Mem. S.A.It. Vol. 90, 318

Memorie della

Binaries in Gaia perspective

DR3 teaser

D. Pourbaix

FNRS, Institut d'Astronomie et d'Astrophysique – Université Libre de Bruxelles, Bld du Triomphe, B-1050 Brussels, Belgium, e-mail: pourbaix@astro.ulb.ac.be

Abstract. The third *Gaia* data release (*Gaia* DR3) will be the first in which binary models will be considered in the processing of the positions, velocities, and fluxes. This will likely change the values available (based on the single star assumption) but will provide some additional parameters directly related to the binary nature of the source. The present status of the binary pipeline, based on a few thousand cases, is illustrated here for the very first time.

Key words. astrometry – (stars)binaries: spectroscopic – techniques:spectroscopy – (stars)binaries: eclipsing – techniques:photometry

1. Introduction

Despite the frequency of binaries in the Galaxy, the frequency of detected binaries in the final *Gaia* catalogue is anticipated to be rather low. For instance, the further the system, the smaller the astrometric signature of the companion. However even a tiny fraction of 1.6 billion objects makes a lot of systems to do science with.

Since day one of the *Gaia* Data Processing and Analysis Consortium (DPAC), binaries have had their dedicated pipeline which, so far, has not produced any published result yet. As a matter of facts, some of the plots in this paper are presented for the very first time outside DPAC. They aim at giving the reader an overview of what can be expected in the third data release (DR3), anticipated for the second half of 2021. The presentation will try to be as fair as possible, showing the goodies and warning about some limitations and potential pitfalls.

2. Motivations and limitations

In *Gaia* DR2 (*Gaia* Collaboration et al. 2018b), three distinct astrometric models have been adopted: the default 5-parameter single star (in which the colour of the object was supplied for the photometric pipeline), the 6-parameter single star plus pseudo-colour (the colour was assigned through astrometry and then the single star solution was derived based on that colour), and 2-parameter position model (for poor solutions). In the latter case ($\sim 360\times10^6$ objects), the average position was simply derived: no parallax and no proper motion. The other two models ($\sim 1.3 \times 10^9$ objects) include the parallax and proper motion and only differ on whether the colour of the object was adopted (from the photometric pipeline) or also fitted, thus meaning that, so far, nowhere the binary nature of the source has been accounted for.

A close inspection of the post DR2 Hertzsprung Russell diagram (*Gaia*

Collaboration et al. 2018a) revealed the binary nature of some sources by an increase of their luminosity: their parallax is likely accurate but there are nevertheless two sources of photons observed as a single entity by *Gaia*. These objects often lie approximately 0.75 mag above the main stellar locus, consistent with pairs of twins. The fact this secondary stellar locus can be noticed so clearly tells us right away that twins are common and their parallax is already correct despite being derived with the wrong physical model. Indeed, if their parallaxes were randomly wrong, their secondary locus would be spread out.

Obviously, not all binaries are twins and it is therefore fair to wonder how imposing the single star model affects the resulting parallax and proper motion. For instance, the Hipparcos (ESA 1997) end of mission results were sometime improved years after their original publication by adopting an alternative model (e.g., Pourbaix et al. 2003; Fekel et al. 2005). For Hipparcos, an input catalogue was used to tune the processing: some ground based results were adopted to ease the astrometric fit. However, some known binaries were nevertheless processed with the single star model without biasing their parallax and/or proper motion (e.g., Griffin 1980). On the contrary, some other binaries processed as single would see their astrometric fit substantially revised by adopting the known ground based solution but the remaining parameters would not be constrained enough to firmly accept that new solution (e.g. HIP 116360).

So, clearly, there are binaries which deserve a binary model and some others that do not. Among the first group, a range of submodels can be considered (trend, acceleration, Keplerian, non-Keplerian, etc.) depending on the amplitude, the fraction of the orbit covered by the observations, etc. The goal of the *Gaia* team is not to always supply the full orbital model but instead to provide the most robust model which fits the observations well enough. So, one cannot exclude that your favourite binary will keep being processed as single in *Gaia* DR3+ but, if so, be confident that model was the most appropriate to describe the observations.

3. Binary models in DR3

3.1. Disclaimer

The results presented in this section are based either on positions to be used for DR3 but processed with an early version of the astrometric pipeline and on radial velocities and fluxes available at the time of DR2 (i.e. the time interval covered by the observations is only 1.5 yr). These results, for any of the three types of input (positions, RV, and fluxes), are therefore preliminary and only aim at offering a feeling of what will be available eventually.

The absence of input catalogue means that the identification of the candidate binaries results from either a poor single star fit (astrometry), a variability of the radial velocity, or a specific shape of the light curve. These conditions are not mutually exclusive so whenever possible a combination of different solutions will also be considered. For the astrometry, $100,000$ objects brighter than 15th magnitude in G and a goodness of fit larger than 5 ($F2$, the Wilson & Hilferty's cube root transformation, follows a N(0, 1) distribution Wilson & Hilferty 1931; Stuart & Ord 1994) have been selected, i.e. well above the large majority of points in terms of $F2$ even if $F2$ does not meet its theoretical behaviour yet (Fig. 1). For the photometry, the regular DPAC pipeline was used to identify some eclipsing binaries. That identification was then used as a proxy to identify spectroscopic ones. In total, $6,000$ spectro-eclipsing candidates have been processed.

Even if the third data release will come at least seven years after the beginning of the nominal mission, the results will be based on $\sim 1,000$ days only. Even if the *Gaia* precision is much better than the Hipparcos one, the time coverage of *Gaia* DR3 will thus match the complete Hipparcos mission.

3.2. Astrometric binaries

An astrometric binary is an object which is seen as single but for which the successive positions are not consistent with those of a single star (Lindegren et al. 2012). Since DR2, the notion of 'seen as single' has substantially

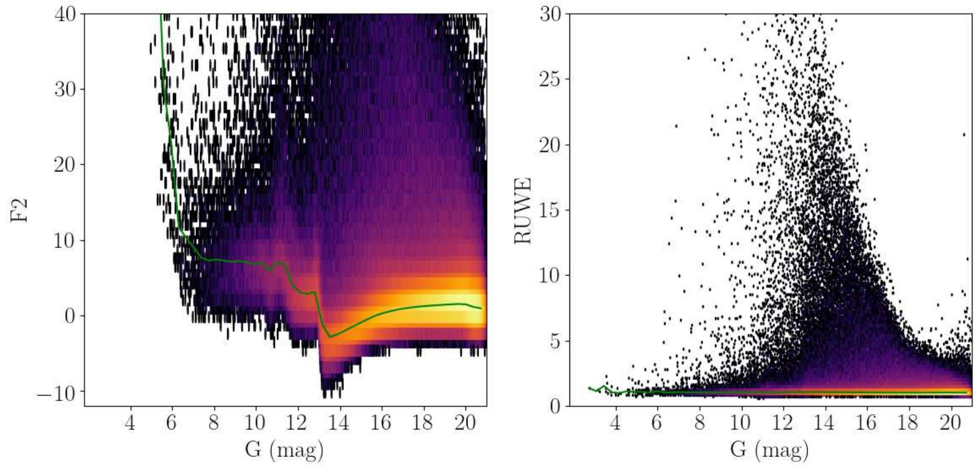

Fig. 1. Distribution of the goodness of fit (F2) and Re-normalised Unit Weight Error in terms of magnitude. The curve represents the median for a given bin of magnitude (the range of G is subdivided in 50 bins).

evolved and the point spread function of a lot of the original singles turn out to exhibit multiple peaks. It is still too early to know to what degree the individual peaks will be identified and whether each of them will behave like a single star or not. So, some poor DR2 fits will be changed either into pairs (or even more) of genuine single stars or in pairs exhibiting some reciprocal motion.

What a poor single star fit is is also difficult to assess. In DR2, several indicators are reported (excess noise, Unit Weight Error, GoodnessOfFit, . . .). A few months after DR2, yet another indicator was reported, the Renormalised Unit Weight Error, to better cope with the colour and magnitude terms biasing UWE and F2. RUWE now sits next to DR2 as an additional table in the *Gaia* Archives.

Even though RUWE is strongly peaked at 1 with a very small standard deviation, it keeps exhibiting some much larger values in the G=6–18 range. Using RUWE alone, one would wrongly conclude that the rate of binaries in the range 11–12 is almost twice as high as in the 16–17 interval. Although some binaries are likely characterised by a large RUWE, one should not use RUWE (at least the DR2 one) as a robust indicator of binarity. For DR3, the astrometric (unresolved) binary pipeline will process every point source with a poor single star fit (large RUWE). It is acknowledged that some of these objects will end up being genuine single stars but that is the price to pay not to miss many genuine binaries.

The processing itself consists in a cascade of models with increasing complexity (and thus computing time), depending on the period and variability of the binaries.. The object leaves that cascade as soon as the current model fits the observations well enough (in the χ^2 sense).

3.2.1. Acceleration models

In case of long period binaries (with respect to the time coverage of the observations), some curvature might start showing up on the trajectory. Any linear trend caused by the orbital motion has already been swallowed by the proper motion term of the single star solution. The initial curvature is modelled as the first order time derivative of the proper motion. Once the curvature increases, the second order time derivative of the proper motion is added to the model. These two acceleration models thus hold 7 and 9 parameters respectively. Even though such models give a strong indication of binarity,

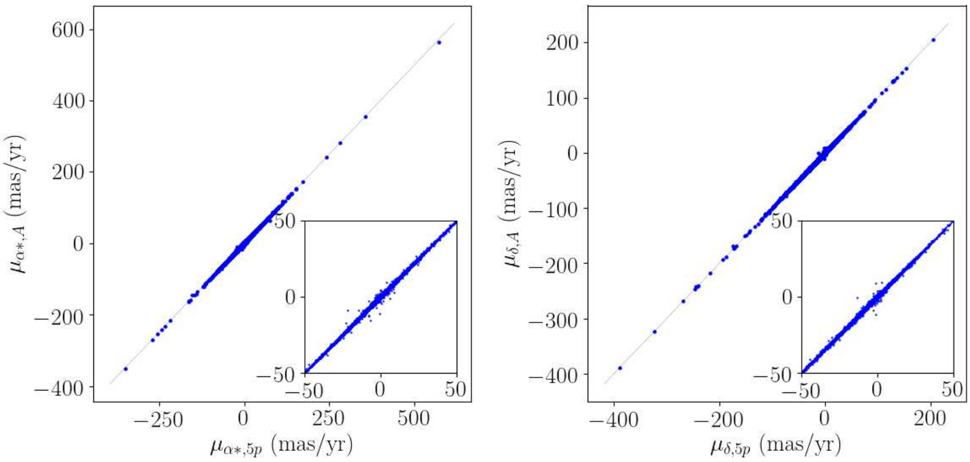

Fig. 2. Revision of the proper motion (in right ascension on the left panel, in declination on the right one) when one of the two acceleration models is adopted instead of the single star one. The x-axis represents the value associated to the single star model. The y-axis is the value of the same parameter when the acceleration model is applied.

they do not offer anything close to any orbital parameter. They can nevertheless change the proper motion of the default single star solution (Fig. 2).

3.2.2. Orbital model

Once there are enough observations (i.e. the time coverage is long enough with respect to the orbital period of the binary), the curvature terms of the previous models can no longer explain the shape of the motion (e.g. the periodicity). The full Keplerian model is then considered, bringing in seven parameters at once on top of the default five astrometric parameters.

Here again, the object leaves the cascade if the fit is good enough in the χ^2 sense. However, one further imposes that the orbital solution is significant enough, i.e. the ratio of the semi-major axis to its uncertainty is large enough. Because the *Gaia* observations are one dimensional (along a scanning direction which keeps changing due to the scanning law), that constraint on the significance of the solution filters out some solutions that are otherwise good in terms of goodness of fit. In the Hipparcos era, such poorly constrained solutions had their eccentricity artificially set to 0 (DMSA/O, ESA 1997). In order to avoid biasing the distribution of orbital parameters by imposing some circular solutions, these poorly constrained solution are discarded.

Substituting the single star model with the orbital one has some impact on the original astrometry. The parallax (ϖ) of the orbital model if often increased with respect to its value with the 5-parameter model (single star), in particular when $\varpi_{5p} < 1$ mas and the period is of the order of 400-500 days (Fig. 3). This is particularly relevant for DR3 as that release will come in two folds, more than one year apart. The first batch (eDR3) will contain the single star solution of all the sources where it is feasible but not necessary where it is appropriate. By the time eDR3 comes out, the pipeline dedicated to the binaries might not even have come to its completion. A quality indicator might nevertheless warns the users about some ϖ_{5p}. When DR3 comes out, some of these parallaxes could be substantially revised.

Very few objects end up flagged as orbital but they are located as expected in a period-eccentricity or alike diagram (Fig. 4). The first feature worth noting is the gap at 1-yr. This is

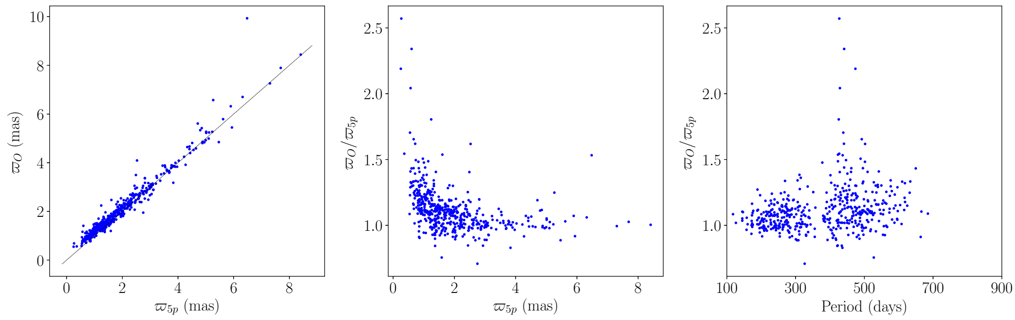

Fig. 3. Effect of adopting the orbital model on the parallax.

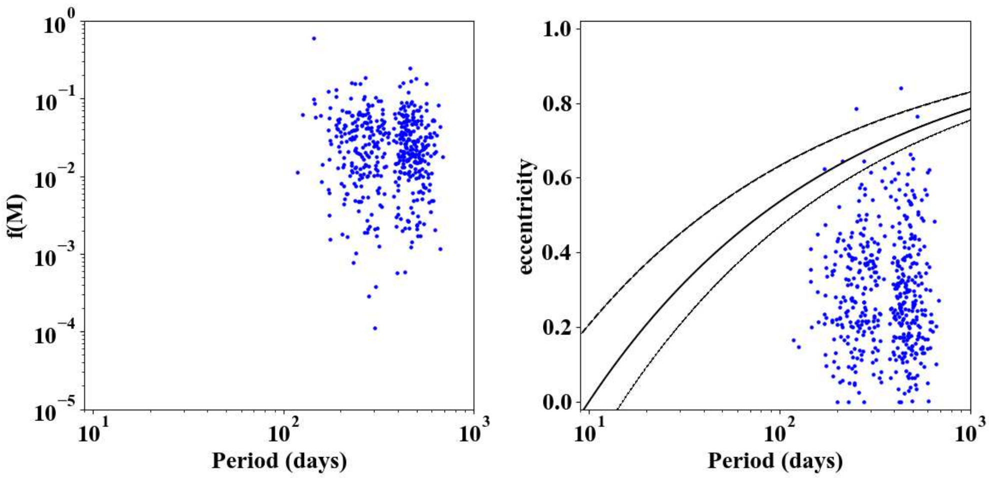

Fig. 4. Mass function and eccentricity versus orbital period for the orbital solutions obtained for a 10^5-sample of objects brighter than 15 in G and with $F2 > 5$. The three curves on the right panel represent the limit of circularisation at 5, 10, and 15 days (Halbwachs et al. 2005).

an artefact of astrometry. The parallax is associated to a 1-year periodic phenomenon. So, the parallactic term swallows any periodicity close to 1-yr, including an astrometric wobble caused by a companion as suggested by Pourbaix & Jorissen (2000) and lately confirmed by Butkevich (2018). That gap will stay throughout the whole mission.

On the left panel of Fig. 4, there are a few system with a low mass function that would often coincide with a sub-stellar companion. Whereas one cannot rule out that it is not the case here, a much more likely explanation lies in the way the mass function is derived. The astrometric orbit corresponds to the orbit of the photocentre around the barycentre of the system. When the two components are very similar, the semi-major axis vanishes and so does the mass function. So, two heavy giant stars can nevertheless have a small astrometric mass function.

The lack of systems with periods above 800 days comes from the constraint on the significance of the orbits combined with the interval covered so far by the *Gaia* observations. By the time DR3 gets out, *Gaia* will have observed for about seven years out of which only the first three have been used to derive the results contained in that data release. Whereas no long period will pop up by DR3, the absence of systems below 200 days might improve in due time.

3.2.3. Alternative astrometric models for variable stars

In case of a binary with a variable component, the photocentre moves according to both the orbital motion and the variability: the so-called Variability Induced Movers (Wielen 1996). For Hipparcos (DMSA/V, ESA 1997), VIM were considered with a limited success (Pourbaix et al. 2003) but also in a very specific situation: the two components were supposed to be in a fixed configuration so the motion was simply the combination of parallax, proper motion, and VIM. With *Gaia* more models are considered for the two (unresolved) components: fixed (same as Hipparcos), straight line, acceleration (2 models as for constant stars), and orbital. In total, up 13 parameters will thus be fitted. *Gaia* DR3 will potentially contain some of these VIM solutions. No results are reported here as the photometric data was not used in our tests.

3.3. Spectroscopic binaries

Unlike Hipparcos (ESA 1997), *Gaia* brought its own spectrometer in space (RVS: the radial velocity spectrometer). Whereas it covers the same area of the sky as the astrometric instrument, the frequency of the observations is nearly half of the astrometric ones (Fig. 5). That difference results from the optical path inside *Gaia*. The wavelength range of RVS and its resolution have been set more than a decade ago (Cropper et al. 2018), cannot be changed either, and contribute to the limitation of the number of spectroscopic binaries that can be detected.

The flux limit (about 17th magnitude) under which the RVS data are not sent to Earth is also set on board. However, the processing team further reduced that magnitude range and the interval of effective temperatures (Sartoretti et al. 2018). Despite all these limitations, RVS has already shown its capabilities in DR2 (*Gaia* Collaboration et al. 2018c).

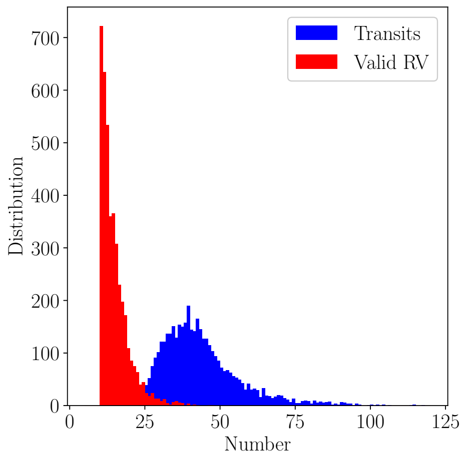

Fig. 5. Distribution of the number of observations: astrometric versus spectroscopic.

It should be remembered that the end of mission precision always assumes the object is single and its epoch velocities can be averaged. The epoch precision is however what matters for the spectroscopic binary detection and we are therefore far from the typical precision reached on ground-based velocities.

The detection of the spectroscopic binaries is based on the constancy of the radial velocity. If the radial velocity varies while the flux does not exhibit any synchronous change (other than possibly those of an eclipsing binary), the object is a good candidate for a spectroscopic binary. Clearly, the constancy detection strongly relies upon the reliability of the uncertainty of the velocity. If it is overestimated, few objects will be flagged as variable (and thus as candidate binaries). On the contrary, if it is underestimated, the number of (spurious) candidate binaries will blow up (with strong consequences on the computing resources required to process all these objects). With the DR2 pipeline, the reliability of the epoch velocity uncertainties could sometime be questioned (Fig. 6).

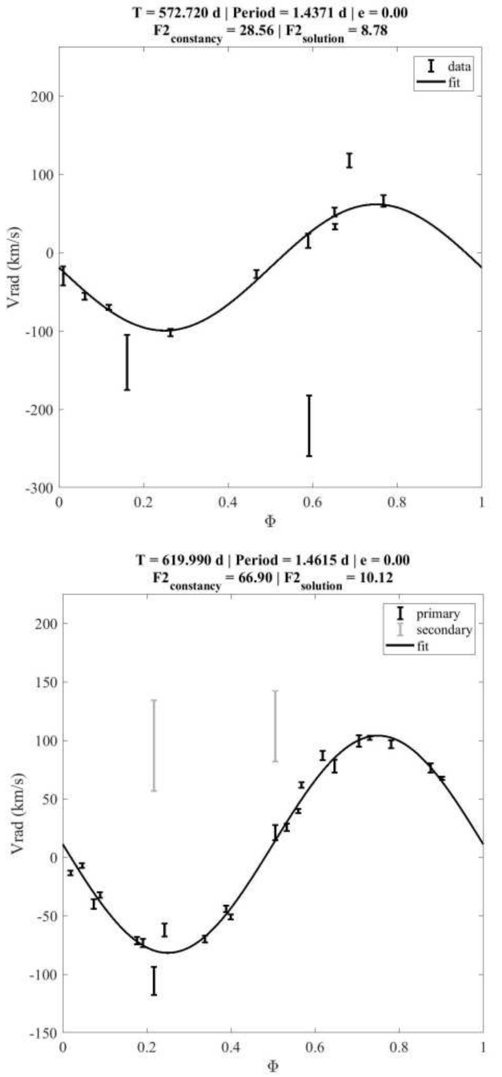

Fig. 6. Representative velocities and the corresponding orbital fits?

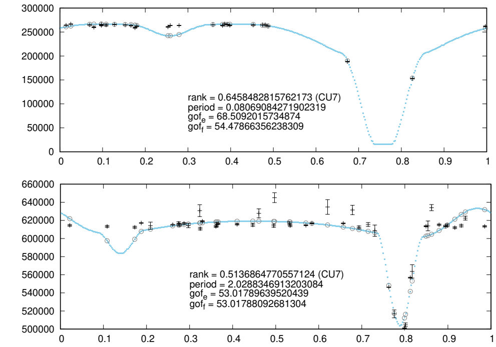

Fig. 7. Representative light curves and the corresponding eclipsing binary fits.

Once an object enters the spectroscopic binary pipeline, several models are considered depending first on how many components were detected in the RVS pipeline (SB1 vs SB2). These models range from polynomials of various degrees to the full Keplerian model (circular or ecliptic). Two such circular solutions are given for illustration purpose in Fig. 6. The bottom panel shows an double-lined binary for which only an SB1 solution was derived due to the very few velocities of the secondary available.

3.4. Eclipsing binaries

Eclipsing binaries are the most difficult to identify as not only the flux has to vary but the pattern of the variation is also very specific (Holl et al. 2018). Regardless of (the reliability of) the precision on the flux, even a genuine eclipsing binary could remain undetected due to the distribution of the observations. Some specific features (e.g. limb darkening) of an identified eclipsing system could remain unmodelled because no observation fails at that specific phase and that might be true throughout the whole mission.

Prior to DR2, about one million eclipsing binaries have been identified but have not been fully processed as such nor published (Fig 7). For DR3, despite the additional observations, the eclipsing binary identification will be carried out among that million only, with some much stronger classification criteria applied. So, DR3 will contain no more than one million eclipsing binary solutions. These orbits should hopefully fill the left part of the eccentricity-period diagram left empty by the astrometric solutions (Fig. 8)

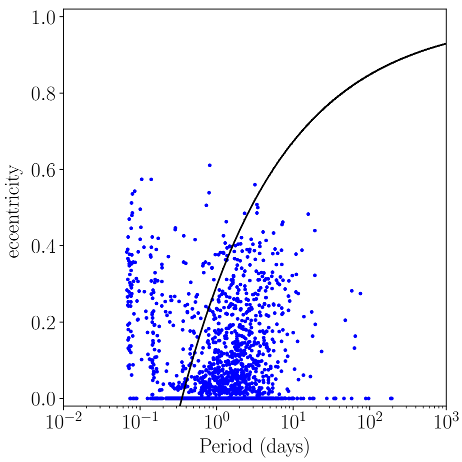

Fig. 8. Eccentricity-period diagram based on 6 000 eclipsing binaries for which radial velocities are also available. The identification is based on DR2 data with the DR2 EB classification pipeline. The curve is the envelop of the SB9 (Pourbaix et al. 2004) systems as in June 2019. The objects on the left of the curve, in particular the eccentric systems with periods shorter than 0.2d, could result from a wrong classification as eclipsing binaries in the first place.

3.5. Combined models

The three datasets (positions, velocities, and fluxes) and processing pipelines are independent, more than one could come up with a solution. If so, there is one more step in the pipeline: a combined solution. Spectro-eclipsing should be rather frequent (when velocities are available) but astro-spectro are not unlikely either (Jancart et al. 2005). When possible, a unique solution using the 2+ datasets is derived, making it more robust as it uses more observations for fewer parameters.

It is nevertheless possible that the two solutions do not match (more likely for a combination involving the astrometry). If so, the multiplicity of the object will be updated and the two solutions will be published next to each other.

4. Conclusion

Despite the preliminary nature of the inputs, it is already clear where one heads on: there will be binary results in DR3 (second half of 2021). Some filtering (not science driven) will likely be necessary prior to publication but some results are already too good to be removed.

Binaries can be seen as validators of some upstream processing: even the single star solutions can benefit from the modelling of binaries. This is true irrespective of the nature of the observations. DPAC teams are making progress everyday so DR3 will not only benefit from more observations, they will also be better calibrated, better corrected for instrumental effects, etc.: DR2- still suffer from some *growing pains*.

Despite all these progresses, there are still way too many uncertainties to make any educated guess about the number of binaries in DR3 (even a rough estimate would be presumptuous). However, just one million of eccentricities and periods would make the $(e, \log P)$ diagram more populated than anything available today.

Acknowledgements. Long term funding by the BELgian federal Science Policy Office (BELSPO) through various PROgramme de Développement d'Expériences scientifiques (PRODEX) grants is acknowledged. This work presents results from the European Space Agency (ESA) space mission *Gaia*. *Gaia* data are being processed by the *Gaia* Data Processing and Analysis Consortium (DPAC). Funding for the DPAC is provided by national institutions, in particular the institutions participating in the *Gaia* MultiLateral Agreement (MLA). The *Gaia* mission website is https://www.cosmos.esa.int/*Gaia*. The *Gaia* archive website is https://archives.esac.esa.int/Gaia.

References

Butkevich, A. G. 2018, MNRAS, 476, 5658
Cropper, M., Katz, D., Sartoretti, P., et al. 2018, A&A, 616, A5
ESA 1997, The Hipparcos and Tycho Catalogues, ESA SP-1200
Fekel, F. C., et al. 2005, AJ, 129, 1001
Gaia Collaboration, Babusiaux, C., van Leeuwen, F., et al. 2018a, A&A, 616, A10
Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2018b, A&A, 616, A1
Gaia Collaboration, Katz, D., Antoja, T., et al. 2018c, A&A, 616, A11

Griffin, R. F. 1980, Observatory, 100, 113
Halbwachs, J. L., Mayor, M., & Udry, S. 2005, A&A, 431, 1129
Holl, B., Audard, M., Nienartowicz, K., et al. 2018, A&A, 618, A30
Jancart, S., Jorissen, A., Babusiaux, C., & Pourbaix, D. 2005, A&A, 442, 365
Lindegren, L., Lammers, U., Hobbs, D., et al. 2012, A&A, 538, A78
Pourbaix, D. & Jorissen, A. 2000, A&AS, 145, 161
Pourbaix, D., Platais, I., Detournay, S., et al. 2003, A&A, 399, 1167
Pourbaix, D., Tokovinin, A. A., Batten, A. H., et al. 2004, A&A, 424, 727
Sartoretti, P., Katz, D., Cropper, M., et al. 2018, A&A, 616, A6
Stuart, A. & Ord, K. 1994, Kendall's Advanced Theory of Statistics, 6th ed., Vol. 1 (Hodder Arnold, London)
Wielen, R. 1996, A&A, 314, 679
Wilson, E. B. & Hilferty, M. M. 1931, Proceedings of the National Academy of Sciences of the United States of America, 17, 684