) contains very precise astrometric and photometric properties for more than one billion sources, astrophysical parameters for dozens of millions, radial velocities for millions, variability information for half a million stars from selected variability classes, and orbits for thousands of solar system objects.
Aims. Before the catalogue was published, these data have undergone dedicated validation processes. The goal of this paper is to describe the validation results in terms of completeness, accuracy, and precision of the various *Gaia* DR2 data.
Methods. The validation processes include a systematic analysis of the catalogue content to detect anomalies, either individual errors or statistical properties, using statistical analysis and comparisons to external data or to models.
Results. Although the astrometric, photometric, and spectroscopic data are of unprecedented quality and quantity, it is shown that the data cannot be used without dedicated attention to the limitations described here, in the catalogue documentation and in accompanying papers. We place special emphasis on the caveats for the statistical use of the data in scientific exploitation. In particular, we discuss the quality filters and the consideration of the properties, systematics, and uncertainties from astrometry to astrophysical parameters, together with the various selection functions.

Key words. catalogs – stars: fundamental parameters – astrometry – techniques: radial velocities – stars: variables: general – minor planets, asteroids: general

1. Introduction

This paper describes the validation of the second data release, *Gaia* DR2, from the European Space Agency mission *Gaia* (Gaia Collaboration 2016, 2018b). The approach followed by this catalogue validation is an external, transverse analysis of the properties of the various contents.

A great variety of catalogue properties are described together with their dedicated validation by Lindegren et al. (2018) for astrometry, Evans et al. (2018) for photometry, Sartoretti et al. (2018) and Katz et al. (2018) for spectroscopic data, Andrae et al. (2018) for astrophysical parameters (AP), Holl et al. (2018) for variable stars, Gaia Collaboration (2018e) for solar system objects, and Mignard et al. (2018) for the reference frame. In addition, science demonstration papers such as Gaia Collaboration (2018a) for the Hertzsprung–Russel (HR) diagram, Gaia Collaboration (2018d) for the Milky Way disc kinematics or Gaia Collaboration (2018c) for the Milky Way satellites have also indirectly contributed much to demonstrate the overall quality of the catalogue, and Gaia Collaboration (2018b) summarises its impressive characteristics. For this reason, special emphasis is placed here on the caveats attached to the data, in order to allow a better exploitation of the catalogue.

We mention here only what is strictly necessary and invite the reader to refer to the above papers or to the *Gaia* online documentation[1] for details. As we describe below, understanding the properties of the catalogue is mandatory for a proper scientific use of the data; the papers mentioned above are important – and inspiring.

This paper is organised as follows. We first describe the general consistency of the data (Sect. 2) and then the completeness of the catalogue from small to large scale (Sect. 3). We discuss in turn astrometric properties, systematics and random (Sect. 4), photometric quality (Sect. 5), spectroscopic results (Sect. 6), AP (Sect. 7), and validation of solar system objects (Sect. 8).

2. Data and general validation tests

2.1. Data integrity and consistency

The data release consists of several data tables. In most of this paper, we focus on the *gaia_source* table with the mean parameters for about 1.7 billion point-like sources. In addition, smaller tables contain the analysis results of light curves for variable sources, Sect. 5.4, and the results for solar system objects, discussed in Sect. 8.

The *gaia_source* catalogue contains positions and G-band mean photometry, both with several auxiliary parameters, for all sources. For a large subset, 1.3 billion sources, it also gives proper motions and parallaxes, again with many auxiliary parameters; for another large subset, 1.4 billion sources, photometry in the $G_{\rm BP}$ or $G_{\rm RP}$-bands; for smaller subsets of between 77 and 161 million sources, various AP; and for a more modest 7 million sources, the radial velocity. Light curves are given for half a million variables (from selected variability classes only, such as RR Lyrae stars, Cepheids, long-period variables, BY Dra-type stars, δ Scuti and SX Phoenicis stars, and short-timescale variables) and 2 million individual CCD observations in 3 30 000 transits of fourteen thousand asteroids.

[1] van Leeuwen et al. (2018), `http://gea.esac.esa.int/archive/documentation/GDR2/index.html`. In this paper, we note the catalogue fields with a special font, e.g. `astrometric_chi2_al`. The description of these fields can be found in Sect. 14 of the catalogue documentation, Hambly et al. (2018).

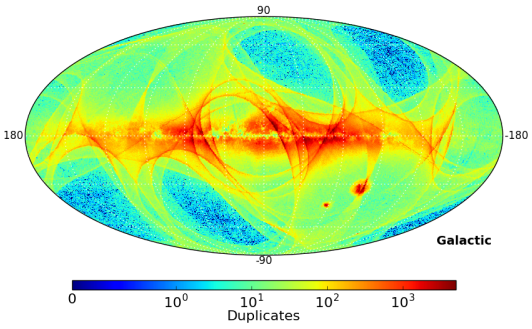

Fig. 1. Map of duplicated sources in Galactic coordinates.

For all preliminary versions of the *Gaia* DR2 catalogue, one of the validation tasks consisted of several basic verification tests in order to check the internal consistency of the data records, for instance, that data fields were present when and only when expected, that fluxes were converted consistently into magnitudes, or that positions were expressed equally well in equatorial, ecliptic, and Galactic coordinates. The fields were corrected when needed for the final catalogue, and the validation results are not reported here.

In addition, the data in the *Gaia* DR2 as a whole generally behave following expectations. This has been established for example by comparing the (clustering) behaviour of multi-dimensional distributions of the observables and their uncertainties for different regions on the sky (symmetric with respect to the disc, and with a similar number of transits/observations) using the Kullback–Leibler divergence statistic (KLD; Kullback & Leibler 1951). Furthermore, comparisons to Galactic models confirm that the global behaviour of most of the data, at a surface level, is as expected.

2.2. Duplicate entries

The *Gaia* data processing is complex, see for example Lindegren et al. (2018, Sect. 2), and has still not reached full maturity. It may therefore happen that the same source is processed twice, but based on disjoint sets of observations. In the published catalogue, only one of the solutions has been kept, and the flag `duplicated_source` has been set, but the removed duplicated solution was made available for validation. Although these duplicates have relatively little effect, decreasing their number for the next data release would nevertheless allow to increase the number of observations per star. Turning weakness into strength, such duplicated sources offered an interesting opportunity for the validation, as we discuss in the various sections below.

The duplicated sources, with two independent solutions in the initial versions of *Gaia* DR2, are found throughout the entire sky (Fig. 1), but because of various details related to on-board as well as on-ground processing, they are not a random subset of the catalogue and more often lie on the bright side, reaching 39% at $G = 10.3$ as shown in Fig. 2. Conclusions based on this subset are therefore not necessarily representative for the full catalogue. This is especially the case for sources without full astrometric solutions, where the quality indicators show poorer results for the duplicate solutions. For sources with full astrometry, on the other hand, the quality is only marginally affected.

Image sizes in *Gaia* in the along-scan direction (AL) are around 0.″2. It follows from the way observations are acquired, however, that sources separated by less than 0.″4–0.″5 cannot be resolved without a dedicated process. Such a process is still not in place, and for *Gaia* DR2, duplicated sources were therefore

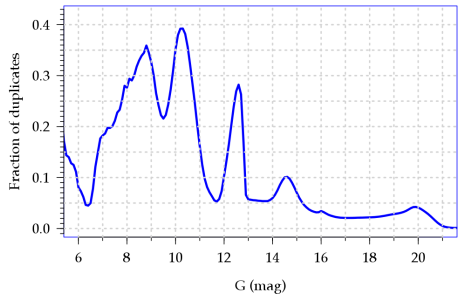

Fig. 2. Fraction of duplicated sources vs. G magnitude. The peaks must be due to a combination of problems in the on-board detections and the cross-match process.

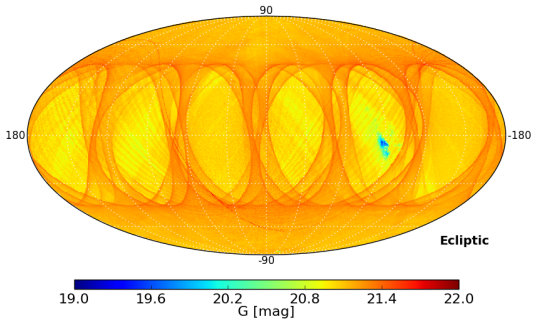

Fig. 3. Sky map in ecliptic coordinates of limiting magnitude: 99th percentile in G.

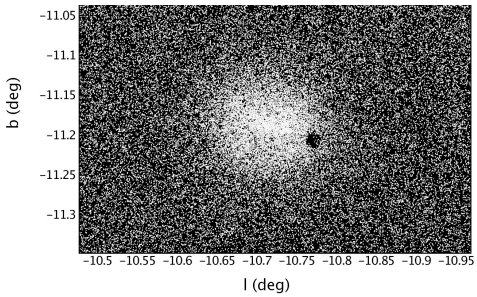

Fig. 4. Only a few regions are underscanned due to the lack of on-board resources, such as here, on the edge of NGC 6541.

defined as solutions separated by less than 0.″4. The average separation within duplicate pairs is 0.″019, which is so small that it shows that the pairs represent basically the same sources, and that resolved double stars can only represent a very small fraction of them.

Contamination by close-by sources may indeed give erroneous solutions, as we discuss in Sect. 4.1. The processing for *Gaia* DR2 rests on the assumption that all sources are isolated point sources. When this condition is not fulfilled, the resulting photometry and astrometry may suffer distortions. The G_{BP} and G_{RP} photometry is especially vulnerable because it is based on aperture photometry of dispersed spectra.

3. Sky coverage and completeness

In this section, the completeness of the *Gaia* DR2 catalogue is described with respect to the actual sky content. The situation is obviously more complicated for the various data that may or may not be available for each source. In this respect, Appendix A details that first the satellite observation, then the various processing steps have built the catalogue content, that is, the fraction for each data category of the total number of sources, and we refer to Gaia Collaboration (2018b) for characteristic figures of the catalogue.

3.1. Limiting magnitude

Figure 3 illustrates the variation in limiting magnitude (99th percentile) across the sky. The map is in ecliptic coordinates in order to emphasize the importance of the scanning law. The brightest limit is found near the Galactic centre, where the star density is very high, and where we have relatively few scans. On the other hand, the faintest limit is achieved near the caustics of the scanning law at ecliptic latitude ±45°, where more observations were made.

3.2. Overall large-scale coverage and completeness

For *Gaia* DR1, several regions suffered from limited on-board resources, which created holes in the sky coverage; these regions are now covered, and only a few remain, such as near the NGC 6541 globular cluster, shown in Fig. 4.

Figure 5 shows the completeness versus OGLE data (Udalski et al. 2008) in some selected fields with different sky density. Because the OGLE spatial resolution is poorer than that of *Gaia*, a comparison with OGLE provides upper limits to the *Gaia* completeness. Compared to *Gaia* DR1 (Fig. 15 of Arenou et al. 2017), the coverage is now much better because the *Gaia* DR2 catalogue is almost complete at $G = 18$, whereas the coverage was lower than 80% for *Gaia* DR1 as soon as the density was above one hundred thousand stars per square degree.

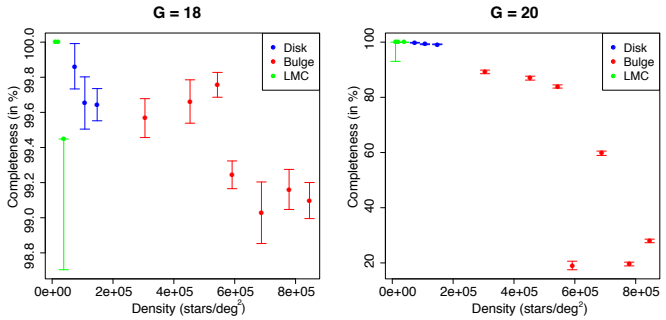

Fig. 5. *Gaia* DR2 completeness vs. some OGLE fields at $G = 18$ and $G = 20$ as a function of the measured density at $G = 20$. The scale of the two plots is very different.

For very crowded regions, we used *Hubble* Space Telescope (HST) observations of 26 globular clusters, which are expected to be complete down to at least $G \sim 24$ and have a spatial resolution comparable to that of *Gaia*. The HST data we employed are the same as were used in Arenou et al. (2017). They were acquired by Sarajedini et al. (2007) with the Advanced Camera for Surveys (ACS) and contain photometry in F606W and F814W filters, which we transformed into *Gaia* G magnitudes through a direct comparison of the magnitudes of the stars in the G, F606W, and F814W passbands. This avoids issues due to variations of metallicity and interstellar extinction. For each cluster, we compared the number of sources in various magnitude slices in the inner (inside 0.5 arcmin) and outer (0.5–2.2 arcmin) regions. The result of all clusters is given in Table B.1, and an example for NGC 6121 (M4) is shown in Fig. 6. The information contained in Table B.1 is also visually represented in Fig. 7, where it is clearly visible that the completeness level depends on both magnitude and local density (for $G < 20$). Overall, the completeness level of *Gaia* DR2 is much higher than in *Gaia* DR1.

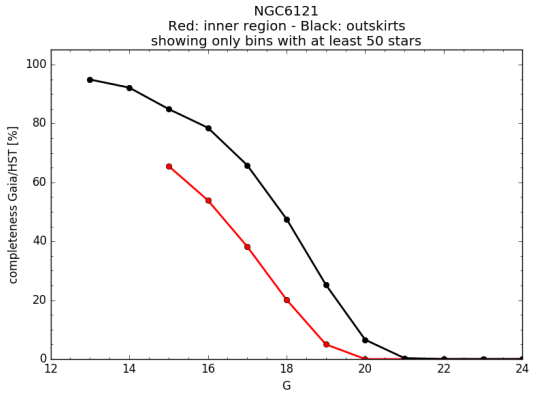

Fig. 6. Completeness level with respect to HST data in the inner (within 0.5 arcmin, red) and outer region (black) of the cluster NGC 6121.

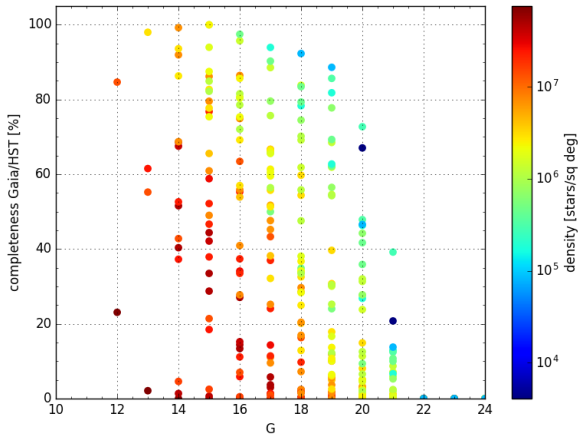

Fig. 7. Completeness levels with respect to HST data in different regions of 26 globular clusters with various local densities, showing the influence of crowding on the completeness.

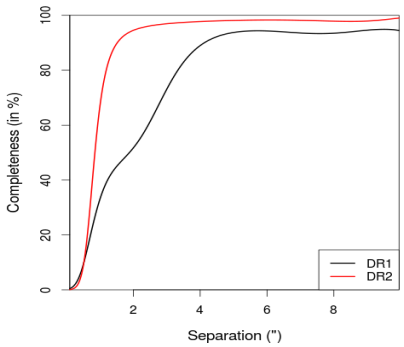

Fig. 8. Improvement of the completeness (in percent) of visual double stars from the WDS catalogue as a function of the WDS separation between components, from *Gaia* DR1 (black) to *Gaia* DR2 (red).

3.3. Small-scale completeness of Gaia DR2

One first way to check the spatial resolution of the *Gaia* catalogue is to use known double stars, using the Washington Visual Double Star Catalogue (WDS; Mason et al. 2001). Figure 8 shows the completeness as a function of the separation between visual double stars as provided by the WDS. This shows that the completeness starts to decrease at around 2″, while it was around 4″ in *Gaia* DR1.

The small-scale completeness can also be evaluated from the distance distribution between source pairs in the whole catalogue. Figure 9 shows distributions in two test fields: a dense field near the Galactic plane and a sparse field at −60° Galactic latitude. The dense field contains 456 142 sources in a circle of radius 0.5°, while the sparse field has 250 092 sources within a radius of 5°. The sparse field has therefore a surface density 200 times lower than the dense field. From *Gaia* DR1 to *Gaia* DR2, the dense field has obtained 56% more sources, whereas the sparse field has only gained 12%. The top panel shows the distributions for the dense field. *Gaia* DR1 (black lower curve) shows a deficit of pair distances smaller than 3.″7 and extremely few below 2″. For *Gaia* DR2 (red upper curve), the deficit does not set in until 2.″2 and gently decreases to zero around 0.″5. For *Gaia* DR1, it was required that all sources had a known colour, but this requirement has been waived for *Gaia* DR2. This difference explains the gain in angular resolution as illustrated by the blue dashed curve in the middle, showing the distribution for *Gaia* DR2 sources with known colour. This curve shows the same features as the curve for *Gaia* DR1, but lies slightly higher as the result of a gain of 10% in the number of sources. The specific distances at which the resolution changes are related to the size of the data acquisition windows, which is 0.″7 × 2.″1 for the point images and 3.″5 × 2.″1 for the BP, RP low-dispersion spectra. The situation for the sparse field is dramatically different for separations below 2″, where we now see a strong peak of binaries. A population of binaries must also be present in the dense field, which is dominated by a much larger number of remote sources, however.

Based on Fig. 9b, the binaries might have been expected to grow towards small separations, down to the angular resolution, ∼0.″12. The question therefore arises where the missing binaries with a 0.″12 < ρ < 0.″5 separation lie. Because there was

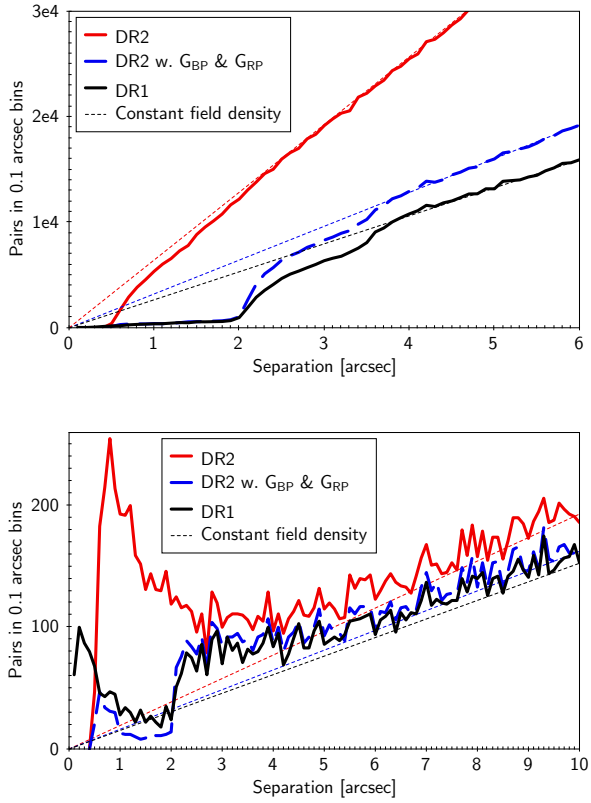

Fig. 9. Histogram of source pair separations in two test fields for *Gaia* DR2 sources, for *Gaia* DR2 sources with G_{BP} and G_{RP} photometry, and for *Gaia* DR1 sources. *Top panel*: dense field at $(l, b) = (-30°, -4°)$. *Bottom panel*: sparse field at $(l, b) = (-100°, -60°)$. The thin dotted lines show the relation for a random distribution.

no special treatment for non-single objects, the missing binaries were handled as single objects, which may sometimes have corrupted astrometry or colour photometry and produced either sources with only two astrometric parameters, or the spurious solutions we discuss in Sect. 4.1.

3.4. High proper motion stars

When we searched for known high proper motion (HPM) stars, we found that 17% of the SIMBAD HPM stars with a proper motion higher than 0.6 arcsec yr^{-1} are missing in *Gaia* DR2. These are preferentially bright stars.

Gaia DR1 missed many more HPM stars because the cross-match of the observations to the sources relied on ground-based proper motions. The cross-match of *Gaia* DR2 is much more independent of a star catalogue, providing a significant improvement. Further improvements are already in place for the future.

4. Astrometric quality of *Gaia* DR2

We have mainly checked the astrometric quality of *Gaia* DR2 for sources with the full five-parameter astrometric solution and focused on the parallaxes and proper motions. The remaining 360 million sources, with only two published parameters, are either fainter than 21 mag, have only a few transits, or gave very bad fits to the five-parameter model (binaries, diffuse objects, etc.). The quality of this group is much lower than for the rest of the sources, and the group therefore is of limited interest. We also checked the reference frame, which was aligned to the ICRF3-prototype, but we do not mention our results here as they are in full agreement with Mignard et al. (2018) and Lindegren et al. (2018), to which we refer.

4.1. Spurious astrometric solutions

Good astrometric solutions can only be obtained if there are many scans that are well spread in scanning angle and in a sufficiently long period of time[2] (Lindegren et al. 2018). In some parts of the sky, this fundamental requirement was not met during the 21 months of observations used in *Gaia* DR2 astrometry. Solutions in these areas will be more susceptible to disturbances introduced by a close-by source, for instance. Especially difficult are areas where one or two scan directions dominate, and even more so if one of these directions is perpendicular to the direction to the Sun and therefore insensitive to parallax. In future data releases, which will be based on longer time series, this problem will diminish.

An obvious way to check for problematic astrometric solutions is to search for significantly negative parallaxes. The top panel of Fig. 10 shows the sky distribution of the 113 393 sources with parallaxes below -10 mas and $\varpi/\sigma_\varpi < -10$. They are clearly concentrated in the dense areas of the Galactic plane and the Magellanic Clouds, especially in some areas delineated by scan patterns. In the centre panel, showing the same, but for 439 020 positive parallaxes ($\varpi > 10$ mas and $\varpi/\sigma_\varpi > 10$), the same patterns are visible, but with a uniform background of supposedly well-behaved astrometric solutions. Finally, the bottom panel shows the 254 007 positive parallaxes after the quality filters defined in Eqs. (1) and (2) were applied. The same filters reduce the number of negative parallaxes to just 298. We conclude that some sky areas contain sources with spurious astrometry, and that these poor solutions may equally well contain a negative or a positive parallax.

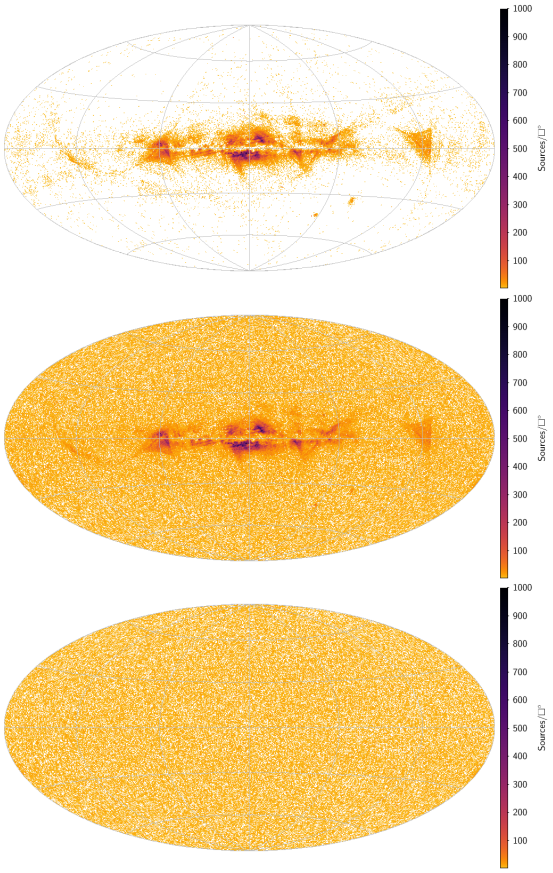

Fig. 10. Sky maps in Galactic coordinates showing sources with parallaxes with uncertainties smaller than 10% and significantly different from zero. *Top panel*: parallaxes below -10 mas. *Centre panel*: parallaxes higher than $+10$ mas. *Bottom panel*: parallaxes higher than $+10$ mas after the quality filters in Eqs. (1) and (2) are applied.

Proper motions are also affected by poor solutions. For example, the proper motions of 6189 stars are higher than 500 mas yr^{-1} in *Gaia* DR2, of which only 70% are known in SIMBAD. Selecting only the stars with `visibility_periods_used` > 8, that is, with a better astrometric quality, this number increases to 93%, which shows that the non-selected objects are probably spurious.

The detrimental effect of the spurious solutions clearly appears in HR diagrams (Fig. 11a) or proper motion diagrams (Fig. 11b). These figures also show that quality filters can fortunately be devised to clean the samples.

The filter used in these plots is the same as defined by Gaia Collaboration (2018a) for their study of the HR diagram and in Appendix C of Lindegren et al. (2018), their Eqs. (C-1) and (C-2). We define

- χ^2 = `astrometric_chi2_al`
- ν = `astrometric_n_good_obs_al -5`
- $u = \sqrt{\chi^2/\nu}$
- E = `phot_bp_rp_excess_factor`[3]

we accept solutions fulfilling the conditions

$$u < 1.2 \times \max(1, \exp(-0.2(G - 19.5))) \quad (1)$$

[2] In this respect, the Catalogue field `visibility_periods_used` indicates the number of groups of observations separated from other groups by at least 4 days.

[3] `phot_bp_rp_excess_factor` is the ratio of the sum of $G_{\rm BP}$ and $G_{\rm RP}$ fluxes over the G flux and should be around one for normal stars.

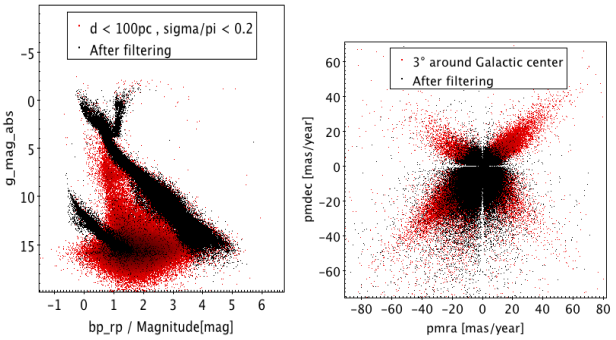

Fig. 11. HR diagram of stars closer than 100 pc (*left panel*) and proper motion diagram near the Galactic centre (*right panel*), with (black) or without (red) filtering of spurious solutions with Eqs. (1) and (2). In both cases, a 20% relative uncertainty limit on the astrometric parameters has also been applied (which generates a "void cross" at the origin on the right).

and

$$1.0 + 0.015(G_{\rm BP} - G_{\rm RP})^2 < E < 1.3 + 0.06(G_{\rm BP} - G_{\rm RP})^2 \qquad (2)$$

By rejecting large χ^2, Eq. (1) helps filtering contamination from double stars, astrometric effects from binary stars, and also from calibration problems. As surprising as it seems, the photometric filtering defined in Eq. (2) cleans the spurious astrometric solutions even more efficiently by detecting perturbations due to close-by sources, but it mostly cleans the faint stars (that would otherwise have been rejected by a filtering on the photometric precision).

To realise that filtering does not come cheap and understand why it could not have been applied for the production of the *Gaia* DR2 catalogue, the fraction of remaining sources is illustrative: in Fig. 11a, filtering keeps 39% of the sources, and only 26% survive in Fig. 11b. The filter of Eq. (2) has the greatest effect; if it had been applied alone, it would have kept only 30% of the sources. Obviously, the fraction of filtered data depends on magnitude and on parallax and proper motion, and it introduces additional selection effects.

Other filters may of course be defined depending on the application. For instance, replacing Eq. (2) by

$$\texttt{visibility\_periods\_used} > 8 \qquad (3)$$

may at first sight appear to remove many good solutions without sufficient observations, but it would be preferable for the bulge proper motions because it would increase the remaining data from 26 to 72% in Fig. 11b while probably cleaning the poor solutions well enough as the influence of crowding on photometry is perhaps not an issue here; but this filtering would not be sufficient for the HR diagram, Fig. 11a, because it would leave too many spurious intrinsically faint stars. If external colours are used for the HR diagram, however, criterion Eq. (2) may not be necessary either. To summarise, the quality filters to be applied may typically be either Eqs. (1) and (2) or Eqs. (1) and (3), depending on whether a photometric filtering is needed.

To end on a positive note, if the fraction of rejected sources may appear very high at first, the probability of a poor solution when taking a star at random is quite low. Spurious solutions produce high astrometric values: selecting high proper motion stars will preferentially select spurious proper motions; making an HR diagram with nearby stars alone will select high parallax values, with a larger fraction of spurious ones. In some other random sample, however, robust statistics may be enough to mitigate their effect.

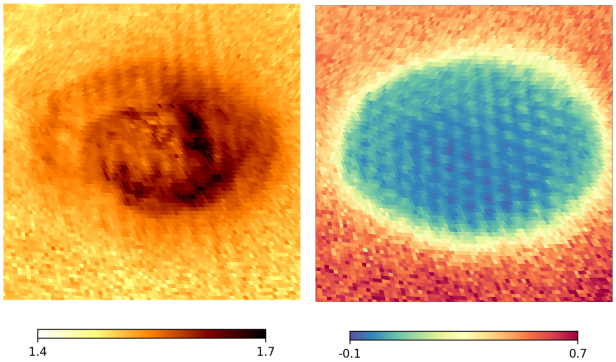

Fig. 12. In the direction of the LMC, the median of the pseudo-colour (*left panel*) shows a similar banding effect as for the median parallaxes (in mas, *right panel*).

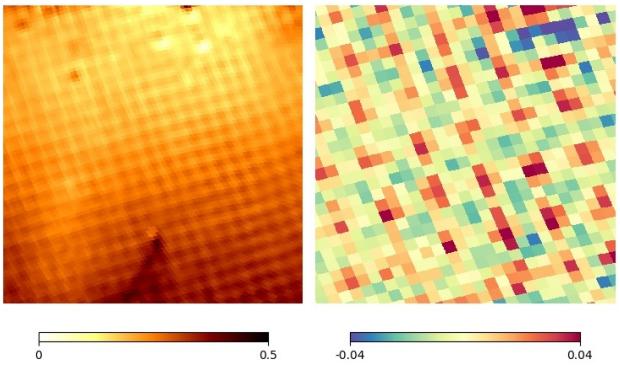

Fig. 13. Small-scale systematics: map of median parallaxes (mas) in a 10° field centred on $(l, b) = (0°, -12°)$ (*left panel*). Residuals (mas) of median parallaxes in field $(1°, -7°)$, size 3° for stars brighter than $G = 17$ only, after subtracting a 0.7° running median (*right panel*).

4.2. Small-scale systematics

Lindegren et al. (2018) showed that spatial correlations are present in the astrometry, producing small-scale systematic errors. In scientific applications, this means that the average parallax or proper motion in a small field will be biased if the systematic error is not accounted for. In practice, they limit the asymptotic precision gain on samples of stars to $\sqrt{\rho}$, i.e. much larger than the expected $1/\sqrt{N}$, where ρ is the correlation between the N sources.

Although probably present throughout the whole sky, these correlations can be more easily seen in fields that mostly consist of distant stars, where the true parallax is small, compared to the parallax error, as for example in dSphs (Gaia Collaboration 2018c), in the direction of the LMC (Fig. 12b) or the bulge (Fig. 13). In the latter field, the scanning law pattern is clearly visible. The systematics have at least a ≈0.02 mas RMS over a ≈ 0.6° period, and they are present for faint as well as for brighter stars. This banding pattern, which produces systematics for parallaxes and proper motions at small angular scale, is more difficult to handle, however, because it changes shape, orientation, and amplitude across the sky.

4.3. Large-scale systematics

The systematics on a larger scale can be studied using QSOs. Compared to the thorough QSO selection made in Lindegren et al. (2018), we used a simpler approach and employed sources

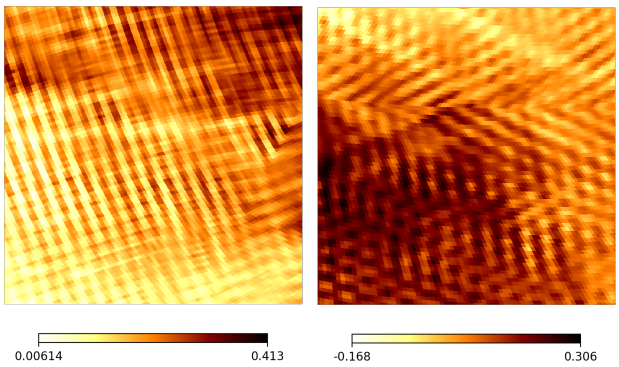

Fig. 14. Correlation $\rho(\varpi, \mu_\delta)$ towards the bulge (*left panel*) and $\rho(\alpha, \delta)$ towards the LMC, same fields as Figs. 13a and 12, respectively.

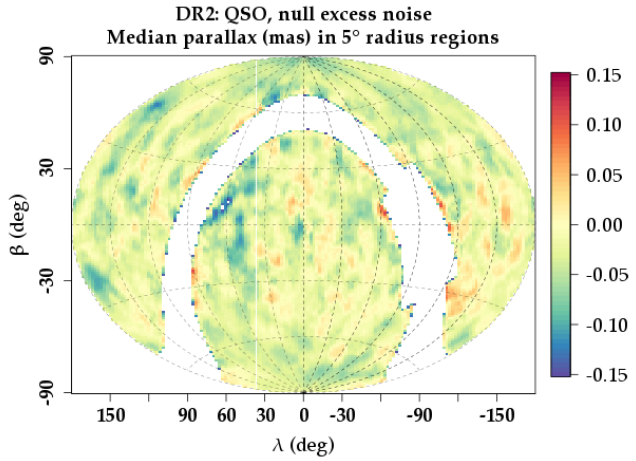

Fig. 15. Variations in QSO parallaxes (mas) in 5° radius fields, ecliptic coordinates. In addition to a ≈−0.03 mas median zero-point, large-scale variations also appear in ecliptic longitude.

identified as QSO in the *Gaia* DR2 catalogue[4]. However, we kept only those with a parallax over uncertainty smaller than 5 in absolute value and a null astrometric excess noise[5] to avoid bad astrometric solutions caused by perturbations by nearby sources, for example. Non-QSO may remain in the sample, however, and we did not pay attention to the results near the Galactic plane where there are not enough sources. Only 220 178 QSO remain in this sample.

In order to estimate how the systematics are spatially represented, median parallaxes have been computed in random regions defined by a given radius and with at least 20 QSOs. Compared with a similar plot created for *Gaia* DR1 (Arenou et al. 2017, Fig. 24), Fig. 15 shows an improvement by a factor at least 3 for the amplitude of the systematics.

There is a significantly negative global zero-point (≈−0.03 mas, cf. Table 1), and some variations appear on a larger scale. For example, in a 5° radius around $(\alpha, \delta) = (191°, 50°)$, the median parallax is -0.105 ± 0.031 mas, to be compared with $+0.028 \pm 0.027$ mas in a 5° radius region around $(\alpha, \delta) = (30°, -8°)$, that is, there can be up to a 0.13 mas peak-to-peak variation of the parallax systematics throughout the sky. A statistical study of these angular correlations has been done by Lindegren et al. (2018, Sect. 5.4). Large-scale systematics are of smaller but not negligible amplitude.

4.4. Global systematics from external comparisons

As in Arenou et al. (2017), we analysed the parallax systematics using a comparison to many external catalogues.

Direct comparison of the parallaxes was made with HIPPARCOS (van Leeuwen 2007), VLBI (Reid & Honma 2014), HST (Benedict et al. 2007; Benedict & McArthur 2015), and RECONS (Henry & Jao 2015) parallaxes (using their database as of January 2018).

Distance moduli were compiled for distant stars, distant enough so that the uncertainty on their parallax is 10 times smaller than that of *Gaia*. Distance moduli from variable period-luminosity relations were obtained for RRab RR Lyrae and fundamental mode Cepheid stars using both GCVS (Samus' et al. 2017) variables and variables directly provided by *Gaia*, using both the supervised classifications and the SOS component of the variability pipeline (Eyer et al. 2017). We used the 2MASS (Skrutskie et al. 2006) magnitudes, using the extinction-independent magnitude $K_{J-K} = K - \frac{k_K}{k_J - k_K}(J - K)$, with k the extinction coefficients, and the period-luminosity relation of Muraveva et al. (2015) for RRLyrae (using the metallicity information from the *Gaia* light curve when available, assuming −1 dex with a dispersion of 0.6 dex otherwise) and of Fouqué et al. (2007) for Cepheids.

Distance moduli were also compiled from spectroscopic surveys, here APOGEE DR14 (Holtzman et al. 2015) and LAMOST DR2 (Luo et al. 2015), using K_{J-K} and Padova isochrones (Bressan et al. 2012; CMD 2.7). A catalogue of distances of SEGUE K giants (McConnachie 2012) was also used. In contrast to the *Gaia* DR1 validation, we no longer provide comparisons with asteroseismic distances because only a few stars have a distance information that is significantly better than that of *Gaia*.

Very distant stars, for which the true parallaxes can be considered to be almost zero, were also compiled through Milky Way satellites confirmed members, mostly using their radial velocities (RVs) for dSph. For the LMC and SMC, the bright subset for which we could use the *Gaia* RVs to confirm their membership was also tested (called LMC/SMC Vr in Table 1). Finally, parallaxes of confirmed QSOs were tested from the ICRF2 (Fey et al. 2015), RFC2016c[6], and LQRF (Andrei et al. 2009) catalogues.

More details about the construction of these catalogues are provided in the online catalogue documentation (Antoja et al. 2018; Sect. 10.4). The results of the comparison are summarised in Table 1. All the catalogues point towards a global zero-point bias in the parallax of about −0.03 mas, with sky variations illustrated by the dSph members (see also Gaia Collaboration 2018c).

For most of the tests, parallax variations with magnitude, colour, and pseudo-colours[7] have been found, depending on whether we consider the weighted mean differences or the normalised differences, indicating correlations with the uncertainty estimates (see Sect. 4.6.3). The strongest correlation of the differences with colour and magnitude is seen with APOGEE. Here, the difference is larger for the redder sources, which are

[4] Catalogue field `frameRotatorObjectType` equal to 2 or 3, that is, a source assumed to be extragalactic and used to determine the reference frame.

[5] The `astrometric_excess_noise`, see Lindegren et al. (2018), expresses the astrometric goodness of fit (GoF) in angular measure. This is the quantity that would be needed to be added to the measurement uncertainties to make the weighted sum of squared residuals equal to or smaller than the degree of freedom. It is >0 only for poor fits.

[6] `http://astrogeo.org/vlbi/solutions/rfc_2016c/`

[7] The `astrometric_pseudo_colour` is an astrometrically determined effective wavenumber given in μm^{-1}, see Lindegren et al. (2018, Sect. 3.1).

Table 1. Summary of the comparison between the *Gaia* parallaxes and the external catalogues.

Catalogue	Nb	Outliers	$\langle G \rangle$	ϖ difference	ϖ uwu
HIPPARCOS	62 484	0.1%	8.3	*−0.118 ± 0.003*	*1.25 ± 0.003*
VLBI	40	2.5%	8.2	−0.07 ± 0.03	*1.9 ± 0.2*
HST	51	*33%*	11.7	−0.01 ± 0.02	*2.1 ± 0.3*
RECONS	432	3%	12.6	*−0.71 ± 0.06*	*1.69 ± 0.06*
GCVS RR Lyrae	197	2%	14.9	*−0.033 ± 0.009*	*1.51 ± 0.08*
Gaia RR Lyrae	795	3%	15.6	*−0.056 ± 0.005*	*1.38 ± 0.04*
Gaia Cepheids	1417	2%	15.6	*−0.0319 ± 0.0008*	*1.53 ± 0.03*
APOGEE	5212	2%	13.9	*−0.048 ± 0.002*	*1.44 ± 0.01*
LAMOST	174	9%	14.9	*−0.040 ± 0.005*	*1.50 ± 0.08*
SEGUE Kg	3151	0.2%	16.5	*−0.041 ± 0.002*	*1.10 ± 0.01*
LMC	51 162	1%	19.2	*−0.038 ± 0.0004*	*1.098 ± 0.004*
LMC Vr	319	4%	12.8	*−0.042 ± 0.001*	*1.34 ± 0.05*
SMC	26 404	2%	16.4	*−0.0268 ± 0.0004*	*1.43 ± 0.006*
SMC Vr	114	8%	12.5	*−0.037 ± 0.002*	*1.4 ± 0.1*
Draco	427	0%	19.3	*−0.047 ± 0.008*	*1.08 ± 0.04*
Ursa Minor	78	0%	17.4	*−0.054 ± 0.008*	1.03 ± 0.08
Sculptor	1287	0.3%	19.1	*−0.028 ± 0.006*	*1.11 ± 0.02*
Sextans	375	0.3%	19.3	*−0.09 ± 0.02*	*1.07 ± 0.04*
Carina	864	0%	19.8	*−0.020 ± 0.007*	1.05 ± 0.03
Crater2	63	0%	19.1	*−0.06 ± 0.03*	*0.96 ± 0.09*
Fornax	2659	0.4%	18.8	*−0.052 ± 0.004*	*1.18 ± 0.02*
CVnI	51	0%	20.0	-0.030 ± 0.08	0.91 ± 0.09
LeoII	123	0%	19.5	0.05 ± 0.05	1.0 ± 0.06
LeoI	292	0.7%	19.6	*−0.23 ± 0.05*	*1.30 ± 0.05*
Phoenix	81	0%	20.6	0.09 ± 0.07	1.07 ± 0.08
all dSphs	6300	0.3%	19.0	*−0.044 ± 0.002*	*1.13 ± 0.01*
ICRF2	2347	0.3%	18.8	*−0.031 ± 0.003*	*1.16 ± 0.02*
RFC2016c	3523	0.3%	18.9	*−0.031 ± 0.002*	*1.15 ± 0.01*
LQRF	79 631	0.04%	19.1	*−0.0322 ± 0.0008*	*1.088 ± 0.003*

Notes. We present the total number of stars used in the comparison (Nb) as well as the percentage of outliers excluded (at 5σ, in italic if higher than 10%) as well as the median G of the sample. The parallax differences ($\varpi_G - \varpi_E$, in mas) and unit-weight uncertainty (uwu) that needs to be applied to the uncertainties to adjust the differences are indicated in italic when they are significant (p-value limit: 0.01).

also the faintest, which may be due to systematics linked to the isochrones used. For the Cepheids, variations with astrometric excess noise and GoF are present, indicating possible contamination with binaries.

The analysis of about 200 clusters, including open clusters (OCs) and globular clusters (GCs), also shows a residual zero-point in parallaxes. We computed the differences between the actual DR2 value and the reference value for all the stars in the clusters using the DAML (Dias et al. 2014) and MWSC (Kharchenko et al. 2013) catalogues. The difference depends on the reference catalogue, on the distance of the cluster, and on the colour of the stars. On the average, for the whole cluster sample, the residual zero-point is $\varpi_{Gaia} - \varpi_{\rm reference} = -0.067 \pm 0.12$ mas for MWSC and -0.064 ± 0.17 mas for DAML. Figure 16 shows the comparison for the two catalogues.

A comparison with literature values for the proper motion of clusters is presented in Fig. 17. The DAML and MWSC proper motion catalogues present significant differences for the same clusters. While average zero-point differences are smaller than 1 mas yr^{-1}, the dispersion around this value can be on the order of 3–4 mas yr^{-1} (see Dias et al. 2014; for a discussion). This is reflected in the comparison with *Gaia* DR2 proper motions. We find that the residual zero-point is $\mu_{Gaia} - \mu_{\rm reference} = 0.51 \pm 0.17$ mas yr^{-1}, and 0.25 ± 0.12 mas yr^{-1} for $\mu_\alpha \cos\delta$ and μ_δ for the MWSC, while the analogous quantities for DAML are 0.0 ± 0.19 and 0.41 ± 0.18 mas yr^{-1}. These values are consistent with the differences between the two catalogues. On the basis of this comparison, we have no evidence for a significant residual proper motion zero-point in *Gaia* DR2.

4.5. Managing systematics

For samples on a small spatial scale, one first concern is how to evaluate the systematics. Figure 14a compared to Fig. 13a and Fig. 14b compared to Fig. 12b show that, locally, some hint of astrometric non-uniformity may perhaps be indicated by local variations of the astrometric correlations.

Although significant variations of the parallax zero-point with magnitude and colours is probably present, for example for the QSO parallaxes versus $G_{\rm BP}-G_{\rm RP}$ colour in Fig. 18a, the trend is most obvious with the astrometric pseudo-colour, Fig. 18b: about 0.05 mas peak-to-peak for QSO. The amplitude is even much larger on a subset of sources in the LMC direction[8], Fig. 18d. This cannot be due to contamination by foreground sources, as the parallax peak-to-peak variation with $G_{\rm BP}-G_{\rm RP}$ (0.05 mas, Fig. 18c) is one order of magnitude smaller than with pseudo-colour (0.6 mas, Fig. 18d). The pseudo-colour

[8] A sample of 1.56 million sources in a 3° radius around $(\alpha, \delta) = (80°, -69°)$, keeping only those with null astrometric excess noise and $(\varpi, \mu_\alpha, \mu_\delta)$ within 4σ of (0.02–0.03, 1.8, 0.2), which accounts both for the average LMC values and the average DR2 parallax zero-point.

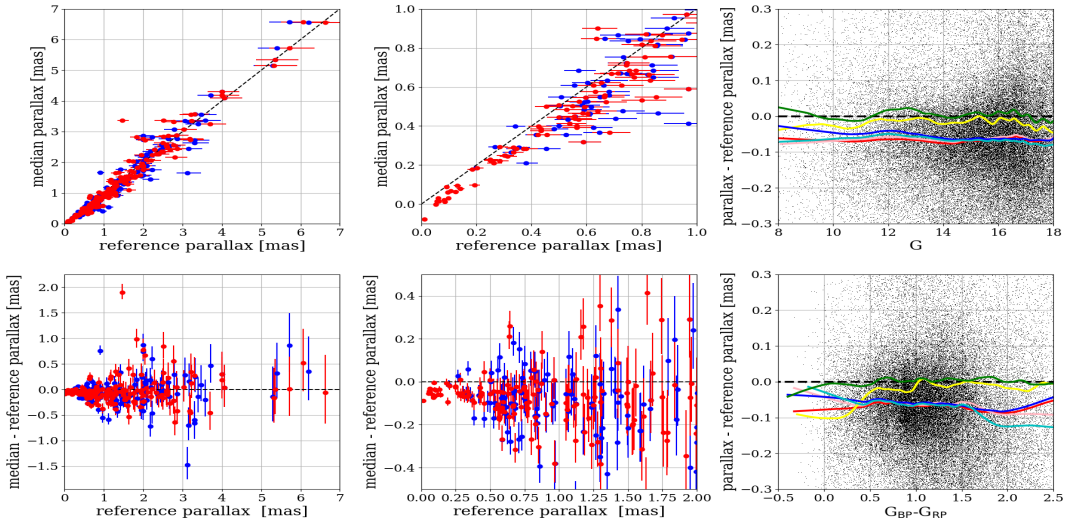

Fig. 16. Parallax comparisons with clusters using the MWSC and DAML catalogues. *Upper left panel*: cluster median DR2 parallaxes compared with literature: MWSC (red dots) and DAML (blue dots). *Upper central panel*: same as left panel, restricted to $\varpi < 1$ mas. *Upper right panel*: parallax differences for the cluster star sample as a function of G (black dots). *Lower left panel*: parallax differences for the whole sample. *Lower central panel*: same as left panel, for $\varpi < 2$ mas. The *right panels* are analogous for the whole star sample (black dots). *Lower right panel*: parallax differences as a function of the colour ($G_{\rm BP} - G_{\rm RP}$). In the *right panels*, lines show the smoothing for the two reference catalogues for different distances. The red line indicates the whole MWSC, the pink line shows MWSC OCs with a distance $d > 1000$ pc, the yellow line represents MWSC OCs with $d < 500$ pc, the blue line shows all OCs in DAML, the cyan line shows DAML with $d > 1000$ pc, and the green line represents DAML with $d < 500$ pc.

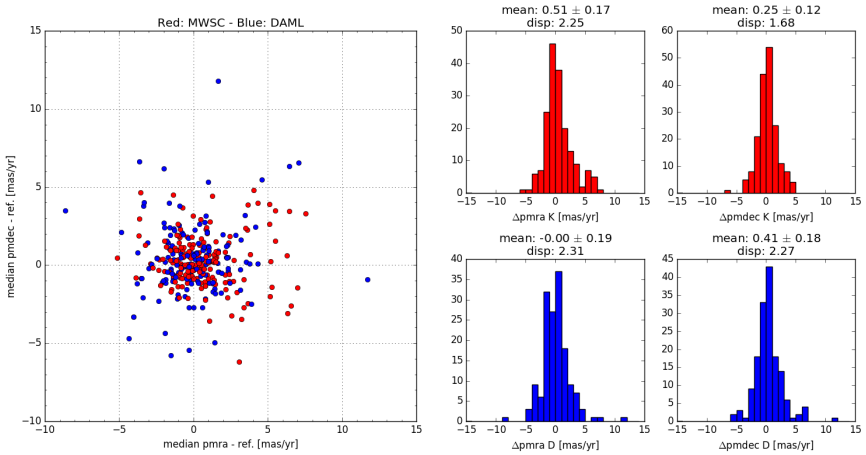

Fig. 17. Difference between the median DR2 proper motion of the clusters in μ_δ vs. $\mu_\alpha \cos\delta$ for DAML (blue) and MWSC (red) sample (*left panel*). Distribution of the differences in $\mu_\alpha \cos\delta$ and μ_δ for MWSC (*upper right panel*) and analogously for DAML (*lower right panel*).

has absorbed a fraction of the astrometric systematics, as can be seen from their spatial variations on the LMC, Fig. 12a. Pseudo-colour variations, when they are not representative of the colour variations themselves, may then help to detect astrometric systematics.

It could also be tempting to "correct" the parallaxes for the global zero-point. Individually, this would be inadvisable, first because its contribution level is generally below the random error; second, as seen above, the local (Sect. 4.2), regional variations (Sect. 4.3), or colour and magnitude effects may be larger than the global zero-point. However, for a sample that is well distributed over the whole sky and that is used for a luminosity calibration, for instance, the zero-point may be corrected, or, better, solved for, as mentioned by Lindegren et al. (2018).

4.6. Uncertainties of the astrometric random errors

4.6.1. Distribution of the astrometric errors

The astrometric error distribution, at least for faint sources, can be studied using the QSO. We used sources with `frame_rotator_object_type` equal to 2 or 3, keeping only those with a parallax over uncertainty smaller than 5 in absolute value and keeping even those with non-zero excess noise (488 805 sources). The statistical distribution of the errors (parallax over uncertainty) can then be directly seen, and the deviation from normality beyond 2σ that was present in *Gaia* DR1 has now disappeared in *Gaia* DR2 (Fig. 19), the errors being now much more Gaussian. This legitimates the use of the normal distribution in likelihood functions where the astrometric errors appear.

4.6.2. Internal comparisons

A simple test on astrometric precision is a comparison of parameters for the duplicated source pairs mentioned in Sect. 2.2. Figure 20 shows histograms for the normalised differences of right ascensions and parallaxes for sources brighter than 17 mag. The properties for declinations are similar to those for right ascensions, and the proper motion components show features similar to the parallaxes. For sources with the full five-parameter astrometric solution for both solutions, the comparison in Fig. 20

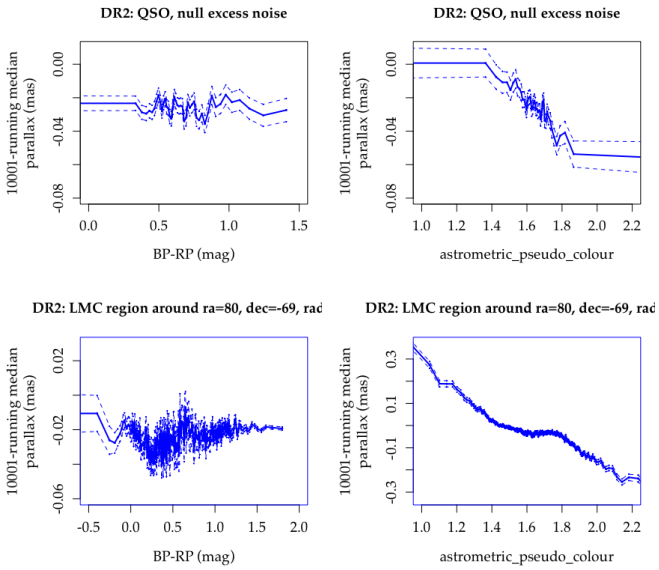

Fig. 18. Running median of parallaxes (mas) on 10001 points with $\pm1\sigma$ standard error. No large parallax systematics are linked to the (G_{BP}−G_{RP}) colour when we use QSO (*top left panel*) or a sample of stars in the LMC direction (*bottom left panel*). However, a significant correlation of the parallaxes with the pseudo-colour is present for QSOs (*top right panel*), and it is much larger in the LMC direction (*bottom right panel*).

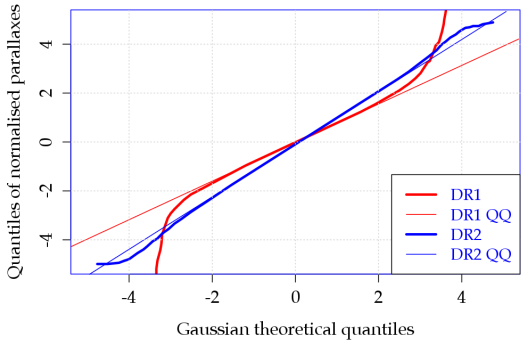

Fig. 19. Normalised QSO parallaxes truncated to ±5. As seen using a Gaussian Q–Q plot, the parallax errors of *Gaia* DR2 QSOs (blue) are now much closer to the Gaussian (0,1) diagonal than DR1 parallaxes (red). The thin lines represent the lines passing through the first and third quartiles of the corresponding samples.

only suggests that formal uncertainties are slightly underestimated, at perhaps 10%. However, for sources where the full solution for some reason failed for one or both solutions, the differences are non-Gaussian and show very large wings. As discussed above, Sect. 4, these sources may be binaries, show structure, or have only a few observations. As mentioned in Sect. 2.2, the duplicated sources are more affected than the average sources with only two astrometric parameters. This may explain why the uncertainties are so strongly underestimated for this specific subset.

The formal uncertainties on the parallaxes have also been studied using the dispersion of the negative parallax tail estimated by deconvolution (see Arenou et al. 2017, Sect. 6.2.1 for details). The unit-weight uncertainties[9] are shown Fig. 21 for several illustrative subsets of the *Gaia* DR2 catalogue, as a

[9] We call unit-weight uncertainty (uwu) the factor (ideally one) that needs to be applied to the published uncertainties to be equal to the estimated dispersion of the astrometric parameters.

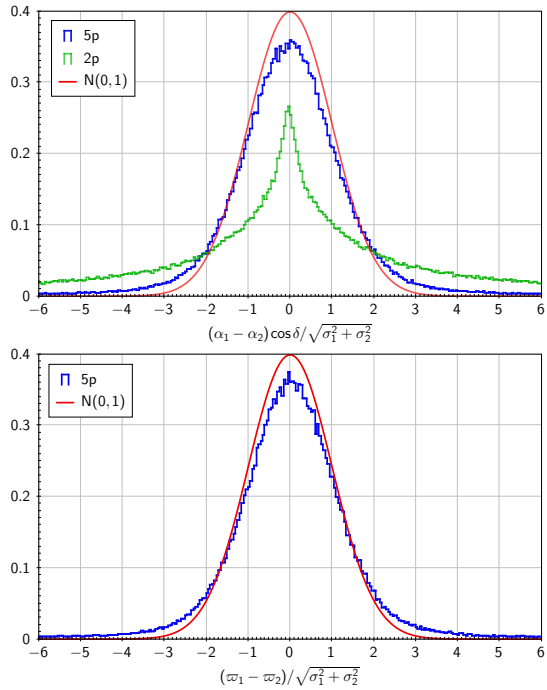

Fig. 20. Histograms for the normalised differences for right ascensions (*top panel*) and parallaxes (*bottom panel*) for duplicate source pairs brighter than G = 17 mag. A normal distribution is overplotted.

function of these uncertainties. Although the formal uncertainties primarily increase with magnitude (see e.g. Lindegren et al. 2018, Fig. 9), they also increase with the astrometric excess noise. Non-zero excess noise can be due either to the non-single character of the source or to imperfect calibrations (Lindegren et al. 2012). In addition, `phot_bp_rp_excess_factor` is an indication of binarity or duplicity (contamination) in dense fields, or of extended objects (Evans et al. 2018). From top to bottom (largest unit-weights to smallest ones), the subsets with non-zero excess noise and large colour excess factor most probably represent the contribution from non-single stars and extended objects, respectively, showing that the added excess noise was not enough to compensate for the actual dispersion. Below these sources lie the subsets in the direction of the Galactic centre and LMC, which probably show the effect of contaminated sources in dense fields. Lower still lie the duplicated stars, which mostly consist of single stars for small uncertainties, with a possible contribution of binaries for larger uncertainties. For faint, average stars, the unit-weight is too large by only about 15%. Then the QSO uncertainties look most realistic, as is also shown in Fig. 22.

For all subsets, the unit-weight increases towards small uncertainties (i.e. for stars with magnitude between 13 and 15), which could be underestimated by about 40%. This probably originates from the reweighting that has been applied to the uncertainties (Lindegren et al. 2018; Appendix A) to correct a bug found too late in the data processing cycle. It appears that this reweighting correctly improved the uncertainty estimates of stars brighter than G = 13, but had an adverse effect for stars with `astrometric_n_obs_ac` = 0, that is, stars fainter than 13 (stars $13 \lesssim G \lesssim 15$ have the smallest uncertainties); reverting to the previous uncertainties for stars with `astrometric_n_obs_ac` = 0 may then be safer.

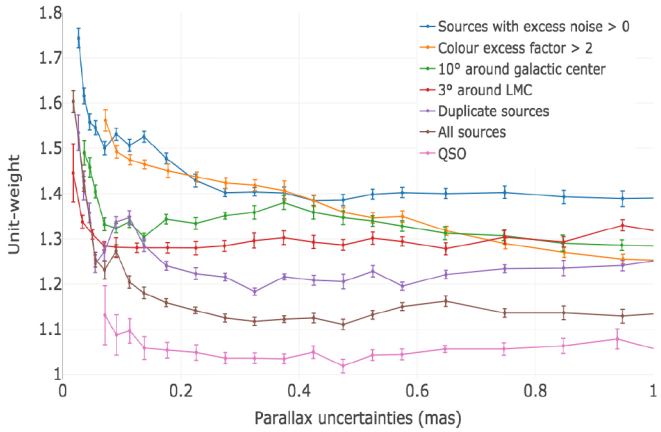

Fig. 21. Unit-weight uncertainties of parallaxes estimated by deconvolution vs. these uncertainties. From top to bottom, sources with non-zero excess noise (blue), `phot_bp_rp_excess_factor` larger than 2 (orange), within 10° of the Galactic centre (green), within 3° towards the LMC (red), duplicated sources (violet), all catalogue sources (brown), and QSO (pink). Only sources with more than eight visibility periods and GoF < 5 have been kept in all subsets, except for the subset with non-zero astrometric excess noise, where no GoF upper limit was applied.

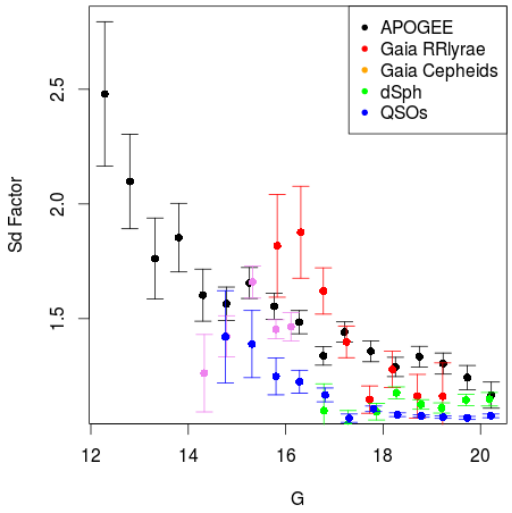

Fig. 22. Unit-weight uncertainty that would need to be applied to the *Gaia* parallax uncertainties to be consistent with the residual distribution versus APOGEE, *Gaia* RRlyrae, and Cepheid distance moduli as well as dSphs and LQRF QSOs.

4.6.3. Comparison to distant external data

The uncertainties have also been tested in a comparison to distant stars or QSOs (Table 1). These comparisons are complicated by the fact that the uncertainties of the external catalogues may not be accurately determined and by the pollution from incorrect identifications for QSOs. Still, the under-estimation of the parallax uncertainties is seen to increase with magnitude in all the tests we conducted, as illustrated in Fig. 22. This is the same trend as shown in Fig. 21, but of a larger amplitude for the reasons just explained. For bright stars, however, the comparison with HIPPARCOS, Table 1, shows that the parallax uncertainties are unlikely to be much underestimated, as HIPPARCOS parallax uncertainties themselves may well have been slightly underestimated.

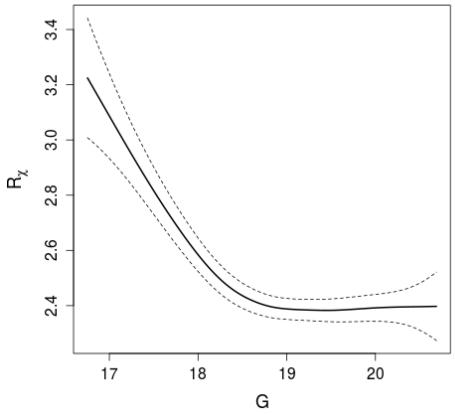

Fig. 23. χ^2 test of the LQRF QSOs proper motions as a function of G magnitude. The residual R_χ should follow a χ^2 with 2 degrees of freedom. The dotted lines correspond to the 1σ confidence interval. The correlation observed here is most likely due to the underestimation of the uncertainties as a function of magnitude.

The variation of uncertainties with magnitude explains why, depending on how the uncertainties are handled in computing the differences with external data, a correlation between the difference and the magnitude is seen or not. The most striking example is produced when a χ^2 test is used: while no significant correlation of the individual proper motions of QSOs is seen with magnitude, the correlation is significant when they are combined through their covariance matrix (Fig. 23). This is also seen in the comparison with the HIPPARCOS proper motions, similarly to what was found in *Gaia* DR1 (Arenou et al. 2017).

4.6.4. Parallax and proper motion precision tested using distant clusters

The astrometric precision has also been estimated using a sample of about 200 OCs and about 20 GCs. We used the DAML and MWSC catalogues as reference values. We calculated the dispersion of parallaxes from the median value for each cluster after normalising the offsets by the nominal uncertainties, selecting only stars with uncertainties on parallax smaller than 2 mas. Figure 24 shows the median absolute deviation (MAD) of the above distribution as a function of the parallaxes for open and globular clusters. Nearby clusters show a clear internal parallax dispersion. However, the MAD does not really converge to one for distant clusters either, as would be expected if the uncertainties on the parallax were correctly estimated. The results suggest that the uncertainties are underestimated. This holds in particularly at the bright end of the star distribution (for $G < 15$). The uncertainties are definitively underestimated for GCs (see Fig. 24): clearly, the high crowding is responsible for the fewer observations per star and the degraded astrometric precision.

A residual parallax trend with colours is shown in Fig. 25 for all the stars in the cluster sample in the blue edge and possibly in the red edge of the colour domain, although with a poor statistics in this case. This could be a consequence of an imperfect chromaticity correction. However, since the majority of stars in our sample are on the main sequence, magnitude, and colour are strongly correlated, and it is hard to distinguish between the two effects.

Figure 26 shows the MAD of the distribution of the proper motions in right ascension and in declination, normalised by the nominal uncertainties as a function of the parallax. Nearby

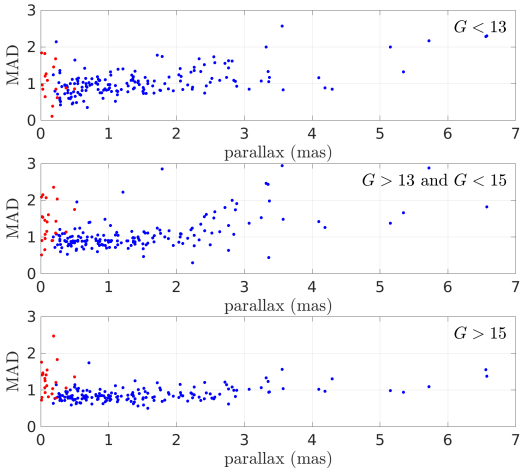

Fig. 24. MAD of the distribution of the parallaxes normalised by the nominal uncertainties as a function of the parallaxes (in mas) for open clusters (blue dots) and globulars (red dots). *Top to bottom panels*: $G < 13$, $13 < G < 15$, and $G > 15$.

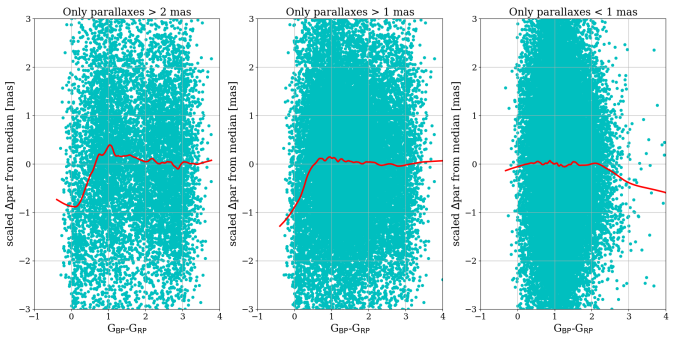

Fig. 25. Distribution of the differences between median cluster parallaxes and single-star *Gaia* parallaxes normalised by the nominal uncertainties for stars with $\varpi > 2$mas (*left panel*), $1 < \varpi < 2$mas (*central panel*), and $\varpi < 1$mas (*right panel*). The red lines are the smoothed distributions.

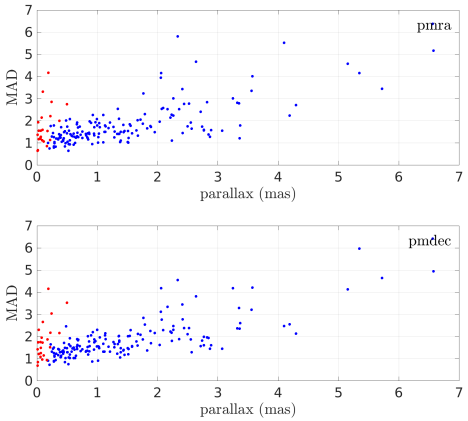

Fig. 26. MAD of the distribution of normalised differences on the proper motions (mas yr^{-1}) in right ascension (*top panel*) and declination (*bottom panel*) as a function of the parallax (in mas) for open clusters (blue dots) and globulars (red dots).

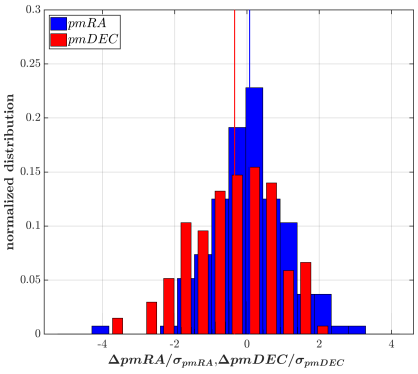

Fig. 27. Normalised differences in Omega Cen cluster proper motions (ra in blue and dec in red) between *Gaia* DR2 data and HST data. The lines represent the mean distribution values.

clusters are affected by intrinsic proper motion dispersion, while distant clusters tend to MAD = 1, but with a large dispersion. At small parallax, all the objects belonging to the tail that have MAD > 1.5 are globulars, implying that proper motion uncertainties are also underestimated in the central regions of this type of cluster.

We compared the proper motions in Omega Cen with external HST data by Libralato et al. (2018), where relative proper motions are available down to very faint magnitudes, and a proper motion zero-point is provided. About 140 stars were found in common. The studied field is located at the outskirts of the cluster and it is not very crowded. We compared the normalised dispersion of the differences in proper motions for the stars in both samples. The normalised dispersion is very close to one both for $\mu_{\alpha*}$ and μ_δ, implying that the proper motion uncertainties are correctly estimated (see Fig. 27).

Finally, we checked the quality of the astrometry for the unresolved photometric binary sequence that is clearly visible in the CMDs (see e.g. Fig. 44) for about 12 OCs selected after visual inspection and located farther away than about 400 pc. This would minimize the effect of the internal velocity dispersion and of the mass segregation. The procedure and a few examples are discussed in Sect. 5.2.2. We derived the deviation from the cluster median for every star in the main sequence and in the binary star sequence in proper motion space. The global distributions are shown in Fig. 28. For all the OCs, the Kolmogorov–Smirnov test does not reject the null hypothesis that the two samples are drawn from the same distribution at the 5% significance level for $\mu_{\alpha*}$ and for the parallaxes, while for μ_δ, the null hypothesis is rejected only for two objects with marginally inconsistent p-values of 4.8% and 3.8%. This shows that the astrometric quality of unresolved binaries with a small magnitude difference is not significantly different from that of single stars.

4.7. Quality indications of the astrometric solution, and outliers

Gaia DR2 includes many quality indicators for the astrometric solution, as described in detail in Lindegren et al. (2018), and we mention only a few here. An obvious indicator is the number of rejected observations (`astrometric_n_bad_obs_al`) as compared to the retained observations (`astrometric_n_good_obs_al`). Figure 29 shows the fraction of outliers throughout the whole sky and in a small area. The sky map shows whole great circles with more than average rejections. This points to specific time intervals of maybe half a day with a specific problem for the astrometric calibration. This is also clear in the zoom, where we see examples of two or three consecutive scans, each 0.7° wide, with some issue.

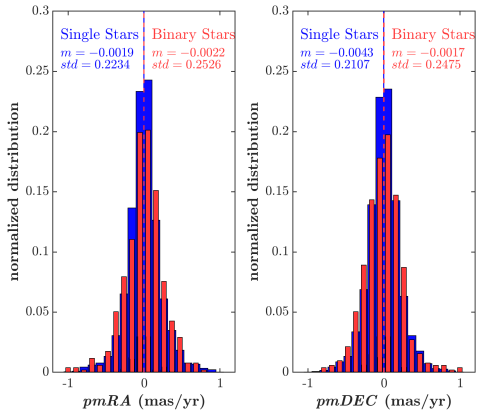

Fig. 28. Distribution of the deviation from the cluster median of the proper motions in the main sequence (in blue) and photometric binaries (in red) for a sample of 12 OCs.

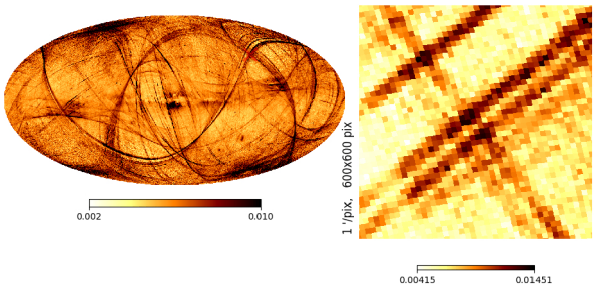

Fig. 29. Mean fraction of outlying CCD measurements throughout the whole sky (*left panel*) and 10° size detail near $(l, b) = (-94.^{\circ}4, 40.^{\circ}7)$. Several bad scans or larger regions have a larger portion of outliers.

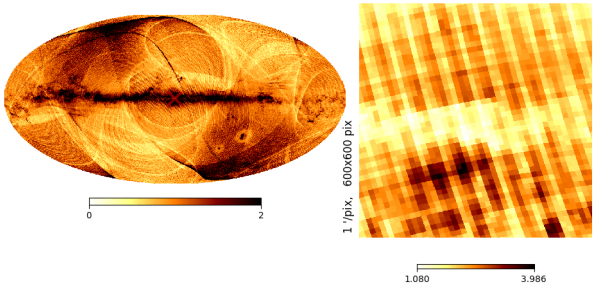

Fig. 30. Mean value of astrometric excess noise divided by parallax uncertainty. *Left panel*: whole sky. *Right panel*: 10° size detail at the Galactic centre (0,0). Imprints of the scanning law are present throughout the whole sky.

Another useful quality parameter is the excess noise, `astrometric_excess_noise`, which expresses in angular measure the insufficiency of the source model to match the observations. This is illustrated in Fig. 30, which shows the excess noise normalised by the parallax uncertainty. It demonstrates that specific zones have a degraded quality of the astrometric solution irrespective of the random errors. It again points to the scanning pattern, but not to specific time intervals because we do not see problematic great circles.

Scanning patterns in quality maps are not a concern by themselves, given that the astrometric solution depends on well-distributed scans. However, we note that similar patterns are visible in the parallax itself, for example, as shown in Figs. 12b and 13.

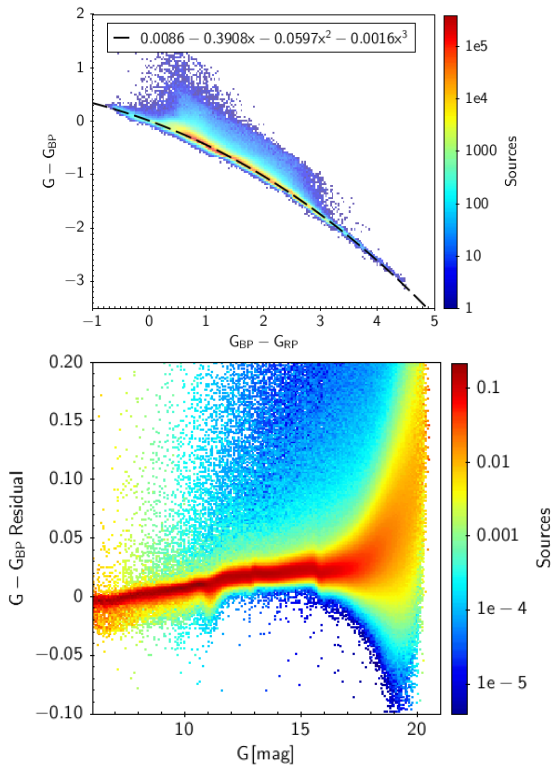

Fig. 31. *Top panel*: $G - G_{\rm BP}$ relation vs. $G_{\rm BP}-G_{\rm RP}$. *Bottom panel*: 2D histogram of the $G - G_{\rm BP}$ residuals after subtracting the colour-dependent relation shown in the *top panel*. Only sources at high Galactic latitudes were used, and the histogram was reweighted to give the same weight to each magnitude interval.

5. Photometric quality of *Gaia* DR2

The photometry in *Gaia* DR2 consists of three broad bands: a G magnitude for all sources, and a $G_{\rm BP}$ and $G_{\rm RP}$ magnitude for the large majority. The photometry and its main validation is described in Evans et al. (2018), and we here merely present some additional tests. Like astrometry, photometry has been greatly improved since *Gaia* DR1 as a result of better calibrations, better image parameter determination, and the availability of colours. We tested the photometric quality of *Gaia* DR2, accuracy and precision, with both internal methods (using *Gaia* photometry only) and by comparisons to external catalogues.

5.1. Photometric accuracy

5.1.1. Internal comparisons

Figure 31b shows a comparison between the G magnitude and the $G_{\rm BP}$ magnitude at high Galactic latitudes. The differences depend on the spectral type, and as a first approximation, the colour dependence (Fig. 31a) was subtracted. The comparison shows a trend with magnitude of a few mmag mag^{-1}, which corresponds to an increasing brightness in $G_{\rm BP}$ for fainter sources. The trend is even stronger for the faintest sources. The small kinks at magnitudes 11, 13, and 15.5 are discussed in Evans et al. (2018) and correspond to changes in the on-board windowing. Although the comparison is presented for $G-G_{\rm BP}$, we cannot distinguish if the bias comes from G, $G_{\rm BP}$, or $G_{\rm RP}$. A $G_{\rm RP}$ comparison would give the same trend and kinks because of the way the colour dependence is subtracted.

As discussed in Evans et al. (2018) and in Sect. 5.5.2 of the on-line catalogue documentation, the $G_{\rm BP}$ and $G_{\rm RP}$

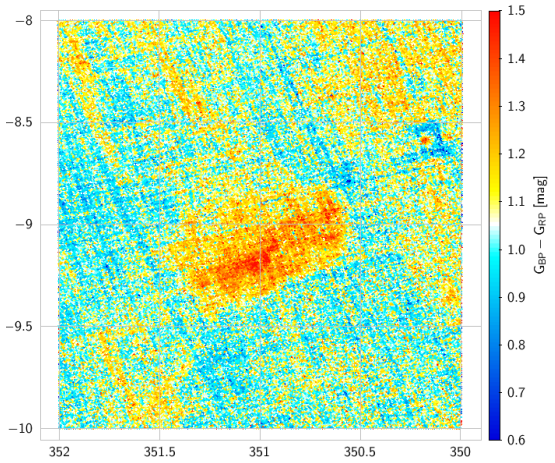

Fig. 32. Median colours, $G_{\rm BP}$−$G_{\rm RP}$, in a dense field (Galactic coordinates) showing artefacts from the scan pattern.

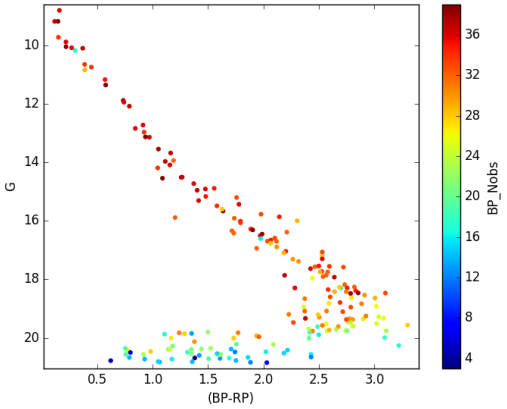

Fig. 33. CMD of Alessi 10, where the colour map shows the number of $G_{\rm BP}$ observations per CCD used in the data analysis, $BP_{\rm Nobs}$.

bands may have an uncorrected flux excess, as indicated by `phot_bp_rp_excess_factor`. This occurs especially in dense fields, for binaries, near bright stars, and for the fainter sources. This excess partly arises because the sky background level is underestimated and partly because overlapping spectra were not deblended. An example of distorted colours in a dense field is given in Fig. 32, which shows artificial patterns originating in the individual scans (the streaks), while the red blob is an area with few stars and therefore probably a real feature (a cloud). Fig. 33 shows the CMD of the cluster Alessi 10 as another example. At the faint end, $G > 18$−19, the main sequence is excessively blue; there are fewer $G_{\rm BP}$ observations for these stars than on average of the cluster.

5.1.2. Comparisons with external catalogues

We compared *Gaia* DR2 photometry to the HIPPARCOS, *Tycho*-2, 2MASS (Skrutskie et al. 2006), SDSS tertiary standard stars of Betoule et al. (2013), and Pan-STARRS1 (PS1; Chambers et al. 2016) photometry, selecting only low-extinction stars ($E(B-V) < 0.015$) using the 3D extinction map of Capitanio et al. (2017). We also compared to OGLE data in regions of relatively homogeneous extinction. An empirical robust spline regression was derived that models the global colour–colour relation. The residuals from these models are plotted as a function of magnitude in Fig. 34. Comparison with 2MASS shows the effect of the 2MASS J band saturation at $J = 9$ mag of their "Read 2-Read 1" frames, rather than a possible *Gaia* issue, and is therefore not shown here.

In the G band, a strong saturation effect at $G < 6$ is visible in the comparison with HIPPARCOS and *Tycho*-2 (see also Evans et al. 2018). The strong increase of the residuals for the faint stars, seen with SDSS in Evans et al. (2018), is confirmed here with PS1 as well as with OGLE data. A small dip at $G \sim 16$, seen in Fig. 31b, is also present in the comparison with PS1, indicating that it is a feature of the G band alone.

$G_{\rm BP}$ starts to deviate at $G_{\rm BP} \sim 18$ in low-density regions (Fig. 34b) because the sky background level is underestimated. No strong feature is seen in $G_{\rm RP}$.

The small global increase (∼2 mmag mag^{-1}) of the residuals with magnitude in the internal comparison (Fig. 31) is seen with the external catalogues, but is much more difficult to follow because each of the catalogues covers a relatively small interval. Moreover, we applied our internal comparison tests to the external catalogue photometry and also found global variations of this order of magnitude for PS1 and larger for SDSS (up to 10 mmag mag^{-1}).

The variation of the residuals with magnitude is much stronger for the blue stars, as illustrated in Fig. 35. The colour–colour relation of brighter stars is more dispersed than that of faint stars, and the difference versus faint stars decreases with increasing magnitude up to $G \sim 11$, where a jump of around 0.02 mag occurs, much larger than seen in the global Fig. 31. Comparison with 2MASS photometry indicates that the issue lies in the G band, but its cause is not yet known.

Recently Weiler (2018) has studied *Gaia* DR2 photometry for spectrophotometric standard stars finding a trend for G of 3.5 mmag mag^{-1} and an offset between bright and faint stars in $G_{\rm BP}$ of 20 mmag. These findings are in agreement with our Figs. 31 and 35.

5.2. Photometric precision

5.2.1. Internal comparisons

The duplicated sources, see Sect. 2.2, have been used for a simple test of the published uncertainties for the three broad-band magnitudes. It was found that even for pairs of good astrometric quality, that is, for two full astrometric solutions, the uncertainties appear underestimated. The normalised magnitude differences are best understood if an uncertainty floor of 2.3 mmag is added in quadrature to the magnitude uncertainties (see the on-line documentation, Antoja et al. 2018, Fig. 10.12). This test was made on a bright subset ($G < 17$ mag), and as described in Sect. 2.2, the duplicated sources need not be representative for the catalogue as such. The apparent inconsistencies between the magnitudes are not yet understood, and we therefore refrain from any recommendation regarding the use of a floor for the magnitude uncertainty.

5.2.2. Photometric precision using clusters

The quality of the *Gaia* DR2 photometry using open and globular clusters was verified for stars brighter than $G \sim 18$, limiting our diagnostics to clusters with little extinction. Cluster members were always selected using *Gaia* DR2 proper motions and parallaxes. We assessed the quality of the $G_{\rm BP}$/$G_{\rm RP}$ photometry by estimating the width of the sequence in the CMD of clusters with secure membership. A downside of having to rely on secure membership is that it is difficult to provide diagnostics for

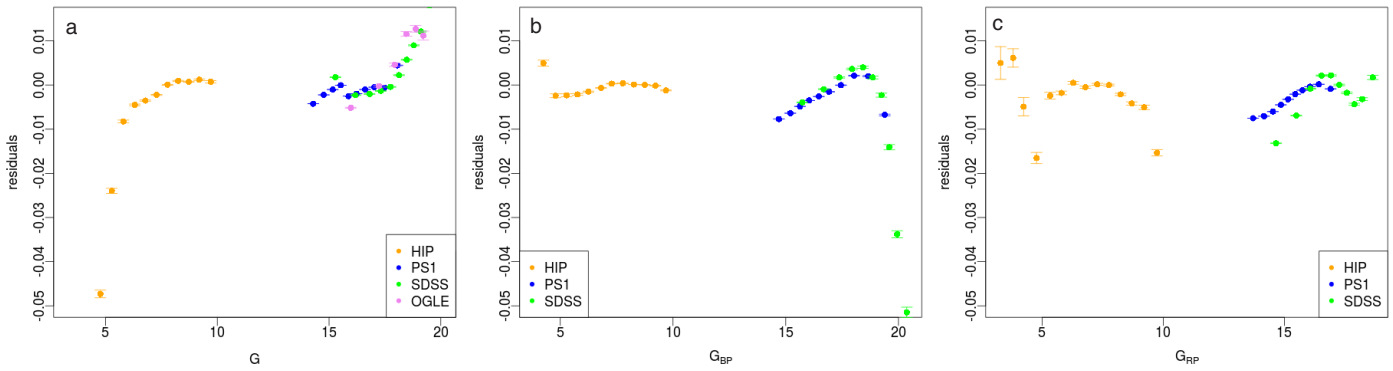

Fig. 34. *From left to right panels*, G, G_{BP}, and G_{RP} photometry versus external photometry: HIPPARCOS (orange), SDSS (green), PS1 (blue), OGLE (magenta). *Panel a*: $G - r$ residuals of the global $G - r = f(g - i)$ spline for SDSS and PS1, $G - V$ residuals of the global $G - V = f(V - I)$ spline for OGLE. *Panel b*: G_{BP}−g residuals of the global G_{BP}−$g = f(g - i)$ spline. *Panel c*: G_{RP}−z residuals of the global G_{RP}−$z = f(g - i)$ spline. For HIPPARCOS the residuals are computed versus the $X - Hp = f(V - I)$ spline, where X denotes respectively G, G_{BP}, and G_{RP}. The zero point of those different residuals is arbitrary.

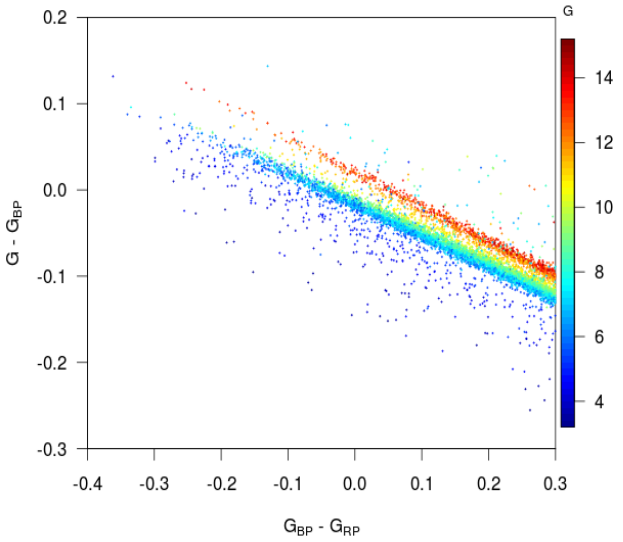

Fig. 35. Colour–colour relation for hot stars, using low-extinction stars ($E(B - V) < 0.015$) with $\varpi/\sigma_\varpi > 10$ and $M_G < 2.5$, colour-coded according to the mean G magnitude.

stars fainter than $G \sim 18$, as the photometry/astrometry (and thus our ability to distinguish cluster stars from field stars) strongly decreases in quality at that magnitude. Binary stars were first selected and removed. We used a LOWESS fitting (Cleveland 1979) to follow the sequence, and removed binary star candidates by clipping out sources with G_{BP} fluxes two error bars lower and G_{RP} fluxes higher than the fitted relation. Then we derived the dispersions of the relations G_{BP} and G_{RP} versus G. We obtained very clean sequences for 12 OCs. An illustration is shown in Fig. 36 for NGC 2682.

The typical dispersion in both G_{BP} and G_{RP} is of the order of 0.02 mag. We restrict our analysis to stars in the un-evolved part of the main sequence, to avoid evolutionary effects. Scaling the difference to the fitted relation by the individual error bars of each star, we find a unit-weight uncertainty of 1.3 on average for G_{BP} and 1.5 for G_{RP}. Because of effects such as rotation, magnetic field, and stellar activity, the main sequence has a natural width that is difficult to estimate since it may vary from one cluster to another. This means that what we derive is an upper limit to the uncertainties on the photometry. Our result suggests that the uncertainties on magnitudes in both filters are correctly estimated or only slightly underestimated. In a few cases, we detected a wide main sequence where a comparison with extinction maps (Schlegel et al. 1998) strongly suggests differential extinction across the field (see e.g. Fig. A.8 of Gaia Collaboration 2018a).

The quality of the photometry is substantially degraded in the inner regions of globular clusters (inside the core radius) because of high crowding. This effect is absent in the external regions. Fig. 37 gives an example of the magnitude/colour shift between the inner and outer regions of globular cluster NGC 5286.

5.3. Photometric quality indicators and outliers

An extensive discussion on the G_{BP}/G_{RP} flux excess factor, `phot_bp_rp_excess_factor`, is provided in Sect. 8 of Evans et al. (2018). As it is sensitive to contamination by close-by sources in dense fields, binarity, background subtraction problems, and extended objects, Evans et al. (2018; Eq. (1)) recommend to use it, with a colour term such as in Eq. (2), to filter the photometry from outliers. As we showed in Sect. 4.1 above, this also has a beneficial effect for astrometry.

To avoid the background issues for the faint G_{BP} stars, the $G - G_{RP}$ colour can be used. However, counter-intuitively, in crowded areas with stars that have a roughly similar spectral type (e.g. a selection of distant bulge stars that consists of mainly red giants), the CMD using G_{BP}−G_{RP} colour distribution will look reasonable (Fig. 38a), but not the CMD using $G - G_{RP}$ (Fig. 38b). The reason is that the contamination flux is present in both the G_{BP} and G_{RP} bands, as they are integrated over the same spatial scale, which averages out in the G_{BP}−G_{RP} colour, while it will be absent from G band, which is derived from a narrow image profile-fitting, and this leads to a strong artificial reddening tail. Filtering with Eq. (2) is needed, especially in crowded area, even if G_{BP} is not used. This filter removes almost all stars fainter than $G > 16$ in Fig. 38.

5.4. Variability

The occurrence of variability along with outliers in the time-series photometry can strongly affect the mean magnitudes derived for variable sources by the *Gaia* photometric processing. In order to study this effect, we have compared the two independent estimates of the G mean magnitude provided in the *Gaia* DR2 archive for variable stars of RR Lyrae and Cepheid types, namely, `phot_g_mean_mag` listed in the

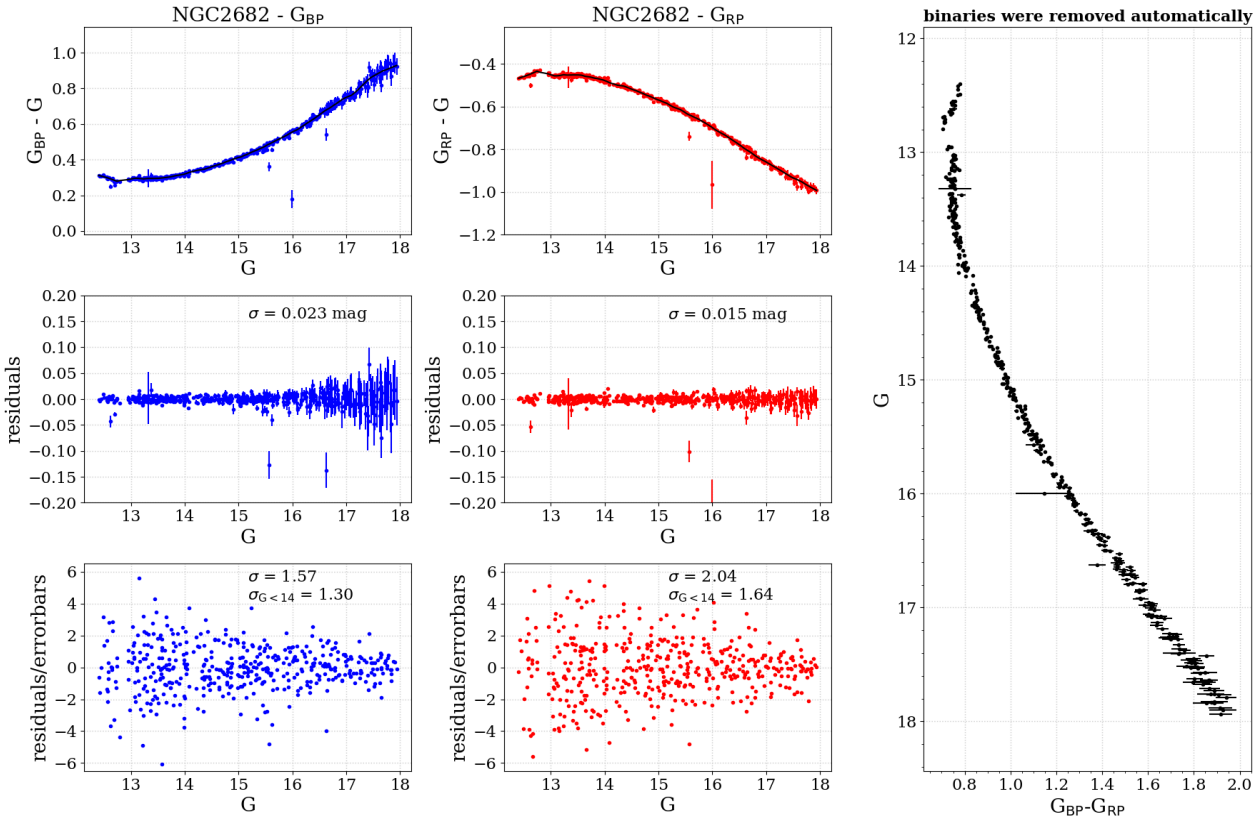

Fig. 36. CMD of NGC 2682 after the binary sequence was removed. *Left panel*: G_{BP}. *Middle panel*: G_{RP}. *Right panel*: G vs. $G_{BP}-G_{RP}$. *Top panel:* XP-G vs. G. *Middle panel*: residuals. *Bottom panel*: normalised residuals.

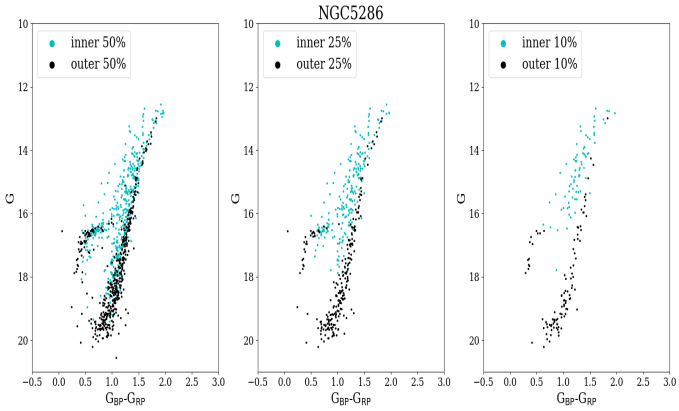

Fig. 37. CMD of globular cluster NGC 5286 inside a radius of 2.2′. The *left panel* shows inside (cyan)/outside (black) of 1.4′, the *central panel* plots the data inside 0.89′(cyan) and outside 1.75′(black), and the *right panel* shows the CMDs inside 0.55′and outside 2.0′.

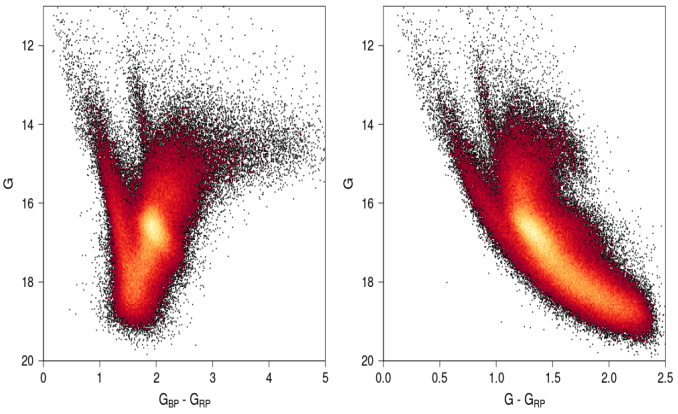

Fig. 38. Bulge CMD around the Sagittarius window (l = 1.6°, b = −2.65°) using a) $G_{BP}-G_{RP}$ or b) $G-G_{RP}$ colour, with photometric precision better than 5% for G_{BP} and/or G_{RP} and 2% for G, but without the Eq. (2) filter.

gaia_source table and `int_average_g`, which are provided for the same stars in the variability tables. The `phot_g_mean_mag` mean magnitudes are the result of the *Gaia* photometric processing, which is described in detail in Evans et al. (2018), while the `int_average_g` mean magnitudes are computed as part of the specific processing of RR Lyrae stars and Cepheids, which takes into account the variability of these sources. The `int_average_g` mean magnitudes are derived from the Fourier models that best fit the time-series data of the sources (Clementini et al. 2016). Furthermore, the outlier rejection procedures applied to estimate the `phot_g_mean_mag` and `int_average_g` mean magnitudes are different. Nevertheless, the two measurements of the mean G magnitudes are in good agreement for the large majority of stars.

For a small fraction of variables, eight RR Lyrae stars and six Cepheids, the two mean G magnitudes differ by more than 1 mag. We have visually inspected the time-series photometry of these 14 variables and found that their datasets contain faint outliers that significantly deviate from the majority of the photometric measurements. As an example, the time-series photometry of the RR Lyrae variable WY Scl and the Cepheid UY Car are presented in Fig. 39. The intensity-averaged mean G magnitude of WY Scl is `int_average_g` = 13.04 mag, while `phot_g_mean_mag` = 15.25 mag. The upper panel of

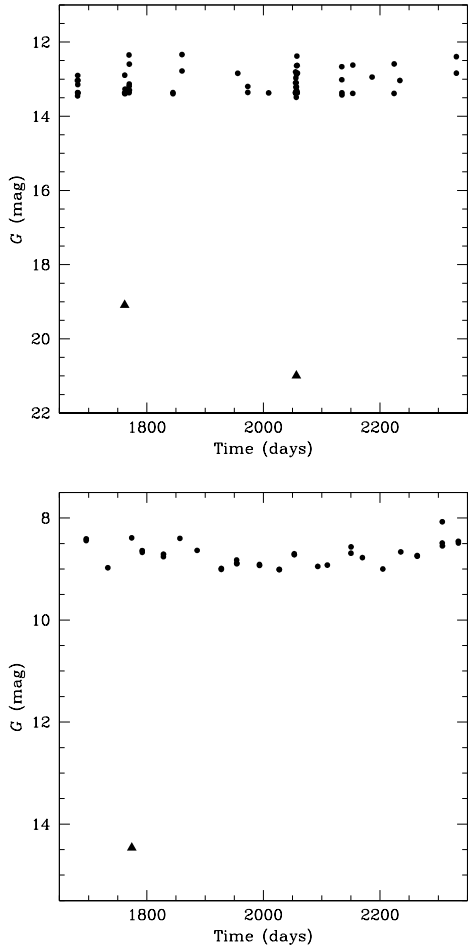

Fig. 39. Time-series G-band photometry of the RR Lyrae variable WY Scl (*upper panel*) and the Cepheid UY Car (*bottom panel*). Black dots and triangles represent the measurements we used to calculate the `int_average_g` values and rejected outliers, respectively.

Fig. 39 shows that the determination of the `phot_g_mean_mag` was affected by two faint outliers (triangles) that were instead rejected in the estimation of the `int_average_g` magnitude. The same issue also affects UY Car (bottom panel of Fig. 39), for which `phot_g_mean_mag` = 14.19 mag, while `int_average_g` = 8.69 mag.

The reason for this discrepancy is the estimation of the mean G magnitudes `phot_g_mean_mag`: the outlier rejection procedure is still not optimal for variable sources (Evans et al. 2018), while it is efficient for constant stars. The large spread of the measurements caused by the variability makes the estimation of the sigma used in the rejection difficult and overestimates its value. Hence, outliers may fall within a few sigma from the median value, and therefore they are not rejected. Furthermore, the weighted mean value is calculated in the flux space, where the fainter observations have smaller uncertainties, and consequently higher weights than brighter values. Thus, if there are faint outliers in the photometric dataset of a variable source, they will most likely drag the estimated weighted mean flux and magnitude towards the faint tail of the distribution.

This issue has affected the estimation of the mean magnitude of RR Lyrae itself, a relatively bright star ($V \sim 7.12$ mag) that has given its name to the whole class of RR Lyrae variables. The mean G magnitude of this star provided in *gaia_source* table `phot_g_mean_mag` = 17.04 ± 1.57 mag, which is ~10 mag fainter than the true value. The reason is, as in the previously described cases, a faint outlier in the G-band time series that drags the distribution towards the faint end. Unfortunately, no `int_average_g` mean magnitude is provided for the star in the variability tables: because of the limited number of measurements available during the variability processing, the star was rejected by the algorithm based on the Fourier parameters of the light curve (Clementini et al. 2018). An incorrect value of `phot_g_mean_mag` for RR Lyrae itself caused an incorrect estimation of the magnitude-dependent term applied in the astrometric instrument calibration, and consequently, an incorrect estimation of parallax for this star in *Gaia* DR2: $\varpi = -2.61 \pm 0.61$ mas.

To summarise, the method for determining the mean magnitudes `phot_g_mean_mag` of the photometric processing worked properly for the large majority of stars and produced incorrect results only for a small fraction of variables. It will be further improved in *Gaia* DR3.

6. Radial velocity

We refer to Katz et al. (2018) for a description of the RV data in *Gaia* DR2 and their extensive validation. This publication describes the stars that did not pass the quality filters to be published in *Gaia* DR2, thus affecting the completeness of the RV data. For instance, only stars with an RV uncertainty $\leq$20 km s^{-1} have been published in *Gaia* DR2. In addition, the publication has been restricted to stars with effective temperature between 3500 and 7000 K because of a degraded performance of the RV and the restricted grid of templates. A sky map of the completeness is also provided in Katz et al. (2018), showing the expected decrease in dense areas where there are conflicts between acquisition windows. The completeness also depends on the initial list of sources observed by the *Gaia* spectrograph to be published for *Gaia* DR2, whose quality was very dependent on the sky region. Katz and collaborators also report several considerations on the global zero-point (see also below), comparison with external data, and the precision and accuracy of the data as a function of magnitude, stellar properties, sky position, etc.

6.1. Accuracy

The RVs have been compared to external catalogues: GALAH DR1 (Martell et al. 2017), RAVE DR5 (Kunder et al. 2017), APOGEE DR14 (Holtzman et al. 2015), GES DR3 (Gilmore et al. 2012), SIM (Makarov & Unwin 2015), and a dedicated compilation of several smaller catalogues, UMMSV, composed of Soubiran et al. (2018), Famaey et al. (2005), Mermilliod et al. (2008, 2009), Nidever et al. (2002), Nordström et al. (2004), Worley et al. (2012), and Chubak et al. (2012). The results are summarised in Table 2. The overall differences can be due to either *Gaia* and/or the external catalogue. Similar comparisons are presented in more detail in Katz et al. (2018). A global zero-point offset of between 0.1 and 0.3 km s^{-1} is found with respect to all catalogues, including a global increase of this offset with magnitude. The other correlations are catalogue-dependent and therefore not discussed further.

The data for duplicate sources have been removed from *Gaia* DR2, but have also been used beforehand for validation purposes. Similar to previous sections, here we consider the RV related data for the duplicate sources to test for internal consistency. However, we have to bear in mind that this sample may not be representative of the entire RV dataset. We found 100 406 pairs of duplicated sources with RV data (see Sect. 2.2). The templates for each component of the pair of duplicates show that

Table 2. Summary of the comparison between RVs and external catalogues.

Catalogue	Nb	Outliers	$\langle G \rangle$	RV difference	RV uwu	Correlations
GALAH	571	1%	11.9	*0.16 ± 0.02*	*1.29 ± 0.04*	
SIM	1927	4%	9.4	*0.24 ± 0.006*	*1.12 ± 0.02*	$G_{BP}-G_{RP}$
APOGEE	60 282	2%	12.3	*0.24 ± 0.002*	*1.285 ± 0.004*	G, $G_{BP}-G_{RP}$, T_{eff}, log g
RAVE	373 755	3%	11.4	*0.27 ± 0.002*	*1.480 ± 0.002*	G, $G_{BP}-G_{RP}$, T_{eff}, log g, [Fe/H]
GES	2201	3%	12.7	*0.13 ± 0.02*	*1.33 ± 0.02*	G
UMMSV	6843	4%	7.5	*0.15 ± 0.003*	*1.38 ± 0.01*	G, $G_{BP}-G_{RP}$

Notes. The total number of stars used in the comparison (Nb), the percentage of outliers (at 5σ) excluded, and the median G of the sample are presented. The RV difference (*Gaia*-Ext, in km s^{-1}) and the unit-weight uncertainty (uwu) that need to be applied to the data to adjust the differences are indicated in *italics* when they are significant (p-value limit: 0.01). Significant correlations of the differences with other parameters are indicated in the last column and may also originate from the external catalogue.

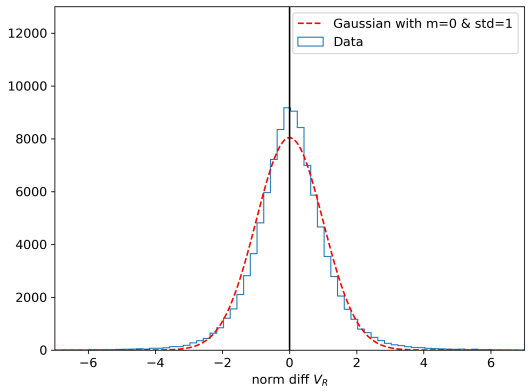

Fig. 40. Blue histogram: differences between the RV within pairs of duplicate sources, normalised by their uncertainty. The red dashed curve is a Gaussian fit forced to have mean = 0 and standard deviation = 1, but free height.

the same template has been used in 40% of the cases for log g, 86% of the cases for [Fe/H], and 41% for T_{eff}.

Overall, the RV data coming from duplicate sources are consistent. Under the assumption of Gaussian errors, the differences between the measurements of RV for the components of the pairs, when normalised by the uncertainties, should yield a Gaussian distribution centred at 0 and with dispersion equal to 1. Fig. 40 shows that the data (blue histogram) follow the expected distribution well (red curve). The normalised median and robust dispersion of the data are 0.023 ± 0.0036 (in the sense that the eliminated sources show slightly higher RVs) and 0.91, respectively.

This suggests a very small bias, again significantly below the random errors. It does not seem to correspond to differences in the templates used, and may rather originate from stars with only few observations, which is usually the case for one of the components of duplicate pairs. However, another estimation of the mean and dispersion can be obtained by fitting a Gaussian to the histogram, giving a mean of −0.08 and standard deviation of 0.87. This again indicates the overall good internal consistency of the data.

6.2. Precision of radial velocities

As indicated in the section above, the core of the distribution of the normalised differences of duplicates has a 0.87 dispersion consistent with the 0.91 value obtained using a robust estimate of the full distribution. The robustness is needed to mitigate the effect of the few binaries that must be present within the duplicates and produce outliers. This internal comparison thus points to a pessimistic estimate of the uncertainties.

Conversely, all the comparisons with external catalogues (Table 2) indicate some underestimation of the RV uncertainties. However, they are a combination of *Gaia* and external catalogue uncertainties and of the intrinsic RV variation that is due to binarity or duplicity in dense fields. The comparisons with the cleanest catalogues (e.g. with multi-epoch RV measurements) indicate the smallest underestimation, and suggest that the *Gaia* RV uncertainties are probably not significantly underestimated.

7. Astrophysical parameters

In this section, we review some of the key features we found during the validation of the AP using different approaches, namely open clusters and internal or external data. Some of these features are also reported in Andrae et al. (2018), who devoted a large part to the AP validation. In Sect. 7.1, we show the results we found for the effective temperature T_{eff}. In Sect. 7.2, we focus on the extinction A_G and reddening $E(G_{BP} - G_{RP})$, and in Sect. 7.3, we give details of the validation of the radius and luminosity. Finally, in Sect. 7.4, we use the duplicate sources in internal releases to validate the AP.

7.1. Temperature

As a first internal consistency test, we plot in Fig. 41 the comparison between the T_{eff} provided in *Gaia* DR2 and the effective temperature template used to derive the RV of the star by comparing the linear fit to the data and the 1:1 correspondence line. Taking into account that, first, the effective temperature template is largely unaffected by the extinction and should not be used as an estimate of the effective temperature of the star and, second, that the extinction could not be used when deriving the effective temperature (Andrae et al. 2018), we see that the effective temperature in *Gaia* DR2 is underestimated.

This is also apparent from Fig. 42, where we explore the relation between colour/temperature and radius as a function of Galactic latitude for a thin slice at Galactic longitude $l = 90°$. When we compare the distributions for low and intermediate latitude bins, we see for example that the cloud of points with log $R \sim 1$, which has colour $G_{BP} - G_{RP} \sim 1.2-1.4$ and effective temperature $T_{eff} \sim 5000$ K for $30° < |b| < 45°$, changes its colour because of reddening to $E(G_{BP} - G_{RP}) > 1.5$ for $|b| < 15°$. As a result, the derived temperature is artificially shifted to below 4500 K. Similarly, in this bin, the radii of the stars are too large for the lowest temperatures compared to stars at higher galactic latitudes.

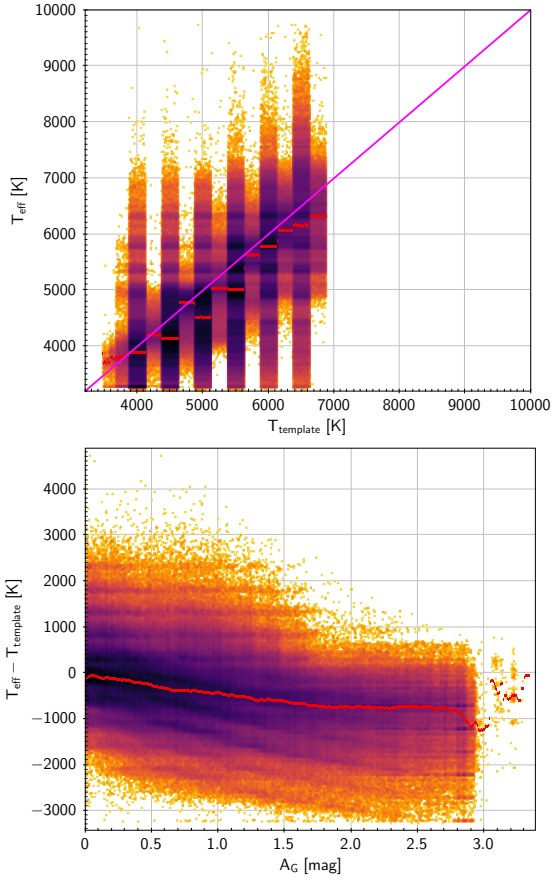

Fig. 41. *Top panel*: T_{eff} estimated from the photometry vs. T_{eff} of the RV template. The red lines show median values, and the diagonal line is the unity line. The template T_{eff} has been randomly smeared ±250 K for clarity. *Bottom panel:* difference between the two temperatures vs. extinction. The red line shows the median.

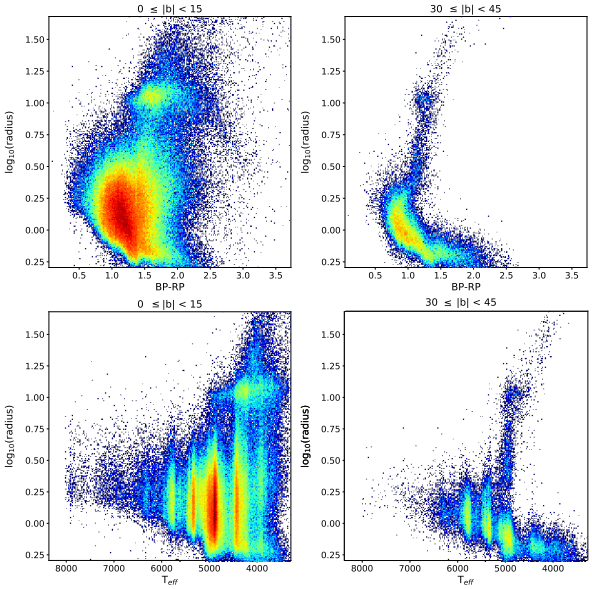

Fig. 42. Log of radius vs. colour (*top panel*) and temperature (*bottom panel*), in the Galactic plane (*left panel*) and intermediate latitude (representative of the behaviour at all latitudes $b \geq 15°$, *right panel*). The temperature of the red clump appears to follow the reddening, and cold stars on the main sequence have too large radii for $|b| \leq 15°$.

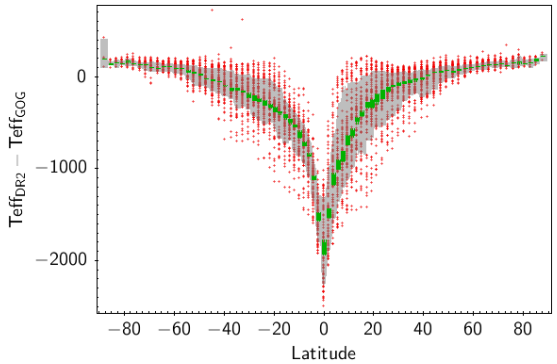

Fig. 43. Difference of the mean T_{eff} (K) between *Gaia* DR2 data and GOG simulation. Each red dot corresponds to the mean difference in a healpix bin at a given latitude (abscissa, in degrees) limited to the magnitude range $16 < G < 17$. The grey area indicates the $Q = 0.15$ to 0.85 distribution quantiles and the green one the central values ($Q = 0.46$ to 0.54).

Consistent with these findings, using GOG (Luri et al. 2014), a *Gaia*-based simulation based on the Besançon Galaxy model (Robin et al. 2012), the influence of extinction suggests that the temperature may have a large bias in the Galactic plane (Fig. 43), but it would be correctly estimated above the plane.

The same feature is observed when we compare our results to the APOGEE DR14 temperatures. The systematic offset is larger than the uncertainties at Galactic latitudes $|b| < 20°$, as presented in Fig. 12 of Andrae et al. (2018).

Then, for about 180 open clusters, we compared the value of T_{eff} with the expectations from PARSEC isochrones by Chen et al. (2014), where magnitudes are calculated with the *Gaia* passbands (see Gaia Collaboration 2018a, for details). We used literature values for the age of each cluster (Kharchenko et al. 2013) and solar metallicity for NGC 2156 and NGC 5316, and the information about NGC 2516 was taken from Jackson et al. (2016). In general, the agreement for clusters located in regions of low extinction is reasonable when the expected temperature is below $T_{eff} \sim 7000–8000$ K. In Fig. 44, we present the distribution of the temperatures for NGC 2360, which has $E(B-V) = 0.07$. This result is consistent with the fact that the temperature T_{eff} was derived under the assumption of $A_G = 0$. Significant deviations of the temperature are expected in moderate-/high-extinction regions, as in the case of NGC 5316, which has $E(B-V) = 0.29$, Fig. 45.

The *Gaia* DR2 T_{eff} was derived by training the regression algorithms with observational templates in the range 3000 K < T_{eff} < 10 000 K (Andrae et al. 2018). This has produced a saturation effect for all stars hotter or cooler than these limits. However, a significant deviation from the expected values is detected already at $T_{eff} =$ 8000 K (see e.g. Fig. 46). A spurious effect of granularity on the T_{eff} distribution is present. This is understood as coming from the inhomogeneities in the T_{eff} training data distribution (see Fig. 18 in Andrae et al. 2018).

7.2. Extinction and reddening

7.2.1. Extinction using open and globular clusters

As explained in Andrae et al. (2018), deriving the extinction and T_{eff} from the G, G_{BP}, and G_{RP} magnitudes has to face the fact that the system is degenerate and these degeneracies lead to large random errors. The consequence is that the A_G values presented in the *Gaia* DR2 catalogue cannot be easily used on a star-by-star

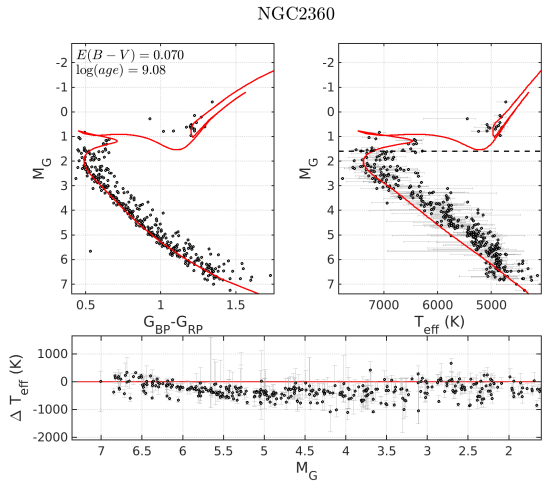

Fig. 44. *Gaia* DR2 $T_{\rm eff}$ (blue points) compared with the expectations from a PARSEC isochrone (red line) for NGC 2360 on the $M_{\rm G}$ absolute magnitude vs. $G_{\rm BP}$−$G_{\rm RP}$ plane (*upper left panel*); on the $M_{\rm G}$ vs. $T_{\rm eff}$ plane (*upper right panel*), and $\Delta\,T_{\rm eff}$ vs. $M_{\rm G}$ (*lower panel*).

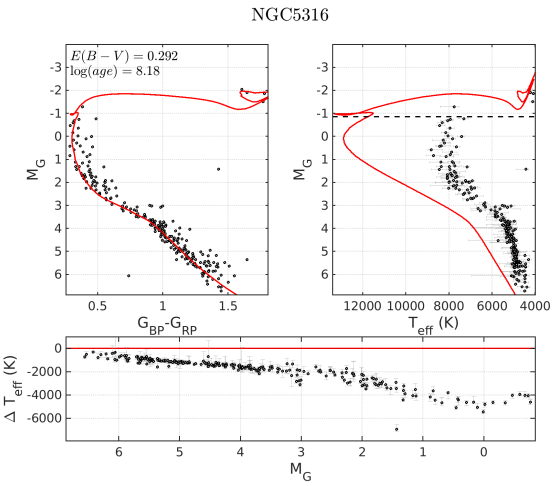

Fig. 45. Same as Fig. 44 for NGC 5316.

basis. An illustration of this is presented in Fig. 47 for the cluster IC 2602. Indeed, A_G presents a large spread inside the cluster, depending on the colour of the stars. Redder stars, either faint main-sequence stars or binaries, always have a higher extinction attributed to them. A similar analysis on globular clusters shows that the majority of stars have extinction values higher than 1. Halo globular clusters are expected to have extinctions in the range $A_G = 0.05$–0.09. It should be pointed out that extinction parameter A_0 is expected to show no variation with temperature, while A_G has a dependence on $T_{\rm eff}$. When $A_0 = 0.3$, we have $(A_G - A_O) > -0.11$ for $(G_{\rm BP} - G_{\rm RP})$ <3 (Danielski et al. 2018). However, the variations presented here cannot be entirely ascribed to this effect.

The non-negativity constraint imposed on extinction and the noise level (with typical uncertainties of about 0.6 mag on the extinction of individual sources) causes a high asymmetry in the probability distribution function for the members of a cluster. As a consequence, their uncertainty distribution is far from being Gaussian. Andrae et al. (2018) recommend to use a maximum likelihood method (ML) to combine the data and derive the most likely value inside the area, while a simple (or weighted) average overestimate the extinction. When we apply

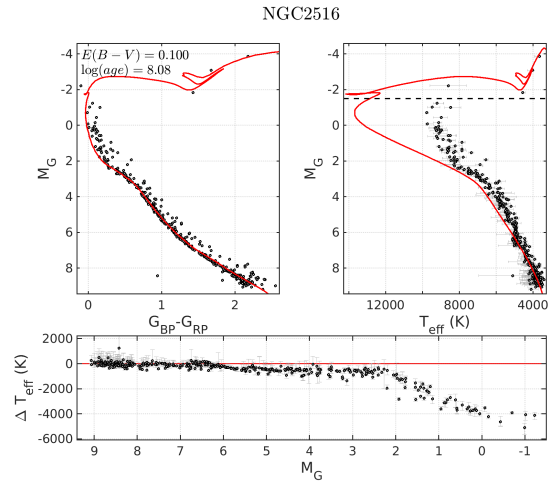

Fig. 46. Same as Fig. 44 for NGC 2516.

this method and filter the best measured stars following the flags in Andrae et al. (2018), we derive for IC 2602 the most likely value for A_G as 0.0 with a 68% probability interval in the range $0 \leq A_G < 0.24$. Applying a simple average instead gives $A_{G,\rm mean} = 0.77$. Kharchenko et al. (2013) quoted an extinction of $E(B-V) = 0.031$, corresponding to $A_G = 0.08$ (see e.g. Fig. 47).

We compared the A_G values derived using the ML method with reference literature data (Kharchenko et al. 2013) for a sample of 100 clusters, including a few halo globulars. Figures 48 and 49 present the results for the disc and the halo sub-sample, respectively.

For the disc stars, we found a general reasonable agreement with literature values, $A_{G,Gaia} - A_{G,\rm ref} = -0.01 \pm 0.02$, but with a high dispersion ($\sigma = 0.31$). In the case of metal-poor populations such as halo globulars, the agreement is less good: $A_{G,Gaia} - A_{G,\rm ref} = 0.10 \pm 0.15$ with $\sigma = 0.56$. We emphasise that comparing literature values with the arithmetic mean A_G of cluster members leads to a general overestimation of about 0.2 mag (in our sample), with differences reaching up to 0.3–0.4 mag in the case of clusters with $A_G < 0.1$ mag. The estimation method recommended in Andrae et al. (2018) is therefore, indeed preferable.

7.2.2. Internal validation of the reddening and extinction

In Fig. 50, we plot Healpix maps for the extinction in G, A_G, and effective temperature, $T_{\rm eff}$, given in *Gaia* DR2 (top and bottom panels, respectively) in Galactic coordinates and with a resolution ~0.9°. As expected, the extinction map traces the large-scale dust structure seen in the Galaxy, decreasing towards high latitudes. The temperature map shows a strong tendency towards cooler stars, where lower temperatures seem confined to the Galactic disc, the anticentre, and the Magellanic Clouds. We emphasise here that while there are certainly differences in the stellar populations at different latitudes, they are unlikely to lead to mean temperature differences as large as these.

Figure 51 shows the histogram of the ratio between the extinction in G, A_G, and the reddening, $E(G_{\rm BP} - G_{\rm RP})$. This ratio peaks around 2, as expected from Jordi et al. (2010). We note, however, a large dispersion towards higher values. Since A_G and $E(G_{\rm BP} - G_{\rm RP})$ are estimated independently of each other, if both are low, then random noise can cause $E(G_{\rm BP} - G_{\rm RP})$ to approach very close to zero such that the ratio $A_G/E(G_{\rm BP} - G_{\rm RP})$ becomes very high. Therefore, caution is necessary when using

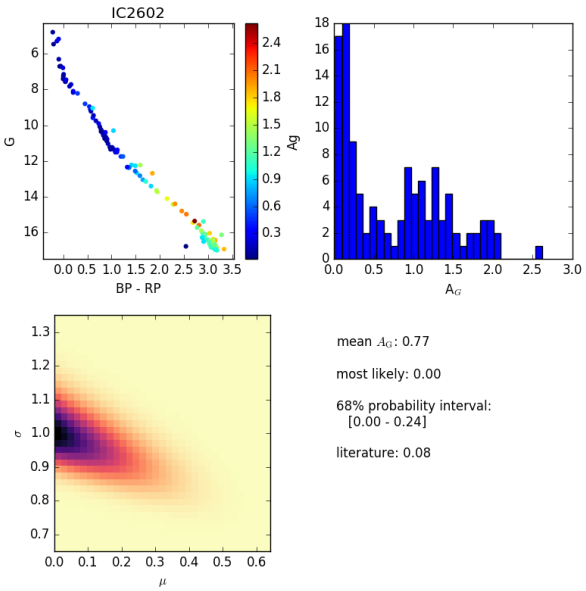

Fig. 47. CMD of IC 2602, where the colours represent the value of the extinction A_G (*top left panel*), A_G distribution in the cluster (*top right panel*), and distribution of μ, the extinction value obtained using the ML method vs. the uncertainty σ (*bottom panel*).

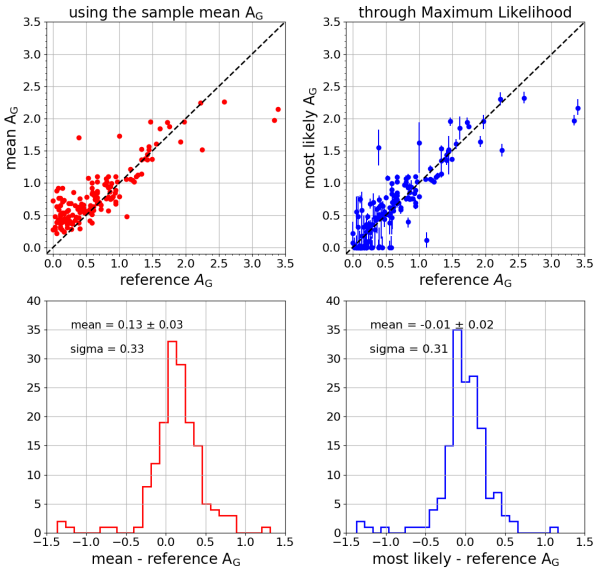

Fig. 48. Mean *Gaia* A_G values vs. literature values (*upper left panel*) and distribution of the differences (*lower left panel*) for OCs. In the *right panels*, we show the analogous plot using the recommended maximum likelihood values of A_G instead of the mean values.

this ratio, especially if it reaches unrealistically high values (i.e. $A_G/E(G_{BP} - G_{RP}) > 3$).

The ratios $A_G/E(G_{BP} - G_{RP})$ have been compared with predictions using spectral energy distributions of solar metallicity stars and the DR2 passbands (Evans et al. 2018) in Fig. 52. The predicted ratios are about 2 for temperatures higher than about 4000 K and decrease to about 1.2 at 2000–3000 K. Similar trends were present with the nominal passbands in Jordi et al. (2010). Instead, the computed ratios do not show the decrease at temperatures below 4000 K, demonstrating an issue with the extinction parameter. Fig. 7 in Andrae et al. (2018) shows that the training set does not have enough models for which $A_G/E(G_{BP} - G_{RP})$ could become ∼1.5 or lower.

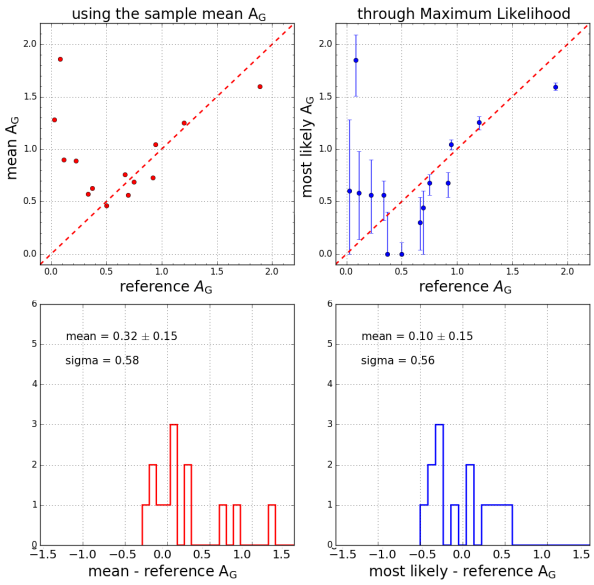

Fig. 49. Same as Fig. 48 for halo globulars.

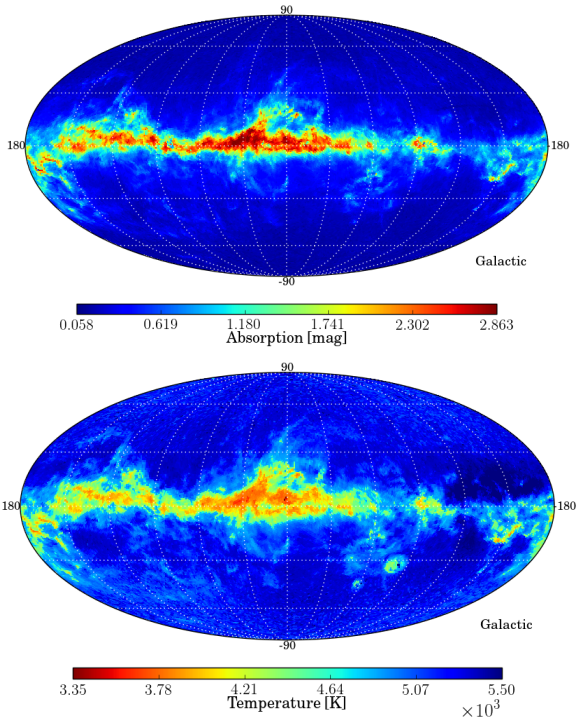

Fig. 50. Healpix maps (level 6, i.e. resolution ∼0.9°) of median extinction A_G (*top panel*) and effective temperature (*bottom panel*) in Galactic coordinates.

7.2.3. Extinction using external catalogues

The comparison to external data is complicated at the low end by the non-negativity constraint, and at the upper end by saturation due to the training grid boundaries. These effects can be seen in Fig. 53, which compares A_G to A_V determined for the APOGEE DR14 by Queiroz et al. (2018). The few outliers that remain at $A_V \sim 0$ with $A_G > 1.5$ indicate that the outlier filtration detailed in Andrae et al. (2018) is imperfect.

The uncertainties, provided as percentiles, are difficult to use on these highly skewed uncertainties. This is illustrated by a sample of low-extinction stars at high Galactic latitudes or within

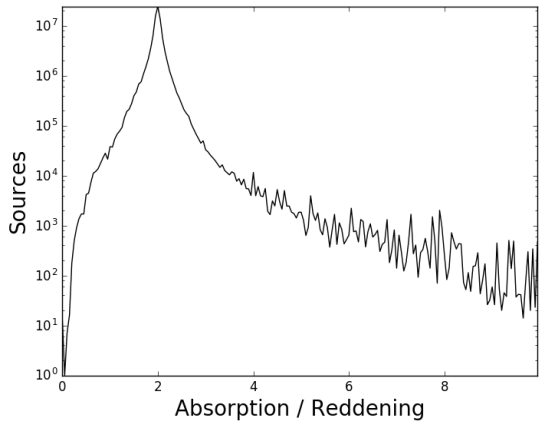

Fig. 51. Histogram of the ratio between the extinction in G, A_G, and the reddening, $E(G_{BP} - G_{RP})$.

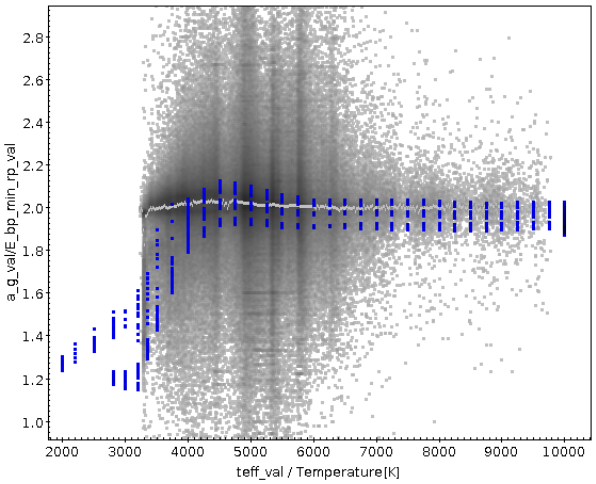

Fig. 52. Ratio $A_G/E(G_{BP} - G_{RP})$ as a function of T_{eff} as derived in DR2 (grey points). The grey line is a running median. Blue dots show the prediction using the stellar energy distribution of solar metallicity stars and the DR2 passbands (Evans et al. 2018).

the Local Bubble (within 50 pc). Not only does the estimated extinction reach high values in these samples (Andrae et al. 2018, Sect. 6.5), but the 16th percentile even exceeds 0.05 mag for 60% of the stars. This shows that the percentiles are not accurate enough to be used as estimates of the individual uncertainties; still, Andrae et al. (2018) showed that they are useful for outlier filtration.

Selecting stellar types using the extinction and colour excess requires caution. For example, when we attempt to select OB stars, the global overestimation of the extinction that is due to the non-negativity constraint moves many cool stars into the hot star colour range. Consequently, such a selection cannot be made with the *Gaia* data alone, and external photometry is required. We tested on low-extinction stars that even after applying a colour-colour cut based on 2MASS photometry, as done in Gaia Collaboration (2018d), which removes the coolest stars from an OB sample, 20 times more stars are found than expected, simply because of the large errors of the DR2 extinctions and because the hottest stars are less numerous than cooler stars (see the online documentation; Antoja et al. 2018, Sect. 10.2.7.2).

7.3. Luminosities and radii

The radii and luminosities are computed using the temperatures, with A_G set to 0.0 mag. They therefore show the same issues as described above. The radius may be recalculated from any estimate of A_G using Eq. (6) given in Andrae et al. (2018), however.

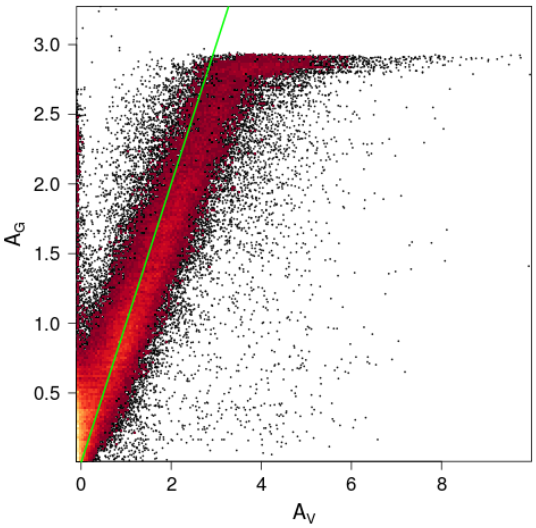

Fig. 53. Comparison between *Gaia* DR2 A_G and A_V determined on APOGEE DR14 stars by Queiroz et al. (2018). In green we show the one-to-one relation. An overdensity of stars with overestimated extinction is seen at low extinction, which is caused by the positivity constraint, and saturation at high extinction is seen as a result of the training grid boundaries.

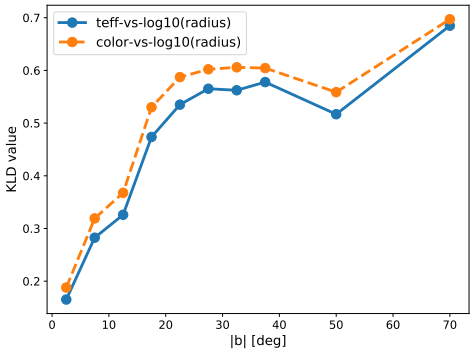

Fig. 54. Quantitative comparison of the distributions of $\log R$ vs. T_{eff} (and $G_{BP}-G_{RP}$ colour) as a function of Galactic latitude, as measured by the KLD statistic. The data are consistent with the expectation that the stellar populations are similar for all Galactic latitudes, except for $|b| \lesssim 20°$, where the behaviour is markedly different, revealing systematic issues with the determinations of radii and temperatures at low latitudes.

Given the expectation that stars at different latitudes will not vary dramatically in their intrinsic properties, we expect the 2D distributions in $\log R$ and T_{eff} to be roughly independent of Galactic latitude. The degree of similarity of the distributions can be quantified using the Kullback–Leibler divergence (KLD; Kullback & Leibler 1951; see also Arenou et al. 2017, Sect. 5.1). This is shown in Fig. 54, which clearly indicates that the distributions in $\log R$ vs. T_{eff} for $|b| \lesssim 20°$ are significantly different from those at higher $|b|$, where they resemble each other (i.e. the KLD value remains approximately constant). On the other hand, the KLD values obtained when computing the distribution of stars in the space of $\log R$ vs. $G_{BP}-G_{RP}$ vary especially strongly with latitude for low $|b|$, at least partly as expected because of reddening (see also Fig. 42).

Tests using asteroseismic targets have been presented in Andrae et al. (2018) and are not repeated here. We present in

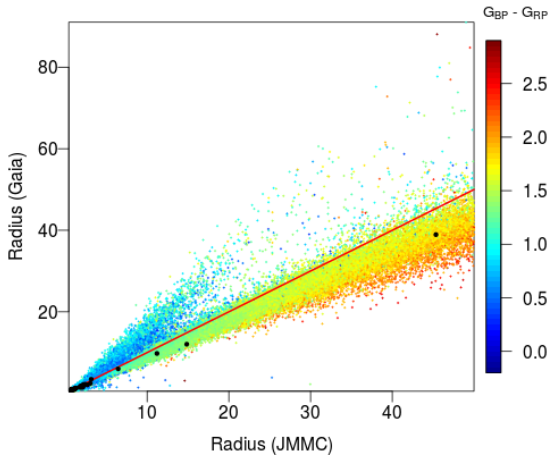

Fig. 55. Radius comparison with the JMMC stellar diameter catalogues for stars with relative parallax uncertainty smaller than 10% and $G > 6$. JSDC v2 is colour-coded with $G_{BP}-G_{RP}$ colour, JMDC in black. The red line corresponds to the one-to-one relation.

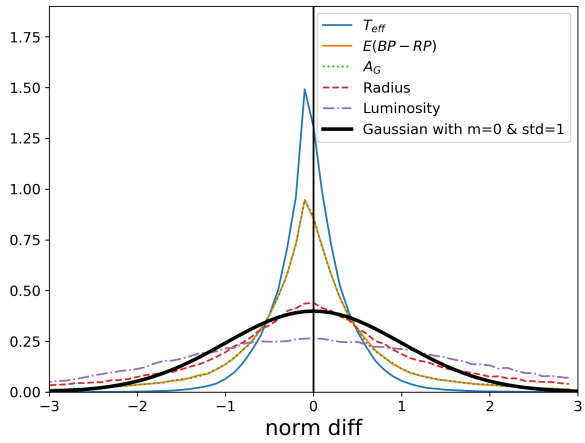

Fig. 56. Density of the normalised differences of pairs of duplicate sources, as in Fig. 40, for the effective temperature (solid blue line), extinction (dotted green line), $E(G_{BP} - G_{RP})$ (solid orange line), radius (dashed red line) and luminosity (dot-dashed purple line). The normal distribution with 0. mean (shown as the vertical black line) and dispersion 1 is shown as the thick solid black line. The curves for extinction and reddening overlap perfectly.

Fig. 55 a comparison with the JMMC stellar diameter catalogue (JSDC v2 from Bourges et al. 2017, selecting stars with $\chi^2 < 2$) and the JMMC measured stellar diameter Catalogue (JMDC; Duvert 2016) for stars with relative parallax uncertainties smaller than 10% with $G > 6$. The trail of stars for which *Gaia* overestimates the radius corresponds to hot stars, which are beyond the APSIS $T_{\rm eff}$ data training range and show strong effects from extinction and a temperature cooler than expected. A global underestimation of the radius is seen, as presented in Andrae et al. (2018). We checked that this still holds when we select only low-extinction stars. The relative underestimation increases with increasing radius.

7.4. Precision on the AP using duplicate sources

We verified as in Sect. 6 for the duplicate sources described in Sect. 2.2 whether the AP of the two components of duplicated sources are consistent. *Gaia* DR2 provides the 16th percentile and 84th percentile of the probability density function for each of the AP. We adopted as the uncertainty for each of the parameter half of the difference between the upper and lower percentiles, although it is known that the extinction errors are far from normal. Using other uncertainty estimators instead did not change the following results.

In Fig. 56, we plot the differences for duplicate sources of the five AP provided in *Gaia* DR2, normalised by their uncertainty. The data are very symmetric, but it is known that the non-negativity constraint on A_G or $E(G_{BP} - G_{RP})$, for example, causes their errors to be asymmetric. Most probably, the errors for each component of a duplicate pair are slightly correlated so that the differences of the errors between pairs can be randomly positive or negative. The normalised distribution then appears leptokurtic (as a result of the lower errors), but with a long tail (as a result of the upper errors), because the adopted uncertainty was the difference between the upper and lower percentiles; a robust width for the normalised A_G indeed gives about 0.5, but the standard deviation is close to 1, as expected. For the various APs, $T_{\rm eff}$ seem to have overestimated uncertainties, but they seem underestimated for the luminosity. For the other parameters, the uncertainties look as expected.

8. Solar system objects

8.1. Data

Gaia DR2 contains astrometry for 1 977 702 CCD observations of 14 099 solar system objects (SSOs), and as additional information, it also provides asteroid magnitudes in the G band for a selected 52% of the observations, obtained as a result of the validation process described in Gaia Collaboration (2018e).

The main goal of the validation of SSOs has been achieved, that is, to show that the asteroid astrometry is very similar to the expected performances, especially in the optimal range of brightness $G \sim 12-17$, where the typical accuracy per CCD observation is at a sub–mas level.

The validation approach to asteroid astrometry has been based on an orbit determination process that was used to assess the data quality. The orbit determination is a set of procedures to compute the orbit of an object: it uses an orbit as initial guess (well known or computed with different procedures), and then it fits an orbit on the available observations. We used the least-squares method and the differential corrections algorithm (the core of the orbit determination) to fit orbits on 22 months of *Gaia* observations, starting from the already well-known orbits of each object.

For *Gaia* DR2, we selected an initial sample of 14 124 objects, which covers all the various categories of SOSs. To assess the quality of the data, we employ in the orbit determination process a high-precision dynamical model. We added the contribution of 16 massive asteroids and Pluto, and we used a relativistic force model including the contribution of the Sun, the planets, and the Moon. While all these precautions are sufficient in the usual orbit determination process, they were not enough for the *Gaia* observations. To properly deal with *Gaia* asteroid astrometry, it is fundamental to appropriately take into account the following:

– *Gaia* astrometry is given in barycentric coordinate time (TCB).
– The error model contains the correlations in $\alpha\cos\delta$ and δ, which are strong in the epoch astrometry and crucial in the orbit determination.

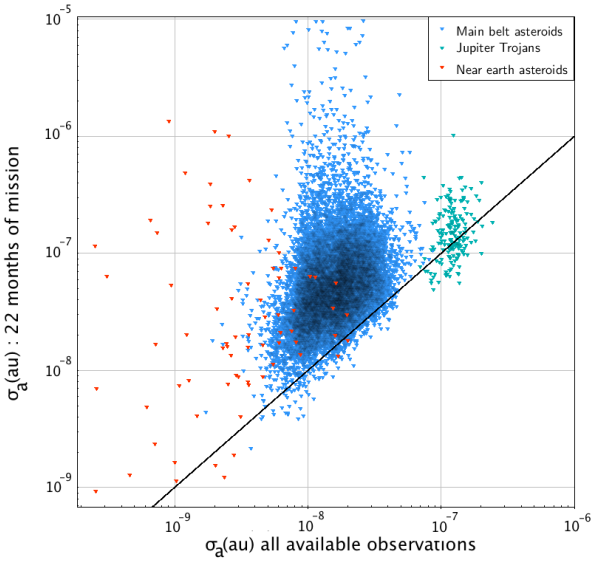

Fig. 57. Quality of the orbit determination measured by the post-fit uncertainty of the semi-major axis (au) for the solar system objects contained in *Gaia* DR2 with respect to the currently available uncertainty. The colours represent the different categories included in *Gaia* DR2: near-Earth asteroids, main-belt asteroids, and Jupiter trojans.

– The positions (α and δ) of the asteroids given in *Gaia* DR2 have been corrected using a full relativistic model, but the light deflection assumes that the object were at infinite distance. In the validation process, we also apply a further correction to take into account the finite distance.

The entire process is described in Gaia Collaboration (2018e), including the computation of the residuals on the equatorial reference frame ($\alpha \cos\delta, \delta$) and on the (AL, AC) plane, and their use for outlier rejection. As a result of this procedure, we discarded 25 objects and 1% of the observations, obtaining the sample published in *Gaia* DR2.

8.2. Orbits

The orbits are of course a secondary product of the validation process of the processing of SOSs. They represent the final outcome of the entire procedure. Since the time span covered by the *Gaia* DR2 observations is quite short (compared to the time span of the hundreds of observations available today), we expect that the quality of the orbits is limited on average. Nevertheless, as shown in Fig. 57, some asteroids in *Gaia* DR2 have already reached a quality in their orbits that is equivalent to ground-based data (and in 350 cases, the orbit is even better determined using *Gaia* observations alone).

9. Conclusions and recommendations

We have described the results of the validation tests that were applied to the second *Gaia* data release as an indirect quality control of the catalogue before its publication. With such a complex mission and the great diversity of sources, less than two years of observations and not yet optimal calibrations, the catalogue content cannot be perfect. While the overall quality of the data is excellent, the user should consider the following recommendations, depending on the specific application, in order to make optimum use of the data:

– First, in applications requiring the cleanest possible dataset, spurious solutions need to be filtered. Section 4.1 summarizes the filters suggested for this purpose.
– Because only a single-star model has been used and calibrations are sub-optimal, the quality indicators of the solutions (`astrometric_excess_noise`, `astrometric_gof_al`, `astrometric_n_bad_obs_al`) may be used to discard other potential outliers.
– In terms of astrometric systematics, correcting individual parallaxes from the global parallax zero-point (Sect. 4.3) is discouraged. For applications where the zero-point matters, however, and if the samples are well distributed over the sky, in colour, and in magnitude, parallaxes may need to be corrected for (or solved for) the global zero-point.
– In the special case of samples in small (<1°) or intermediate (<20°) regions, the contribution of systematics to the error budget has to be taken into account, see Sects. 4.2 and 4.3.
– Some substantial underestimation of the formal uncertainties has to be taken into account (Sect. 4.6), for example in sample selections that are based on astrometric precision, or for likelihood methods. Re-weighting the most precise uncertainties using Eq. (A.6) of Lindegren et al. (2018) may be useful for $13 \lesssim G \lesssim 15$ stars, or more generally, when `astrometric_n_obs_ac` = 0.
– The full covariance information between astrometric parameters should always be taken into account.

Concerning the photometry, the colours of faint stars in the neighbourhood of bright stars or contaminated in dense regions should be taken with care, and sources with large `phot_bp_rp_excess_factor` may be removed, as recommended by Evans et al. (2018) and applied in Eq. (2). For the variable stars in this release, we recommend to adopt mean magnitudes calculated by the variability processing (`int_average_g`), when they are available, over the values (`phot_g_mean_mag`) in the main catalogue.

For the AP, the extinctions cannot easily be used individually because of their large uncertainties; when an average sample value is required, it is important to follow the maximum likelihood method proposed in Andrae et al. (2018) to derive the most probable value, as a simple average would overestimate the extinction. As both A_G and $E(G_{\rm BP} - G_{\rm RP})$ depend on spectral energy distribution (Jordi et al. 2010; Danielski et al. 2018), combining them on very different spectral types would increase the uncertainties. It is preferable to use temperatures and radii in low-extinction regions to avoid biases. In moderate- to high-extinction regions, external photometry combined with *Gaia* photometry, for instance, using the cross-matches that are directly available within the *Gaia* archive (Marrese et al. 2018), may help to distinguish effects from extinction and temperature.

In summary, the wealth of data provided in *Gaia* DR2 beyond any doubt represents a landmark in the history of the astronomical catalogues. However, it was an impossible task to completely avoid mistakes and shortcomings in the astrometric, photometric, spectroscopic, or classification data in a catalogue with 1.7 billion sources with many intricate data for each, given the short time since the observations were made. Nothing is ever for free: the data cannot be used blindly, and any serious scientific exploitation of the *Gaia* data must understand and take into account the various limitations and caveats that are attached to the various *Gaia* DR2 catalogue contents.

Acknowledgements. Funding for the DPAC has been provided by national institutions, in particular the institutions participating in the *Gaia* Multilateral Agreement: the Centre National d'Etudes Spatiales (CNES), the European Space Agency in the framework of the Gaia project. This research has made an extensive use of Aladin and the SIMBAD, VizieR databases operated at the Centre de Données Astronomiques (Strasbourg) in France and of the software TOPCAT

(Taylor 2005). This work was supported by the MINECO (Spanish Ministry of Economy) through grant ESP2016-80079-C2-1-R (MINECO/FEDER, UE) and ESP2014-55996-C2-1-R (MINECO/FEDER, UE) and MDM-2014-0369 of ICCUB (Unidad de Excelencia "María de Maeztu") and the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement GENIUS FP7 – 606740. We acknowledge the computer resources, technical expertise and assistance provided by the Red Española de Supercomputación and specially the MareNostrum supercomputer at the Barcelona Supercomputing Center. AH, MB and JV acknowledge financial support from NOVA (Netherlands Research School for Astronomy), and from NWO in the form of a Vici grant.

References

Andrae, R., Fouesneau, M., Creevey, O., et al. 2018, A&A, 616, A8 (*Gaia* 2 SI)
Andrei, A. H., Souchay, J., Zacharias, N., et al. 2009, A&A, 505, 385
Antoja, T., Arenou, F., Babusiaux, C., et al. 2018, Gaia DR2 Documentation Gaia DPAC, ESA, 10
Arenou, F., Luri, X., Babusiaux, C., et al. 2017, A&A, 599, A50
Benedict, G. F. & McArthur, B. E. 2015, IAU General Assembly, 22, 2257159
Benedict, G. F., McArthur, B. E., Feast, M. W., et al. 2007, AJ, 133, 1810
Betoule, M., Marriner, J., Regnault, N., et al. 2013, A&A, 552, A124
Bourges, L., Mella, G., Lafrasse, S., et al. 2017, VizieR Online Data Catalog: II/346
Bressan, A., Marigo, P., Girardi, L., et al. 2012, MNRAS, 427, 127
Capitanio, L., Lallement, R., Vergely, J. L., Elyajouri, M., & Monreal-Ibero, A. 2017, A&A, 606, A65
Chambers, K. C., Magnier, E. A., Metcalfe, N., et al. 2016, ArXiv e-prints [arXiv:1612.05560]
Chen, Y., Girardi, L., Bressan, A., et al. 2014, MNRAS, 444, 2525
Chubak, C., Marcy, G., Fischer, D. A., et al. 2012, ArXiv e-prints [arXiv:1207.6212]
Clementini, G., Ripepi, V., Leccia, S., et al. 2016, A&A, 595, A133
Clementini, G., Ripepi, V., Molinaro, R., et al. 2018, A&A, submitted [arXiv:1805.02079] (*Gaia* 2 SI)
Cleveland, W. S. 1979, J. Am. Stat. Assoc., 74, 829
Danielski, C., Babusiaux, C., Ruiz-Dern, L., Sartoretti, P., & Arenou, F. 2018, A&A, 614, A19
de Bruijne, J. H. J., Allen, M., Azaz, S., et al. 2015, A&A, 576, A74
Dias, W. S., Monteiro, H., Caetano, T. C., et al. 2014, A&A, 564, A79 (DAML Catalogue)
Duvert, G. 2016, VizieR Online Data Catalog: II/345
Evans, D. W., Riello, M., De Angeli, F., et al. 2018, A&A, 616, A4 (*Gaia* 2 SI)
Eyer, L., Mowlavi, N., Evans, D. W., et al. 2017, A&A, submitted [arXiv:1702.03295]
Famaey, B., Jorissen, A., Luri, X., et al. 2005, A&A, 430, 165
Fey, A. L., Gordon, D., Jacobs, C. S., et al. 2015, AJ, 150, 58
Fouqué, P., Arriagada, P., Storm, J., et al. 2007, A&A, 476, 73
Gaia Collaboration (Prusti, T., et al.) 2016, A&A, 595, A1
Gaia Collaboration (Babusiaux, C., et al.) 2018a, A&A, 616, A10 (*Gaia* 2 SI)
Gaia Collaboration (Brown, A. G. A., et al.) 2018b, A&A, 616, A1 (*Gaia* 2 SI)
Gaia Collaboration (Helmi, A., et al.) 2018c, A&A, 616, A12 (*Gaia* 2 SI)
Gaia Collaboration (Katz, D., et al.) 2018d, A&A, 616, A11 (*Gaia* 2 SI)
Gaia Collaboration (Spoto, F., et al.) 2018e, A&A, 616, A13 (*Gaia* 2 SI)
Gilmore, G., Randich, S., Asplund, M., et al. 2012, Messenger, 147, 25
Hambly, N., Altmann, M., Arenou, F., et al. 2018, Gaia DR2 Documentation Gaia DPAC, ESA, 14
Henry, T. J. & Jao, W.-C. 2015, IAU General Assembly, 22, 2253773
Holl, B., Audard, M., Nienartowicz, K., et al. 2018, A&A, in press DOI: 10.1051/0004-6361/201832892
Holtzman, J. A., Shetrone, M., Johnson, J. A., et al. 2015, AJ, 150, 148
Jackson, R. J., Jeffries, R. D., Randich, S., et al. 2016, A&A, 586, A52
Jordi, C., Gebran, M., Carrasco, J. M., et al. 2010, A&A, 523, A48
Katz, D., Sartoretti, P., Cropper, M., et al. 2018, A&A, submitted [arXiv:1804.09372] (*Gaia* 2 SI)
Kharchenko, N. V., Piskunov, A. E., Schilbach, E., Röser, S., & Scholz, R.-D. 2013, A&A, 558, A53 (MWSC Catalogue)
Kullback, S. & Leibler, R. A. 1951, Ann. Math. Statist., 22, 79
Kunder, A., Kordopatis, G., Steinmetz, M., et al. 2017, AJ, 153, 75
Libralato, M., Bellini, A., Bedin, L. R., et al. 2018, ApJ, 854, 45
Lindegren, L., Lammers, U., Hobbs, D., et al. 2012, A&A, 538, A78 (the AGIS paper)
Lindegren, L., Hernandez, J., Bombrun, A., et al. 2018, A&A, 616, A2 (*Gaia* 2 SI)
Luo, A.-L., Zhao, Y.-H., Zhao, G., et al. 2015, ArXiv e-prints [arXiv:1505.01570]
Luri, X., Palmer, M., Arenou, F., et al. 2014, A&A, 566, A119
Makarov, V. V. & Unwin, S. C. 2015, MNRAS, 446, 2055
Marrese, P., Marinoni, S., Fabrizio, M., & Altavilla, G. 2018, A&A, submitted (*Gaia* 2 SI)
Martell, S. L., Sharma, S., Buder, S., et al. 2017, MNRAS, 465, 3203
Mason, B. D., Wycoff, G. L., Hartkopf, W. I., Douglass, G. G., & Worley, C. E. 2001, AJ, 122, 3466
McConnachie, A. W. 2012, AJ, 144, 4
Mermilliod, J. C., Mayor, M., & Udry, S. 2008, A&A, 485, 303
Mermilliod, J.-C., Mayor, M., & Udry, S. 2009, A&A, 498, 949
Mignard, F., Klioner, S., Lindegren, L., et al. 2018, A&A, 616, A14 (*Gaia* 2 SI)
Muraveva, T., Palmer, M., Clementini, G., et al. 2015, ApJ, 807, 127
Nidever, D. L., Marcy, G. W., Butler, R. P., Fischer, D. A., & Vogt, S. S. 2002, ApJS, 141, 503
Nordström, B., Mayor, M., Andersen, J., et al. 2004, A&A, 418, 989
Queiroz, A. B. A., Anders, F., Santiago, B. X., et al. 2018, MNRAS, 476, 2556
Reid, M. J. & Honma, M. 2014, ARA&A, 52, 339
Riello, M., De Angeli, F., Evans, D. W., et al. 2018, A&A, 616, A3 (*Gaia* 2 SI)
Robin, A. C., Luri, X., Reylé, C., et al. 2012, A&A, 543, A100
Samus', N. N., Kazarovets, E. V., Durlevich, O. V., Kireeva, N. N., & Pastukhova, E. N. 2017, Astron. Rep., 61, 80
Sarajedini, A., Bedin, L. R., Chaboyer, B., et al. 2007, AJ, 133, 1658
Sartoretti, P., Katz, D., Cropper, M., et al. 2018, A&A, 616, A6 (*Gaia* 2 SI)
Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, ApJ, 500, 525
Skrutskie, M. F., Cutri, R. M., Stiening, R., et al. 2006, AJ, 131, 1163
Smart, R. L. & Nicastro, L. 2014, A&A, 570, A87
Soubiran, C., Jasniewicz, G., Chemin, L., et al. 2018, A&A, 616, A7 (*Gaia* 2 SI)
Taylor, M. B. 2005, ASP Conf. Ser., 347, 29
Udalski, A., Szymanski, M. K., Soszynski, I., & Poleski, R. 2008, Acta Astron., 58, 69
van Leeuwen, F. 2007, A&A, 474, 653
van Leeuwen, F., de Bruijne, J. H. J., Arenou, F., et al. 2018, Gaia DR2 Documentation, Gaia DPAC, ESA
Weiler, M. 2018, A&A, in press DOI: 10.1051/0004-6361/201833462
Worley, C. C., de Laverny, P., Recio-Blanco, A., et al. 2012, A&A, 542, A48

Appendix A: *Gaia* DR2 general completeness

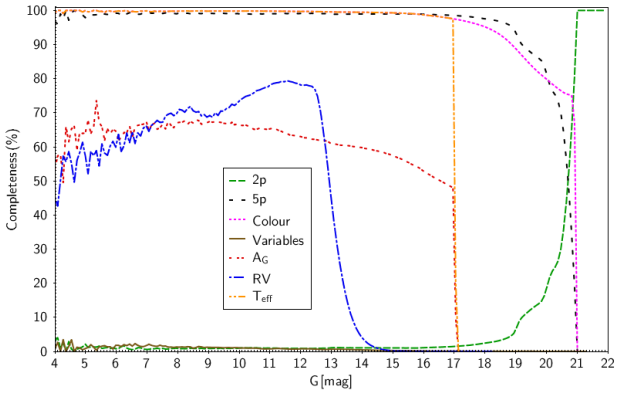

Fig. A.1. Fraction of *Gaia* DR2 sources having two (green) or five astrometric parameters (black); having colour (pink), temperature (orange), extinction (red), RV (blue); or being flagged as variable (brown).

One of the most important properties of a catalogue is its completeness. Although in this respect, *Gaia* DR2 represents a major advance compared to *Gaia* DR1 (as shown by the significant increase in the number of sources), it is nevertheless still an intermediate release, and during its processing, a variety of truncations and filters have been applied to the different types of data, which limit its completeness. Therefore, the selection function of the *Gaia* DR2 catalogue is difficult to define, and significantly depends on the type of data.

The *Gaia* DR2 truncations and filters are discussed in detail in Gaia Collaboration (2018b), but we present here a summary for convenience. An overall summary of the properties of *Gaia* DR2 can be found online[10].

Figure A.1 illustrates the completeness in astrometric, photometric and AP for the sources in *Gaia* DR2. The slightly arbitrary magnitude limits imposed in different processes are especially clear.

A.1. Payload limits

The basic limitation for the *Gaia* data is the on-board detection of sources: only sources detected in the sky mapper are tracked on the focal plane and their data stored to be sent to the ground. The on-board capabilities are described in de Bruijne et al. (2015), to which we refer as it extensively describes how the on-board processing has been optimised, and it also reports the resulting selection function for various types of objects. We may simply note that the detection algorithm has been configured to a limiting sky-mapper magnitude of $G \simeq 20.7$, thus setting the essential limit for the *Gaia* sources (see Gaia Collaboration 2016). However, at these faint magnitudes, the detection is not 100% efficient, so that only up to $G \simeq 20$ can the detection be approximately complete.

Furthermore, owing to limitations on the on-board data handling resources, the capability to observe all stars is reduced in crowded regions. In combination with the still-limited data treatment in crowded areas, this means that the survey limit in regions with densities above a few hundred thousand stars per square degree can be as bright as $G = 18$.

A.2. Full catalogue

The above described payload limits, combined with some additional restrictions introduced by the initial data treatment, define the dataset available on ground to the DPAC. This dataset is then processed through several pipelines to produce the final products in the catalogue. The minimum requirement for a source to be published is the availability of a valid position and a G magnitude. The objects with these parameters constitute the full *Gaia* DR2 catalogue, with a total of 1 692 919 135 sources. This catalogue is essentially complete between $G = 12$ and $G = 18$, but is still incomplete at the bright end, with an ill-defined faint magnitude limit, which depends on celestial position. For binaries or double stars, the current resolution limit is about 0.″4 (cf. Fig. 9). In the next sections we describe the details on the astrometric and photometric selection that led to these numbers.

A.3. Astrometry

The astrometric dataset (*Gaia* observed objects with an astrometric solution) is the result of the AGIS processing of the downloaded data (see Lindegren et al. 2018). For this dataset, the results were filtered by requiring that a source had been observed by *Gaia* at least five times (five focal plane transits), and that the astrometric excess noise and the semi-major axis of the position uncertainty ellipse are lower than 20 and 100 mas, respectively. The visibility of a source depends on the position of the sky and is tied to the *Gaia* scanning law; therefore these limitations have a complex effect on the completeness that depends on the sky region.

Even if all the published sources have at least a position, the parallax and proper motions are determined only for sources satisfying the requirement that they are brighter than $G = 21$, that the number of visibility periods used is at least 6, and that the semi-major axis of the five-dimensional uncertainty ellipse is below a magnitude-dependent threshold. Therefore, the *Gaia* DR2 subset of objects with five-parameter astrometry is significantly smaller than the full dataset and is composed of 1 331 909 727 sources. This five-parameter astrometry dataset is not complete at any magnitude, and the relative completeness (with respect to the full catalogue) varies slightly depending on the magnitude limit, as is shown in Figure 2 in Gaia Collaboration (2018b).

An additional limitation is applicable to high proper motion sources. The completeness for these objects has significantly improved with respect to *Gaia* DR1, but about 20% of stars with proper motion >0.6 arcsec yr^{-1} may still be missing.

A.4. Photometry

In addition to the selection set by the astrometric solution, sources without a well-determined value for G do not appear in *Gaia* DR2. The photometry in the G, G_{BP}, or G_{RP} bands is only reported if the source was observed at least twice by *Gaia* in the respective bands, as described in Riello et al. (2018). As in the previous case, since this is tied to the source visibility, these limitations have a complex effect on the completeness that depends on the sky region. Furthermore, due to limitations of the current photometric processing, leading to the so-called "flux-excess factor" (see again Riello et al. 2018), there is a significant fraction of the catalogue (≈300 million) with missing values of G_{BP} and/or G_{RP}.

A.5. Spectroscopy

Mean RVs are available for a subset of 7 224 631 sources, which is even now more than are currently available from ground-based observations. Objects without RVs are those fainter than $G_{RVS} = 12$ (see Sartoretti et al. 2018), as estimated from the magnitudes in the initial *Gaia* source list (IGSL; Smart & Nicastro 2014),

[10] https://www.cosmos.esa.int/web/gaia/dr2

that is, roughly corresponding to $G \simeq 13$, plus brighter objects for which some quality or conditions or limits on effective temperature were not fulfilled (see Gaia Collaboration 2018b). In addition, RV values are not listed for some sources with $|v_{\rm rad}| > 500$ km s^{-1} for which the value was clearly dubious.

As a result, the RV sample shows the distribution depicted in Fig. 1 of Gaia Collaboration (2018b), incomplete at bright magnitudes, slightly incomplete to $G \simeq 13$, and more incomplete for fainter sources. The completeness also depends on the sky position, showing traces of the distribution of the IGSL catalogue, which was used in the spectroscopic processing.

A.6. Astrophysical parameters

In *Gaia* DR2, the astrophysical parameter results are only available for sources brighter than $G = 17$, and among these only for sources for which G, $G_{\rm BP}$, and $G_{\rm RP}$ are available. Further filtering was applied based on the quality of the various inputs to the astrophysical parameter estimation, where particularly strict criteria were applied to the extinction and reddening estimations. We refer to Andrae et al. (2018) for a detailed description of the filters applied. Essentially, $T_{\rm eff}$ is available for practically all sources at $G < 17$ in the temperature range 3000–10 000 K, while estimates of the other astrophysical parameters (AP) are published for about 50% of these sources.

A.7. Variability data

During the variability analysis, a strict internal filtering was applied to the quality of the photometric time series, thus reducing the number of sources flagged as variable, followed by several additional filters to reduce the contamination due to data processing artefacts, confusion with other variables, and to remove sources for which the results of the light-curve analysis were not deemed reliable enough. For details, we refer to Holl et al. (2018), where estimations of the completeness of the global variability sample and the subsamples of different types of variables are provided.

A.8. Solar system object data

Gaia DR2 includes epoch astrometry and photometry for a pre-selected list of 14 099 known minor bodies in the solar system, primarily main belt asteroids. Thus, in this case, the objects are taken from an input list, and the filtering applied only involves the removal of some observations for which the relative flux uncertainty in the G band was larger than 0.1 (this mainly removes observations of the very "fast" objects). In addition, a selection of the observations was removed, as well as some individual sources. We refer to Gaia Collaboration (2018e) for details.

Appendix B: Open and globular cluster sample

Table B.1. Completeness level (in percentage with respect to HST fields) in various magnitude ranges in the inner and outer regions of 26 globular clusters.

Name	Region	*G* magnitudes									
		11–13	12–14	13–15	14–16	15–17	16–18	17–19	18–20	19–21	20–22
LYN07	inner	–	–	–	–	–	56	33	18	5	1
LYN07	outer	–	–	–	–	54	50	35	24	9	2
NGC0104	inner	23	2	0	0	0	0	0	0	0	0
NGC0104	outer	85	55	43	21	7	1	0	0	0	0
NGC0288	inner	–	–	–	–	–	–	60	40	15	1
NGC0288	outer	–	–	–	100	85	79	70	54	28	6
NGC1261	inner	–	–	–	77	55	37	10	1	0	0
NGC1261	outer	–	–	–	100	96	89	62	30	11	2
NGC1851	inner	–	–	40	29	14	3	0	0	0	0
NGC1851	outer	–	–	–	100	86	61	34	14	5	1
NGC2298	inner	–	–	–	–	86	67	30	10	3	0
NGC2298	outer	–	–	–	–	97	90	83	62	36	10
NGC4147	inner	–	–	–	–	–	52	29	12	3	1
NGC4147	outer	–	–	–	–	–	94	78	63	27	7
NGC5053	inner	–	–	–	–	–	–	–	78	44	13
NGC5053	outer	–	–	–	–	–	100	94	82	46	13
NGC5139	inner	–	–	1	1	0	0	0	0	0	0
NGC5139	outer	43	11	5	2	1	0	0	0	0	0
NGC5272	inner	–	–	69	44	27	6	0	0	0	0
NGC5272	outer	–	100	99	86	75	45	17	6	1	0
NGC5286	inner	–	–	52	33	13	4	0	0	0	0
NGC5286	outer	–	–	86	83	69	51	25	6	1	0
NGC5466	inner	–	–	–	–	–	–	–	69	31	7
NGC5466	outer	–	–	–	–	100	99	100	86	48	12
NGC5927	inner	–	–	–	52	37	24	2	0	0	0
NGC5927	outer	–	–	81	75	75	60	28	6	1	0
NGC5986	inner	–	–	–	59	34	14	2	0	0	0
NGC5986	outer	–	–	90	88	81	61	34	10	2	0
NGC6121	inner	–	–	–	66	54	38	20	5	0	0
NGC6121	outer	–	95	92	85	79	66	48	25	7	0
NGC6205	inner	–	–	68	42	15	1	0	0	0	0
NGC6205	outer	89	92	92	80	56	25	7	2	0	0
NGC6366	inner	–	–	–	–	–	–	69	55	32	9
NGC6366	outer	–	–	–	91	90	81	79	69	42	11
NGC6397	inner	–	–	62	49	28	11	2	0	0	0
NGC6397	outer	–	95	89	82	72	56	33	12	2	0
NGC6656	inner	–	–	53	38	11	1	0	0	0	0
NGC6656	outer	75	75	69	61	41	10	1	0	0	0
NGC6752	inner	–	62	37	18	6	1	0	0	0	0
NGC6752	outer	–	98	94	78	57	32	13	3	0	0
NGC6779	inner	–	–	–	–	63	43	16	2	0	0
NGC6779	outer	–	–	94	87	82	76	56	30	13	2
NGC6809	inner	–	–	–	–	–	48	20	4	0	0
NGC6809	outer	–	–	100	96	84	66	37	11	1	0
NGC6838	inner	–	–	–	–	–	67	54	31	8	1
NGC6838	outer	–	–	91	89	72	80	74	56	28	5
NGC7099	inner	–	–	–	47	34	11	2	0	0	0
NGC7099	outer	–	–	–	94	80	66	38	17	5	1
PAL01	inner	–	–	–	–	–	–	–	–	73	39
PAL01	outer	–	–	–	–	–	–	–	–	67	21
PAL02	inner	–	–	–	–	–	–	84	62	24	7
PAL02	outer	–	–	–	–	–	–	92	89	47	14

Notes. Inner: within 0.5 arcmin; outer: 0.5 to 2.2 arcmin.

Appendix C: Acronyms

Table C.1. List of acronyms used in the paper.

Acronym	Description
2MASS	Two-Micron All Sky Survey
AC	ACross scan (direction)
ACS	Advanced Camera for Surveys (HST)
AGIS	*Gaia* Astrometric Global Iterative Solution
AL	*Gaia* ALong scan (direction)
AP	Astrophysical Parameters
APSIS	*Gaia* Astrophysical Parameters Inference System
BP	*Gaia* Blue Photometer
CCD	Charge-Coupled Device
CMD	Colour Magnitude Diagram
DAML	New catalog of Optically Visible Open Clusters and Candidates Dias et al., 2014
DPAC	Data Processing and Analysis Consortium
DR1	*Gaia* Data Release 1
DR2	*Gaia* Data Release 2
EPSL	Ecliptic Pole Scanning Law
GC	Globular cluster
GoF	Goodness of Fit
HIP	HIPPARCOS catalogue
HPM	High Proper Motion
HST	*Hubble* Space Telescope
HealPix	Hierarchical Equal Area isoLatitude Pixelisation
IGSL	Initial *Gaia* Source List
JMMC	Jean-Marie Mariotti Center
JSDC	JMMC Stellar diameters Catalogue, Bourges et al. (2017)
JMDC	JMMC Measured Stellar diameters Catalogue, Duvert (2016)
KLD	Kullback–Leibler Divergence
LMC	Large Magellanic Cloud
MAD	Median Absolute Deviation
ML	Maximum-Likelihood method
MWSC	Milky Way Star Clusters, Kharchenko et al. (2013)
OC	Open Cluster
OGLE	Optical Gravitational Lensing Experiment
PSF	Point Spread Function
Q–Q	Quantile-quantile plot
RAVE	RAdial Velocity Experiment
RECONS	REsearch Consortium On Nearby Stars, Henry & Jao (2015)
RP	*Gaia* Red Photometer
RV	Radial Velocity
SDSS	Sloan Digital Sky Survey
SED	Spectral Energy Distribution
SMC	Small Magellanic Cloud
SOS	Specific Object Studies of the *Gaia* variability pipeline
TDSC	*Tycho* Double Star Catalogue
UMMSV	local RV Catalogue using Soubiran et al. (2018), Famaey et al. (2005), Mermilliod et al. (2008, 2009), Nidever et al. (2002), Nordström et al. (2004), Worley et al. (2012), and Chubak et al. (2012).
uwu	unit-weight uncertainty (ratio of external over internal errors)
WDS	Washington Visual Double Star Catalogue, Mason et al. (2001)