A&A 674, A34 (2023)
https://doi.org/10.1051/0004-6361/202243782

Astronomy & Astrophysics
Special issue

Gaia Data Release 3

Gaia Data Release 3

Stellar multiplicity, a teaser for the hidden treasure

Gaia Collaboration: F. Arenou[1,⋆], C. Babusiaux[2,1], M. A. Barstow[3], S. Faigler[4], A. Jorissen[5], P. Kervella[6], T. Mazeh[4], N. Mowlavi[7], P. Panuzzo[1], J. Sahlmann[8], S. Shahaf[9], A. Sozzetti[10], N. Bauchet[1], Y. Damerdji[12,11], P. Gavras[8], P. Giacobbe[10], E. Gosset[11,13], J.-L. Halbwachs[14], B. Holl[7,15], M. G. Lattanzi[10,16], N. Leclerc[1], T. Morel[11], D. Pourbaix[5,13,†], P. Re Fiorentin[10], G. Sadowski[5], D. Ségransan[7], C. Siopis[5], D. Teyssier[17], T. Zwitter[18], L. Planquart[5], A. G. A. Brown[19], A. Vallenari[20], T. Prusti[21], J. H. J. de Bruijne[21], M. Biermann[22], O. L. Creevey[23], C. Ducourant[24], D. W. Evans[25], L. Eyer[7], R. Guerra[26], A. Hutton[27], C. Jordi[28], S. A. Klioner[29], U. L. Lammers[26], L. Lindegren[30], X. Luri[28], F. Mignard[23], C. Panem[31], S. Randich[32], P. Sartoretti[1], C. Soubiran[24], P. Tanga[23], N. A. Walton[25], C. A. L. Bailer-Jones[33], U. Bastian[22], R. Drimmel[10], F. Jansen[34,⋆⋆], D. Katz[1], F. van Leeuwen[25], J. Bakker[26], C. Cacciari[35], J. Castañeda[36], F. De Angeli[25], C. Fabricius[28], M. Fouesneau[33], Y. Frémat[37], L. Galluccio[23], A. Guerrier[31], U. Heiter[38], E. Masana[28], R. Messineo[39], C. Nicolas[31], K. Nienartowicz[40,15], F. Pailler[31], F. Riclet[31], W. Roux[31], G. M. Seabroke[41], R. Sordo[20], F. Thévenin[23], G. Gracia-Abril[42,22], J. Portell[28], M. Altmann[22,43], R. Andrae[33], M. Audard[7,15], I. Bellas-Velidis[44], K. Benson[41], J. Berthier[45], R. Blomme[37], P. W. Burgess[25], D. Busonero[10], G. Busso[25], H. Cánovas[17], B. Carry[23], A. Cellino[10], N. Cheek[46], G. Clementini[35], M. Davidson[47], P. de Teodoro[26], M. Nuñez Campos[27], L. Delchambre[11], A. Dell'Oro[32], P. Esquej[8], J. Fernández-Hernández[48], E. Fraile[8], D. Garabato[49], P. García-Lario[26], R. Haigron[1], N. C. Hambly[47], D. L. Harrison[25,50], J. Hernández[26], D. Hestroffer[45], S. T. Hodgkin[25], K. Janßen[51], G. Jevardat de Fombelle[7], S. Jordan[22], A. Krone-Martins[52,53], A. C. Lanzafame[54,55], W. Löffler[22], O. Marchal[14], P. M. Marrese[56,57], A. Moitinho[52], K. Muinonen[58,59], P. Osborne[25], E. Pancino[32,57], T. Pauwels[37], A. Recio-Blanco[23], C. Reylé[60], M. Riello[25], L. Rimoldini[15], T. Roegiers[61], J. Rybizki[33], L. M. Sarro[62], M. Smith[41], E. Utrilla[27], M. van Leeuwen[25], U. Abbas[10], P. Ábrahám[63,64], A. Abreu Aramburu[48], C. Aerts[65,66,33], J. J. Aguado[62], M. Ajaj[1], F. Aldea-Montero[26], G. Altavilla[56,57], M. A. Álvarez[49], J. Alves[67], F. Anders[28], R. I. Anderson[68], E. Anglada Varela[48], T. Antoja[28], D. Baines[17], S. G. Baker[41], L. Balaguer-Núñez[28], E. Balbinot[69], Z. Balog[22,33], C. Barache[43], D. Barbato[7,10], M. Barros[52], S. Bartolomé[28], J.-L. Bassilana[70], U. Becciani[54], M. Bellazzini[35], A. Berihuete[71], M. Bernet[28], S. Bertone[72,73,10], L. Bianchi[74], A. Binnenfeld[75], S. Blanco-Cuaresma[76], A. Blazere[77], T. Boch[14], A. Bombrun[78], D. Bossini[79], S. Bouquillon[43,80], A. Bragaglia[35], L. Bramante[39], E. Breedt[25], A. Bressan[81], N. Brouillet[24], E. Brugaletta[54], B. Bucciarelli[10,16], A. Burlacu[82], A. G. Butkevich[10], R. Buzzi[10], E. Caffau[1], R. Cancelliere[83], T. Cantat-Gaudin[28,33], R. Carballo[84], T. Carlucci[43], M. I. Carnerero[10], J. M. Carrasco[28], L. Casamiquela[24,1], M. Castellani[56], A. Castro-Ginard[19], L. Chaoul[31], P. Charlot[24], L. Chemin[85], V. Chiaramida[39], A. Chiavassa[23], N. Chornay[25], G. Comoretto[17,86], G. Contursi[23], W. J. Cooper[87,10], T. Cornez[70], S. Cowell[25], F. Crifo[1], M. Cropper[41], M. Crosta[10,88], C. Crowley[78], C. Dafonte[49], A. Dapergolas[44], P. David[45], P. de Laverny[23], F. De Luise[89], R. De March[39], J. De Ridder[65], R. de Souza[90], A. de Torres[78], E. F. del Peloso[22], E. del Pozo[27], M. Delbo[23], A. Delgado[8], J.-B. Delisle[7], C. Demouchy[91], T. E. Dharmawardena[33], S. Diakite[92], C. Diener[25], E. Distefano[54], C. Dolding[41], H. Enke[51], C. Fabre[77], M. Fabrizio[56,57], G. Fedorets[58,93], P. Fernique[14,94], F. Figueras[28], Y. Fournier[51], C. Fouron[82], F. Fragkoudi[95,96,97], M. Gai[10], A. Garcia-Gutierrez[28], M. Garcia-Reinaldos[26], M. García-Torres[98], A. Garofalo[35], A. Gavel[38], E. Gerlach[29], R. Geyer[29], G. Gilmore[25], S. Girona[99], G. Giuffrida[56], R. Gomel[4], A. Gomez[49], J. González-Núñez[46,100], I. González-Santamaría[49], J. J. González-Vidal[28], M. Granvik[58,101], P. Guillout[14], J. Guiraud[31], R. Gutiérrez-Sánchez[17], L. P. Guy[15,102], D. Hatzidimitriou[103,44], M. Hauser[33,104], M. Haywood[1], A. Helmer[70], A. Helmi[69], M. H. Sarmiento[27], S. L. Hidalgo[105,106], T. Hilger[29], N. Hładczuk[26,107], D. Hobbs[30], G. Holland[25], H. E. Huckle[41], K. Jardine[108], G. Jasniewicz[109], A. Jean-Antoine Piccolo[31], Ó. Jiménez-Arranz[28], J. Juaristi Campillo[22], F. Julbe[28], L. Karbevska[15,110],

⋆ Corresponding author: F. Arenou, e-mail: Frederic.arenou@obspm.fr
⋆⋆ Retired from ESA.
† Deceased.

S. Khanna[69,10], G. Kordopatis[23], A. J. Korn[38], Á. Kóspál[63,33,64], Z. Kostrzewa-Rutkowska[19,111], K. Kruszyńska[112], M. Kun[63], P. Laizeau[113], S. Lambert[43], A. F. Lanza[54], Y. Lasne[70], J.-F. Le Campion[24], Y. Lebreton[6,114], T. Lebzelter[67], S. Leccia[115], I. Lecoeur-Taibi[15], S. Liao[116,10,117], E. L. Licata[10], H. E. P. Lindstrøm[10,118,119], T. A. Lister[120], E. Livanou[103], A. Lobel[37], A. Lorca[27], C. Loup[14], P. Madrero Pardo[28], A. Magdaleno Romeo[82], S. Managau[70], R. G. Mann[47], M. Manteiga[121], J. M. Marchant[122], M. Marconi[115], J. Marcos[17], M. M. S. Marcos Santos[46], D. Marín Pina[28], S. Marinoni[56,57], F. Marocco[123], D. J. Marshall[124], L. Martin Polo[46], J. M. Martín-Fleitas[27], G. Marton[63], N. Mary[70], A. Masip[28], D. Massari[35], A. Mastrobuono-Battisti[1], P. J. McMillan[30], S. Messina[54], D. Michalik[21], N. R. Millar[25], A. Mints[51], D. Molina[28], R. Molinaro[115], L. Molnár[63,125,64], G. Monari[14], M. Monguió[28], P. Montegriffo[35], A. Montero[27], R. Mor[28], A. Mora[27], R. Morbidelli[10], D. Morris[47], T. Muraveva[35], C. P. Murphy[26], I. Musella[115], Z. Nagy[63], L. Noval[70], F. Ocaña[17,126], A. Ogden[25], C. Ordenovic[23], J. O. Osinde[8], C. Pagani[3], I. Pagano[54], L. Palaversa[127,25], P. A. Palicio[23], L. Pallas-Quintela[49], A. Panahi[4], S. Payne-Wardenaar[22], X. Peñalosa Esteller[28], A. Penttilä[58], B. Pichon[23], A. M. Piersimoni[89], F.-X. Pineau[14], E. Plachy[63,125,64], G. Plum[1], E. Poggio[23,10], A. Prša[128], L. Pulone[56], E. Racero[46,126], S. Ragaini[35], M. Rainer[32,129], C. M. Raiteri[10], P. Ramos[28,14], M. Ramos-Lerate[17], S. Regibo[65], P. J. Richards[130], C. Rios Diaz[8], V. Ripepi[115], A. Riva[10], H.-W. Rix[33], G. Rixon[25], N. Robichon[1], A. C. Robin[60], C. Robin[70], M. Roelens[7], H. R. O. Rogues[91], L. Rohrbasser[15], M. Romero-Gómez[28], N. Rowell[47], F. Royer[1], D. Ruz Mieres[25], K. A. Rybicki[112], A. Sáez Núñez[28], A. Sagristà Sellés[22], E. Salguero[48], N. Samaras[37,131], V. Sanchez Gimenez[28], N. Sanna[32], R. Santoveña[49], M. Sarasso[10], M. Schultheis[23], E. Sciacca[54], M. Segol[91], J. C. Segovia[46], D. Semeux[77], H. I. Siddiqui[132], A. Siebert[14,94], L. Siltala[58], A. Silvelo[49], E. Slezak[23], I. Slezak[23], R. L. Smart[10], O. N. Snaith[1], E. Solano[133], F. Solitro[39], D. Souami[6,134], J. Souchay[43], A. Spagna[10], L. Spina[20], F. Spoto[76], I. A. Steele[122], H. Steidelmüller[29], C. A. Stephenson[17,135], M. Süveges[136], J. Surdej[11,137], L. Szabados[63], E. Szegedi-Elek[63], F. Taris[43], M. B. Taylor[138], R. Teixeira[90], L. Tolomei[39], N. Tonello[99], F. Torra[36], J. Torra[28,†], G. Torralba Elipe[49], M. Trabucchi[139,7], A. T. Tsounis[140], C. Turon[1], A. Ulla[141], N. Unger[7], M. V. Vaillant[70], E. van Dillen[91], W. van Reeven[142], O. Vanel[1], A. Vecchiato[10], Y. Viala[1], D. Vicente[99], S. Voutsinas[47], M. Weiler[28], T. Wevers[25,143], Ł. Wyrzykowski[112], A. Yoldas[25], P. Yvard[91], H. Zhao[23], J. Zorec[144], S. Zucker[75]

(Affiliations can be found after the references)

Received 13 April 2022 / Accepted 9 June 2022

ABSTRACT

Context. The *Gaia* DR3 catalogue contains, for the first time, about 800 000 solutions with either orbital elements or trend parameters for astrometric, spectroscopic, and eclipsing binaries, and combinations of these three.
Aims. With this paper, we aim to illustrate the huge potential of this large non-single-star catalogue.
Methods. Using the orbital solutions and models of the binaries, we have built a catalogue of tens of thousands of stellar masses or lower limits thereof, some with consistent flux ratios. Properties concerning the completeness of the binary catalogues are discussed, statistical features of the orbital elements are explained, and a comparison with other catalogues is performed.
Results. Illustrative applications are proposed for binaries across the Hertzsprung-Russell Diagram (HRD). Binarity is studied in the giant branch and a search for genuine spectroscopic binaries among long-period variables is performed. The discovery of new EL CVn systems illustrates the potential of combining variability and binarity catalogues. Potential compact object companions are presented, mainly white dwarf companions or double degenerates, but one candidate neutron star is also found. Towards the bottom of the main sequence, the orbits of previously suspected binary ultracool dwarfs are determined and new candidate binaries are discovered. The long awaited contribution of *Gaia* to the analysis of the substellar regime shows the brown dwarf desert around solar-type stars using true rather than minimum masses, and provides new important constraints on the occurrence rates of substellar companions to M dwarfs. Several dozen new exoplanets are proposed, including two with validated orbital solutions and one super-Jupiter orbiting a white dwarf, all being candidates requiring confirmation. Besides binarity, higher order multiple systems are also found.
Conclusions. By increasing the number of known binary orbits by more than one order of magnitude, *Gaia* DR3 will provide a rich reservoir of dynamical masses and an important contribution to the analysis of stellar multiplicity.

Key words. binaries: general – astrometry – planetary systems – stars: fundamental parameters – catalogs – white dwarfs

1. Introduction

The success of *Gaia* (Gaia Collaboration 2016b), with parallaxes for around 1.5 billion sources, could overshadow the difficulties faced in measuring the first stellar distances. The two millennia during which this research was unsuccessfully carried out were littered with unrelated but equally fundamental discoveries. In particular, *Herschel*, following the suggestion by Ramponi to Galileo in 1611 (Siebert 2005), observed pairs of stars in order to measure their differential parallaxes, but did not succeed. Instead, what he demonstrated for the first time, in 1802, was the existence of orbits for these stars, changing their nature from unrelated double stars to binaries, proving that the law of gravitation was universal.

After Bessel obtained the first convincing parallax measurement in 1838, he also deduced in 1844, from the non-linear proper motion of Sirius and Procyon, that there could exist not only visual binaries but also invisible binaries, nowadays referred to as astrometric binaries. Astrometry and binarity have

therefore been intimately linked from the start. Indeed, it was not until much later, by observing the periodic Doppler shift of Algol's lines, that Vogel correctly deduced in 1889 that this latter was due to its orbital motion, making Algol the first spectroscopic binary. Furthermore, the periodic eclipse of Algol was hypothesised by John Goodricke in 1782, making this star also the first eclipsing binary (Leverington 1995).

Since then, binary stars have been found to be important for deriving the physical properties of stars but also for their fundamental role in stellar evolution; understanding the statistical properties of binary and multiple stars is therefore of utmost importance for developing our knowledge of the Galaxy. The properties of companions down to the substellar regime are also important for understanding stellar formation. Unfortunately, until now, small sample sizes, selection effects, and the absence of the required astrometric precision have complicated the analysis of the various existing ground-based data.

As a large survey, *Gaia* should be in an ideal place to bring a new and much broader perspective to these fundamental topics. What makes *Gaia* so unique is its ability to find, and above all, to parameterise most types of binaries simultaneously, whether they be visual, astrometric, spectroscopic, or eclipsing, and even through stellar parametrisation, with a remarkable homogeneity of epoch, level of calibration accuracy, data reduction, and process organisation.

The *Gaia* precursor, HIPPARCOS, had already discovered and measured double stars (Lindegren 1997), mostly resolved ones but also several categories of unresolved astrometric binaries, which allowed stellar masses to be determined (Söderhjelm et al. 1997; Martin et al. 1997) albeit for only a small number of sources.

With the successive *Gaia* DR1 (Gaia Collaboration 2016a), DR2 (Gaia Collaboration 2018b), and then EDR3 (Gaia Collaboration 2021a), multiple stars were still not handled, with analysing single stars already posing a significant challenge, and these successive releases represent the improvement of calibrations and source analysis. This does not mean that non-single stars were absent. Whether double or binaries, they are indeed present and processing them as single stars seriously degrades their results. Nevertheless, several flags in the *Gaia* catalogue inform us about the potential duplicity, and the combination of these first *Gaia* releases with HIPPARCOS data already allowed the community to detect long-period binaries (Kervella et al. 2019a, 2022; Brandt 2021).

The advent of *Gaia* DR3 (Gaia Collaboration 2023b) now presents impressive new data products among which, to quote only a few, variability (Eyer et al. 2023), radial velocities (RV, Katz et al. 2023), and astrophysical parameters (Creevey et al. 2023) determined using either high-resolution (RVS) or low-resolution data (*BP*–*RP* photometers, De Angeli et al. 2023) for a very large fraction of the catalogue. *Gaia* DR3 also contains the first analysis of the unresolved binary star contents, covering the typical binary classes (astrometric, spectroscopic, photometric) and presented in several tables: two-body orbits, astrometric or spectroscopic accelerations, and variable binaries. These tables contain the orbital or trend parameters of the binaries that have been discovered. Above all, this offers the prospect of deriving the physical properties of the individual components. This should also improve the measurements of these systems in the main catalogue, with better astrometric parameters or systemic radial velocity.

Although the maturity of the analysis of *Gaia* data now makes it possible to obtain, for the first time, a multi-type catalogue of binaries that is much larger than has been compiled – with difficulty – over the previous centuries, it must be stressed that only a small fraction of the binary content of the main catalogue has been analysed for DR3. This data analysis is described in the documentation (Pourbaix et al. 2022)[1] and the articles accompanying this data release, namely Halbwachs et al. (2023), Holl et al. (2023b), Gosset et al. (in prep.), Mowlavi et al. (2023) and Siopis (in prep.).

The purposes of this publication are manifold. It is first intended to describe the possible use cases of the catalogue, illustrating in particular the potential complementarity of the different data processing chains. This is essentially an appetiser that shows the quality of the data, highlighting the basic results that can be readily obtained, in particular estimating masses which were not part of the non-single star tables. In addition, this performance verification paper acts as a final validation step before releasing the data. It is beyond the scope of this publication to explore the data in detail, and we do not intend to compare them with models or to confirm candidates of various kinds, as these will be the goals of scientific exploitation by the community, but we wish to facilitate this exploitation by describing what has been learnt from our analysis of the data so far.

We start by describing the data content. Useful statistical properties are then clarified together with what is known about the selection function. We then focus on orbits, and not on acceleration solutions (for astrometry) or trend solutions (for spectroscopic binaries), and instead propose a catalogue of masses for these orbital solutions. From this, we present an impressionistic panorama of the potential of this catalogue concerning basic statistical properties and candidate sources of various types; for example EL CVn, compact companions, white dwarfs and high-mass dwarfs, and then ultracool dwarfs and substellar companions. Finally, multiple systems are discussed.

2. Data description

2.1. Table contents

The non-single star (NSS) tables are presented by type of solution or period range rather than according to binary type. The first of the four tables, `nss_two_body_orbit`, contains the orbital parameters for all three categories, that is, astrometric, spectroscopic, and eclipsing binary, all being unresolved. The table `nss_acceleration_astro` contains accelerations or derivative of this parameter for sources that have an astrometric motion better described using a quadratic or cubic rather than a linear proper motion. Similarly, the `nss_non_linearspectro` are trend (long-period) solutions of spectroscopic binaries. The solutions in the `nss_vim_fl` table are different in that the photocentre displacement due to the photometric variability of one component of fixed binaries required the correction of the astrometric parameters (variable-induced movers fixed, `VIMF`). A summary of the solutions is given in Table 1.

The astrometric orbits in the `nss_two_body_orbit` table have a `nss_solution_type` name starting with `Orbital*` and the orbital parameters are described in Appendix B.1. The spectroscopic binaries with either one component being parametrised (`SB1`) or both (`SB2`) have their parameters described in Appendix B.2 and short periods may have a circular solution (`nss_solution_type` = `SB1C`). As a source may simultaneously be, for example, an astrometric binary and a spectroscopic binary, combined solutions have been computed in some cases (`nss_solution_type` = `AstroSpectroSB1`). For the same reason, the `EclipsingSpectro` solutions are combinations of eclipsing and spectroscopic solutions. However, when no combination has been performed, then two solutions for the same source may be present in the `nss_two_body_orbit` table;

[1] https://gea.esac.esa.int/archive/documentation/GDR3/index.html

Table 1. Content of the four non-single star tables.

Table	nss_solution_type	Solutions	Description
nss_acceleration_astro	Acceleration7	246 947	Second derivatives of position (acceleration)
	Acceleration9	91 268	Third derivatives of position (jerk)
nss_two_body_orbit	Orbital	134 598	Orbital astrometric solutions
	OrbitalAlternative*	629	Orbital astrometric, alternative solutions
	OrbitalTargetedSearch*	533	Orbital astrometric, supplementary external input list
	AstroSpectroSB1	33 467	Combined orbital astrometric + spectroscopic solutions
	SB1 or SB2	186 905	Orbital spectroscopic solutions
	EclipsingSpectro	155	Combined orbital spectroscopic + eclipsing solutions
	EclipsingBinary	86 918	Orbits of eclipsing binaries
nss_non_linearspectro	FirstDegreeTrendSB1	24 083	First order derivatives of the radial velocity
	SecondDegreeTrendSB1	32 725	Second order derivatives of the radial velocity
nss_vim_fl	VIMF	870	Variable-induced movers fixed

Notes. The number of solutions is larger than the number of sources. The type of solution OrbitalAlternative* indicates solutions that are either OrbitalAlternative or OrbitalAlternativeValidated.

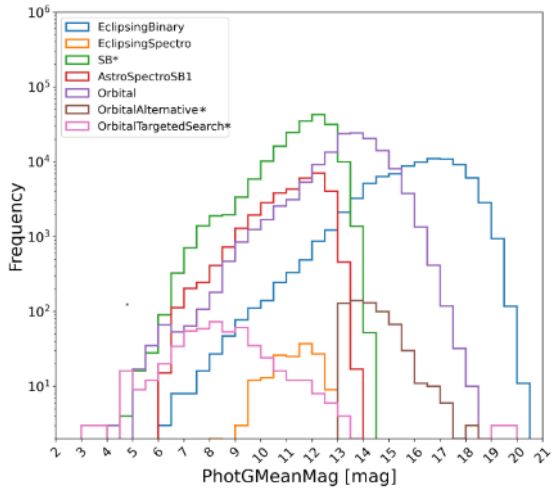

Fig. 1. Magnitude distribution for each solution type in the nss_two_body_orbit table.

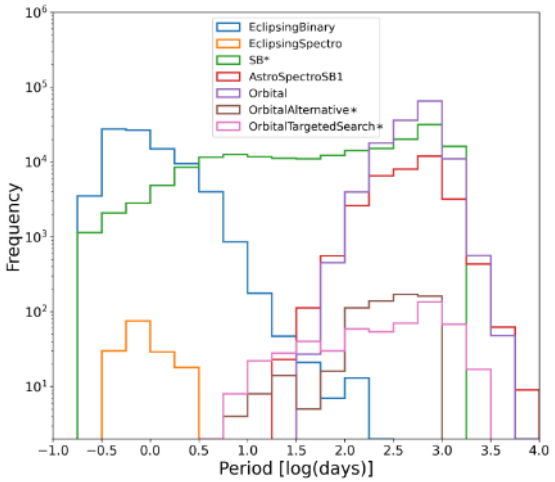

Fig. 2. Number of solutions for each solution type in the nss_two_body_orbit table as a function of period.

that is, a query by source_id may return several solutions. These multiple solutions may indicate either triple systems, or some inconsistency that users may wish to sort out, and then possibly combine these solutions offline.

For the same reason, some sources may also have solutions in several tables simultaneously. To take an example, there are 160 eclipsing binaries that also have a VIMF solution. As the VIMF model should have improved their astrometric solution, and the distance of eclipsing binaries is of interest, this solution should in principle be preferred over the one given in the gaia_source table. This potential multiplicity of solutions for a given source explains why the total number of unique NSS sources is 813 687 while the total number of NSS solutions is larger, 839 098.

The distributions of the various orbital solutions with magnitude are shown Fig. 1. As expected, the brightest are the SB1 and SB2, and consequently also their intersection with astrometric binaries, AstroSpectroSB1, and with eclipsing binaries, EclipsingSpectro. The orbital astrometric binaries (brighter than $G < 19$) peak at $G \approx 14$ while the OrbitalTargetedSearch span the entire magnitude range as the sources were given as input list. The eclipsing binaries are the faintest. We note that the NSS eclipsing binaries are a small subset of the ones detected by photometry (Mowlavi et al. 2023), for which an orbital solution has been computed (Siopis, in prep.); we refer to the much more complete vari_eclipsing_binary table for comparison.

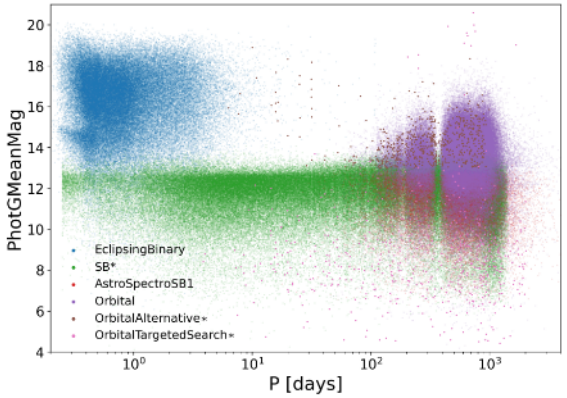

Fig. 3. G apparent magnitude vs. period in the nss_two_body_orbit table.

The distribution of periods, by construction restricted to the nss_two_body_orbit table, is depicted in Fig. 2. The short-period eclipsing and long-period astrometric binaries are nicely bridged by the SBs. Within a few years, *Gaia* has covered the impressive 0.28 – 1500-day period range (99% CI) for thousands of sources, which should prove very valuable for the statistics of binary properties. The coverage in the joint distribution of period and magnitude is qualitatively illustrated in Fig. 3.

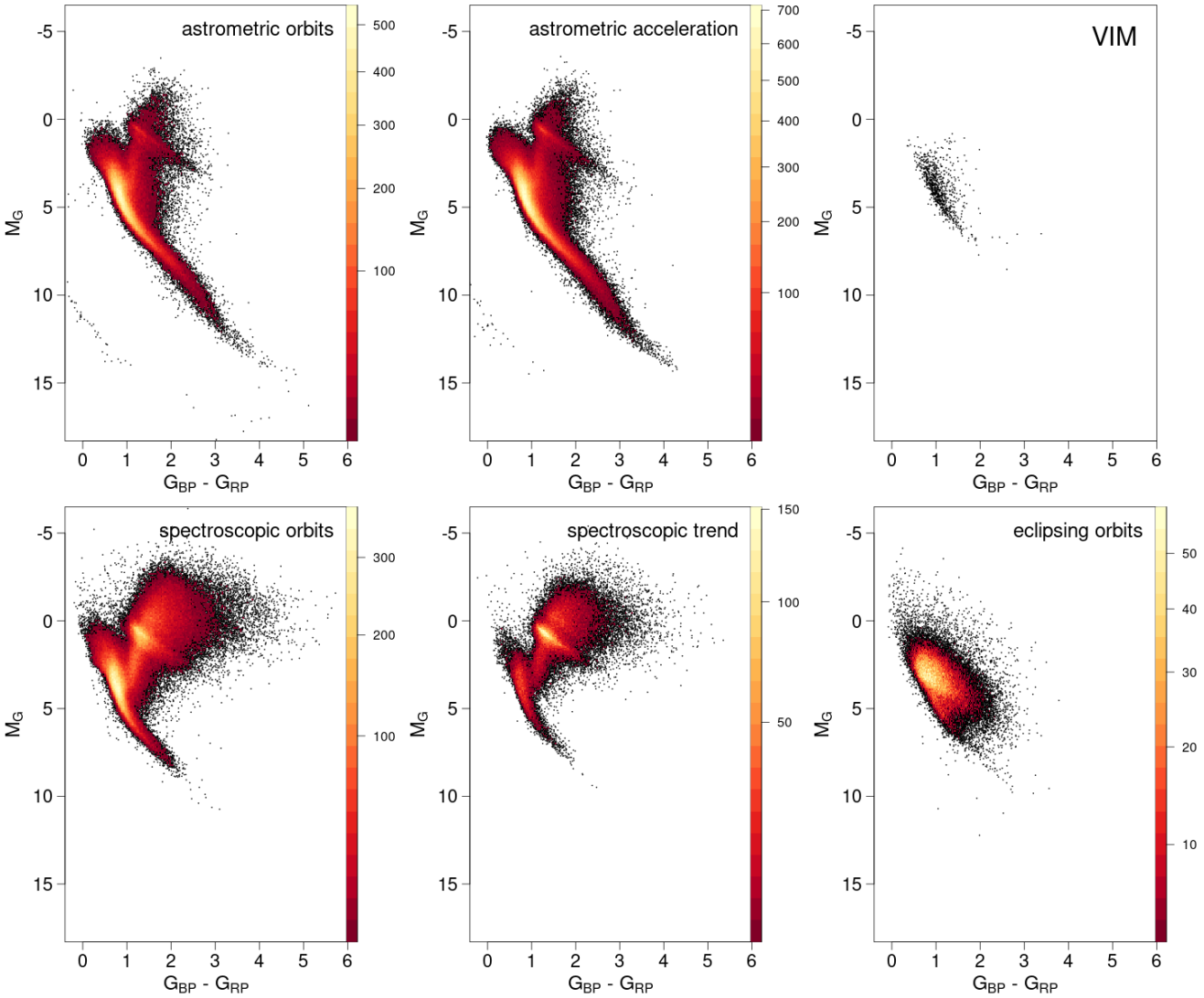

Fig. 4. *Gaia* HRD, uncorrected for extinction, for all NSS solutions with a relative parallax error of better than 20%. No selection is done on the photometric quality. The colour scale represents the square root of the relative density of stars. *Top*: astrometric binaries, (*a*): `all Orbital* solutions plus AstroSpectroSB1`, (*b*): `Acceleration` solutions, (*c*): `VIMF`; *bottom*: Spectroscopic binaries with (*d*): `SB*` orbits and (*e*): `NonLinearSpectro`, (*f*): eclipsing binaries.

The Hertzsprung-Russell diagrams (HRDs) for the various categories are represented Fig. 4 for sources with a parallax signal-to-noise ratio (S/N) >5. We note that the used parallax is that from the NSS solution for what concerns the putative astrometric binaries, while we use the one from the main catalogue for spectroscopic and eclipsing binaries.

2.2. Table construction

Although we refer to the online documentation and the articles accompanying this data release for a detailed understanding of the data processing, it is of interest to describe how the NSS data were obtained, starting with their input data selection, as this is one first key to understanding the NSS selection function.

The basic NSS processing procedure selected its input sources from those that had a poor goodness of fit (GoF) in the upstream results, either in the astrometric or in the spectroscopic processing, or from those that were detected as eclipsing variables; the only exception is the `OrbitalTargetedSearch` (see Sect. 2.2.2), where a predefined source list was given as input to the astrometric orbital fit, irrespective of their actual GoF in the single star solution.

2.2.1. Astrometric binaries: main processing

As *Gaia* DR3 is the first publication of NSS solutions, we decided to limit the content to the most significant ones, this criterion being relaxed for further releases, and the motivation for this is explained below. The definition of the input source list started after *Gaia* DR2, where it was assumed that the sources with `ruwe` > 1.4 represented a reasonable threshold for sources with problematic astrometric solutions[2]. A more recent analysis (e.g. Penoyre et al. 2022) appears to suggest that a lower threshold could have been chosen, but this value also has the advantage that it decreases the processing requirements.

To this `ruwe` > 1.4 criterion, $G < 19$ was added in order to keep the best S/N. The sample defined in this way contains many contaminants, partially resolved rather than unresolved

[2] Cf. Gaia DR2 documentation.

sources. In particular, for a double star with a projected separation between components of between ≈9 mas and ≈0.27″ (Lindegren 2022), depending on the magnitude difference, the epoch position is not exactly on the photocentre[3], meaning that the astrometry of such sources is perturbed and the source is likely to have been selected.

Consequently, the criterion `ipd_frac_multi_peak` ≤ 2 was added to avoid double stars with a large separation and `ipd_gof_harmonic_amplitude` < 0.1 was also added to reject pairs with smaller separations. The `visibility_periods_used` > 11 criterion was also added (>12 for orbital solutions) in order to avoid spurious solutions[4].

However, the sample was still polluted, and so another criterion was added, this time based on photometry, as an attempt to avoid sources with light being contaminated by a neighbour. For this purpose, we made use of the corrected *BP* and *RP* flux excess factor C^* associated to its uncertainty $\sigma_{C^*}(G)$ as defined by Riello et al. (2021, Eqs. (6) and (18)), selecting sources with $|C^*| < 1.645\sigma_{C^*}$ only.

2.2.2. Astrometric binaries: alternative processing

As described by Holl et al. (2023b), alternative orbit determination algorithms have been run on two different input lists. The first one is based on astrometric binaries that could not be successfully modelled by any model in the main processing pipeline, for which a more complex handling was attempted, `nss_solution_type` = `OrbitalAlternative*`. These sources originate from the same list as described in Sect. 2.2.1. The second one is a sample of sources with detected companions published in the literature, `nss_solution_type` = `OrbitalTargetedSearch*`, where all sources have been kept for processing.

2.2.3. Spectroscopic binaries

The selection of the sources that had to be treated by the spectroscopic binary pipeline was based on sources with enough measurements, and a large enough dispersion of these measurements, rejecting stars not having `rv_renormalised_gof` > 4, `rv_nb_transits` ≥ 10, and 3875 K < `rv_template_teff` < 8125 K, or detected as SB2. One may notice that there are more than 6000 sources with a SB solution that have no average radial velocity in the `gaia_source` main catalogue. In that case (as in the other cases where a SB solution is given), the `center_of_mass_velocity` gives the systemic velocity. The absence of a mean RV for what concerns SB2s is normal, as the main spectroscopic processing did not compute this value. For SB1s, it may be useful to note that the computation had not been performed for the sources that were considered either peculiar, potentially SB2, too hot, or with emission lines. Consequently, when some SB results appear doubtful, it may then be useful to check whether `gaia_source.radial_velocity is NULL` for these sources. More details are given in Gosset et al. (in prep.).

[3] We coined the word 'Gaiacenter' in Kervella et al. (2022) by analogy with the 'Hippacenter' defining the actual pointing of the epoch HIPPARCOS observations of double stars (Martin et al. 1997).

[4] Although the DoF is still large enough as there are about 18 astrometric observations per visibility period on average, one may still be cautious with solutions for which there are a low number of visibility periods.

2.2.4. Eclipsing binaries

The input list for candidate eclipsing binaries contained about 2.1 million sources that can be found in the `vari_eclipsing_binary` *Gaia* DR3 table. Their selection is described by Mowlavi et al. (2023); see also the online documentation. The selection of the subset therein for which an orbital solution has been computed is described by Siopis (in prep.).

2.3. Output filtering

Once the first processing results were analysed, it appeared that the cleaning of the input list had still left a very large fraction of spurious solutions. This is why it was decided to keep the most significant solutions for *Gaia* DR3: a general filter was applied to keep those with goodness of fit smaller than 50 and significance >5 (>2 for `OrbitalTargetedSearch*`). The `significance` is computed as the S/N of the semi-major axis for astrometric orbits, (a_0/σ_{a_0}), as the S/N of the acceleration module for acceleration solutions, and as the S/N of the semi-amplitude for spectroscopic binaries, (K_1/σ_{K_1}). Supplementary filtering was applied during the processing or at post-processing level as described for the various models below.

Astrometric binaries: acceleration solutions. One could naively hope that the estimated accelerations would allow us to detect binaries of intermediate period and provide some useful information about the binary, such as the minimum mass producing the given acceleration on the primary. However, the situation appears more complex. The acceleration values themselves are not discussed, and it can be seen that these solutions improve the baseline solution; for example, the giant branch appears slightly thinner for an HRD produced using the parallaxes from the acceleration solution compared to those from the main catalogue.

However, two effects conspire to make the interpretation of the acceleration term unclear. The first one originates from the organisation of the NSS processing: acceleration solutions were attempted before any other solutions, and kept if found to be sufficiently significant with a reasonable GoF. The (unwanted) effect is that some solutions that could have received a full orbit parametrisation were not attempted and appear in the NSS catalogue with an acceleration solution instead. The second effect is that an acceleration term can be significant even for short periods or very long periods. This is demonstrated by the analysis by Lindegren (2022).

The following filtering has been applied (see documentation and Halbwachs et al. 2023, for details): the sources which have been kept are those with significance $s > 20$ and $\varpi/\sigma_\varpi > 1.2\,s^{1.05}$ and GoF < 22 for `Acceleration7` and $\varpi/\sigma_\varpi > 2.1\,s^{1.05}$ and GoF < 25 for `Acceleration9`. Despite this, it is known that a large fraction of the acceleration solutions are not intermediate-period binaries as one would expect, but are rather short- or long-period binaries instead.

Astrometric binaries: `Orbital` solutions. The processing of orbital solutions starts by a period search. Unfortunately, this may lead to the detection of periods related to the scan law, rather than due to some true periodic motion: partially resolved objects with fixed position may give a signal depending on the scanning angle with respect to the orientation of the pair. This problem is fully analysed by Holl et al. (2023a). Consequently, most detected periods below ≈100 days were erroneous, leading to solutions with huge and incorrect mass functions.

To circumvent this problem, the following filtering was adopted (Halbwachs et al. 2023): parallax $S/N > 20\,000/\texttt{period}$, $\texttt{significance}\ s = a_0/\sigma_{a_0} > 5$, and $s > 158/\sqrt{\texttt{period}}$, $\texttt{eccentricity\_error} < 0.182 * \log_{10}(\texttt{period}) - 0.244$.

Astrometric binaries: alternative processing. Aggressive post-processing filtering approaches for both samples produced subsets of solutions that were assigned `OrbitalAlternative*` and `OrbitalTargetedSearch*` solution types, respectively, in the *Gaia* DR3 archive. For both cases, subsamples of sources that passed a variety of validation procedures were further assigned `OrbitalAlternativeValidated` and `OrbitalTargetedSearchValidated` solution types (see Holl et al. 2023b, for details).

Inspection of the `OrbitalAlternative` solutions reveals that the caveat of unrealistically large inferred companion masses at short orbital periods is not entirely removed. A few percent of spurious solutions still likely contaminates this sample.

Spectroscopic binaries. Only the sources with GoF <3, $|\texttt{center\_of\_mass\_velocity}| < 1000$ km s^{-1}, $K_1 < 250$ km s^{-1}, and $\texttt{efficiency} > 0.1$ were kept, where `efficiency` is a measure of the correlation between parameters. One of the most important problems found after processing was the presence of many spurious SB detections with short periods. For this reason, the lower confidence threshold on the period was adapted depending on the period itself: it was set to 0.995 for $P < 1$ d, 0.95 for $P > 10$ d, and $-0.045 \log P + 0.995$ in between. For details on this and other filtering during the spectroscopic processing, see Gosset et al. (in prep.).

Despite all this, the comparison of NSS results with catalogues of known binaries shows that for a few percent of the SB1 solutions, the period may still be incorrect, mainly because of the sparse time sampling. When these sources have both an `SB1` and `Orbital` solution, such cases may be spotted by comparing the respective semi-amplitudes. Short periods with large `ruwe` (e.g. >1.4) are frequently suspect; some may be the inner system of a triple system, but most may instead be some kind of aliases of a longer period.

Inspecting the SB1 solutions, an overdensity of solutions with periods around the precession period (62.97 days) can be noted, in particular by selecting sources with large astrometric excess noise (see Fig. 5). These solutions are spurious and due to some offset in the astrometric coordinates, which generates in turn a spurious variation of the computed epoch radial velocities; as this offset depends on the scanning angle, it occurs with a periodicity linked to the precession of the satellite. The inaccuracy of the astrometric coordinates is most probably due to the fact that they are partially resolved binaries or double stars, which is confirmed by the fact that we also see this overdensity when selecting sources with $\texttt{ipd\_frac\_multi\_peak} > 20$. Holl et al. (2023a) describe the effect of the scanning law on the NSS solutions in more detail.

Spurious SB1 solutions can also be generated by pulsation of the source, as in RR Lyrae and Cepheids. In many cases, the SB1 solution will have the same period as the pulsation, but in other cases, due to the sparse sampling, the pipeline can find a Keplerian fitting solution at a different, typically shorter period. During the release validation, SB1 solutions of sources identified by *Gaia* as RR Lyrae or Cepheids were removed from the release.

Another source of spurious SB1 solutions is contamination from a nearby, brighter star. As explained by Seabroke et al. (2021), and noted by Boubert et al. (2019), the RVS spectrum of sources extracted at a given transit can be contaminated by a nearby source, producing spurious values of the radial velocity. In *Gaia* DR3, the RVS pipeline includes a deblending algorithm, which is nevertheless limited to spectra with overlapping windows (see Seabroke et al. 2021, for details).

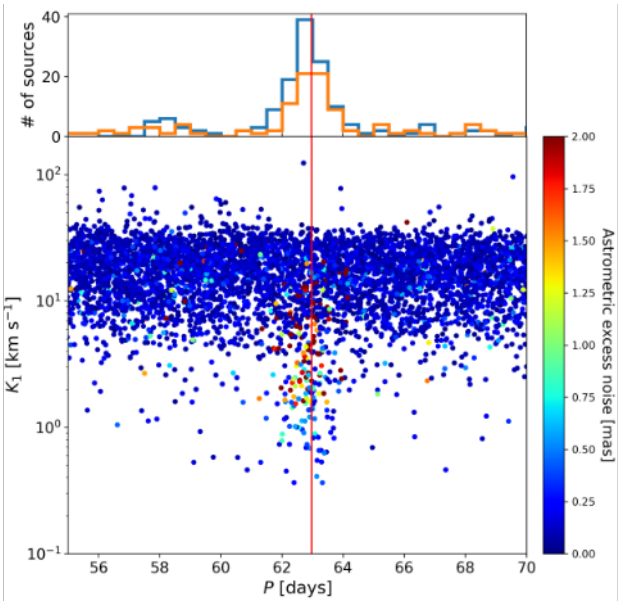

Fig. 5. K_1 semi-amplitude vs. period diagram of `SB1` solutions, colour coded according to their `astrometric_excess_noise`. The diagram shows the presence of an overdensity of solutions at periods near the precession period (marked with a vertical line) with large astrometric excess noise. The histogram at the top shows the density of solutions with astrometric excess noise larger than 1 mas (blue line) and of those with $\texttt{ipd\_frac\_multi\_peak} > 20$ (orange line).

Eclipsing binaries. At post-processing, only the sources with $0.2 < \texttt{efficiency} \leq 1$ and $\texttt{g\_rank} \geq 0.6$ were kept, where the rank is a measure of the quality of the fit. See the online documentation (Pourbaix et al. 2022, Sect. 7.6) for details.

3. Completeness

The resulting NSS dataset is the result of a selection process in three successive steps: (a) the selection of the input list, discussed Sect. 2.2; (b) the sources for which the orbital motion can be preferentially detected by the processing; (c) the filtering done at post processing, indicated Sect. 2.3.

In this section, we give some indications concerning the second step. One main reason for the expected non-uniformity of orbit detections is the number of observations and their temporal distribution. As this is governed by the scanning law of the *Gaia* satellite (see e.g. Fig. 7), this should appear clearly on a sky plot, and this is discussed in Sect. 3.1.

However, even with a given set of observations, all orbits are not perfectly equivalent. First, the period distribution of astrometric orbits shows a prominent lack of solutions around one year, which was obviously expected due to the difficulty in decoupling the orbital from the parallactic effect. There are other more subtle biases depending on the orbit itself; these are discussed Sect. 3.2. The distribution of solutions is finally discussed within the 100 pc horizon at Sect. 3.3 and the completeness is also studied for HIPPARCOS stars Sect. 4.2.2.

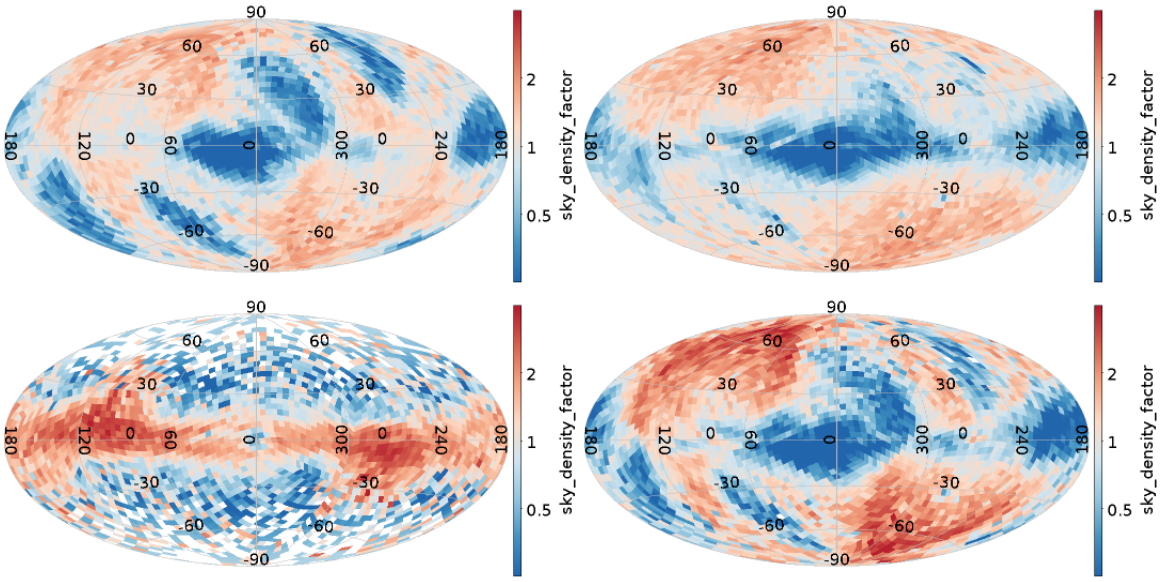

Fig. 6. Sky density factor (Galactic coordinates, healpix level 4, log scale, see text) illustrating the excess or deficit factors of NSS sources compared to their sky average value. *Panel a*: `SB*`, *panel b*: `Acceleration`, *panel c*: `EclipsingBinary`, *panel d*: `Orbital*`.

3.1. Sky distribution

Over the sky, the distribution of the various solution types shows the expected higher density along the Galactic plane together with a larger excess at high ecliptic latitudes around $l \pm 100°$. The latter is due to a larger number of observations, and therefore to a larger probability of detecting periodically variable motions.

However, this tells us little about whether or not the (in)completeness is uniform over the sky. Although we may have for example more eclipsing binaries among young stars, let us assume for a moment that F, the true (unknown) fraction of binaries, is uniform over the sky, and that our NSS samples are roughly complete up to some given magnitude $G_{\rm max}$.

Dividing up the sky in healpix (Górski et al. 2005) level 4 equal-area pixels, we note N_j the number of sources up to $G < G_{\rm max}$ in the full *Gaia* catalogue in a given healpix cell j, and n_j the corresponding number of NSS of a given type up to $G < G_{\rm max}$. With $f = \mathrm{med}(n_j/N_j)$, the empirical median of the ratio over the sky as estimate of F, we call 'sky density factor' $d_j = \frac{n_j}{f\,N_j}$. This factor gives the up or down factor of the average NSS fraction and should be a noisy value around 1 if F is approximately constant over the sky.

Figure 6 shows the sky density factor for several solution types truncated up to a reasonable $G_{\rm max}$ value in healpix level 4 pixels. As this density factor may be attributed to the number of observations available, Fig. 7 presents the ratio of useful observations over the sky for photometry and astrometry, with the same scale for comparison purposes.

For all types of binaries, the expected deficit of sources near the Galactic centre can be seen because of both high density and poor coverage. The distribution of spectroscopic binaries ($G_{\rm RVS-max} = 12$, Fig. 6a) is also as expected with a smooth pattern of regions with higher numbers of field-of-view transits. The non-uniformity is less expected for eclipsing binaries ($G_{\rm max} = 18$, Fig. 6c) with a slight excess at the anticentre and an excess – larger than expected from the number of transits – around $l \pm 100°$ towards high ecliptic latitudes. For acceleration solutions ($G_{\rm max} = 15$, Fig. 6b), there is a deficit in the Galactic plane and an excess at high ecliptic latitudes. This is worse for `Orbital` solutions ($G_{\rm max} = 15$, Fig. 6d), which may be due to the fact that orbital solutions require a greater number of observations than acceleration solutions as there are a greater number of parameters to determine. Again, the sky density factor is relative to the average over the sky, meaning that an excess in some regions may rather indicate a deficit in the rest of the sky. Some or part of the above features of the astrometric solutions likely originate from the input source selection, where sources suspected to be resolved doubles were excluded, which is more frequent in the Galactic plane.

3.2. Astrometric orbit detection sensitivity as a function of orbital inclination

Gaia is observing sources with a cadence and scan angle ψ determined by its scanning law. Depending on whether a binary system is seen face-on (inclination $i = 0°$ or $i = 180°$) or edge-on ($i = 90°$), the detection probability of the astrometric orbit varies. An edge-on orbit that is oriented north-south and is being observed only with 1D astrometry along the east-west axis is undetectable. This extreme example does not occur for *Gaia*, but it illustrates that we can expect continuous variation as a function of inclination angle, with edge-on orbits having the lowest detection probability.

To obtain an empirical estimate of the expected dependency, we simulated 50 000 circular orbits ($e = 0$, $\omega = 0$) with the following fixed parameters: distance 20 pc; period of 500 days; primary mass 1 $\mathcal{M}_\odot$; companion mass of 1 $\mathcal{M}_{\rm Jup}$, and hence a semi-major axis of $a_0 = 0.059$ mas for the orbit of the host. The ascending node Ω was uniformly distributed. We simulated inclinations such that $\cos i$ is uniformly distributed, as expected for isotropic orbit orientation in space.

To each orbit we assigned a realistic *Gaia* DR3 time sampling with associated scan angles randomly retrieved from approximately 1000 real sources distributed over the entire sky with the aim of averaging scan-law-dependent effects. We then computed the rms dispersion of the AL signal $w_{\rm k1}$, Eq. (B.6), caused by the astrometric orbit only, that is, neglecting proper and parallactic motion. This dispersion shows a clear

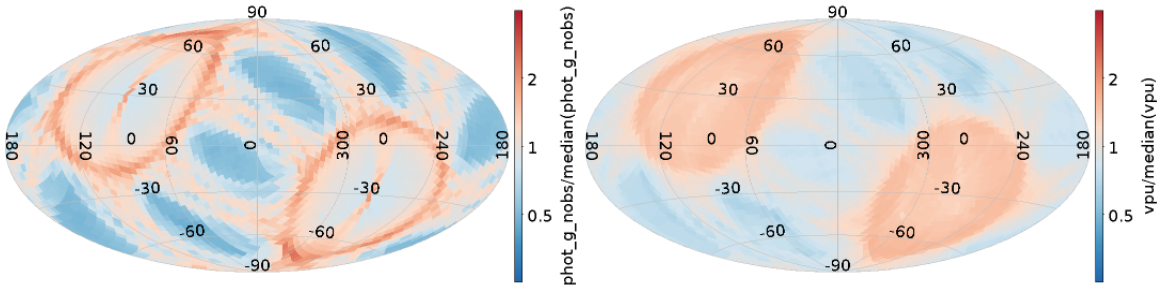

Fig. 7. Ratio of the number of photometric observations over their median values for $G < 18$ (*left*) and the ratio of the number of visibility periods used in astrometry over their median values for $G < 15$ (*right*), with the same colour scale (from $\frac{1}{4}$ to 4) as Fig. 6.

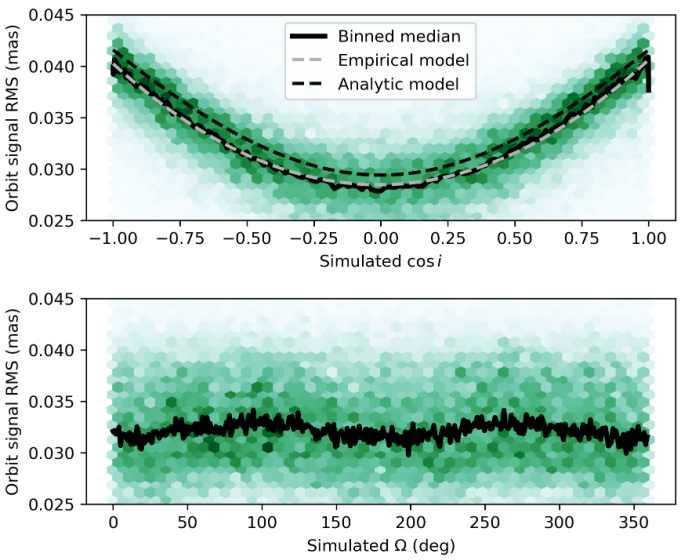

Fig. 8. Density histograms of simulated orbit signal dispersion as a function of $\cos i$ (*top*) and Ω (*bottom*). The black solid curve shows the running median value. *Top panel*: the empirical and analytic models are shown as dashed grey and black lines, respectively. Edge-on orbits have $\cos i = 0$ and face-on configurations have $|\cos i| = 1$.

dependence on inclination angle (see Fig. 8), with the expected minimum for edge-on orbits. The vertical scatter is caused by the variation in the number of assigned *Gaia* observations and their scan-angle distribution for a particular time-series realisation. The dependence on ascending node (Fig. 8, bottom) is much weaker but noticeable. Because we limited our simulation to circular orbits, there is no dependence on the argument of periastron.

In Appendix C we analytically derive the following expression, Eq. (C.10), for the rms of $w_{\rm kl}$ as a function of $\cos i$, which is valid for one-dimensional (1D) along-scan (AL) observations as used for DR3 (Lindegren et al. 2021) under the assumption of circular orbits and random scan angles:

$$\mathrm{rms}(w_{\rm kl})(x) = \frac{a_0}{2}\sqrt{1+x^2} \quad \text{with} \quad x = \cos i. \tag{1}$$

This dependency is shown as 'Analytic model' in Fig. 8. A fit with a quadratic polynomial is also shown as 'Empirical model'. The analytic model reproduces the data very well, both in absolute amplitude and shape, except for a small amplitude offset which probably reflects the fact that the *Gaia* scan angles are not random and are sometimes restricted in range.

Because the amplitude of the orbit signal is the principal factor in deciding whether an orbit can be detected[5], there is no need to simulate the complete processing chain (Halbwachs et al. 2023; Holl et al. 2023b). Our simulations demonstrate that the signal of a face-on orbit is $\sqrt{2}$ larger than that of an edge-on orbit, which means that the former is more likely to be detected.

3.3. The Gaia catalogue of nearby stars

A clean catalogue of 331 312 sources within 100 pc of the Sun (Gaia Collaboration 2021b, GCNS) was published together with *Gaia* EDR3. This catalogue would represent a useful subset for the completeness analysis.

As the NSS parallaxes of `Orbital` or acceleration solutions may supersede the EDR3 ones, it is of interest to first analyse their potential impact on the GCNS content. One finds 116 orbital sources that would now enter GCNS using the following query:

```
SELECT NSS.source_id, GS.phot_g_mean_mag,
NSS.parallax,
NSS.parallax_error, GS.parallax as gs_parallax,
GS.parallax_error as gs_parallax_error
FROM user_dr3int6.nss_two_body_orbit NSS,
user_dr3int6.gaia_source GS
LEFT JOIN external.gaiaedr3_gcns_main_1 GCNS ON
NSS.source_id = GCNS.source_id
WHERE GCNS.source_id IS NULL
AND NSS.source_id = GS.source_id
AND NSS.parallax > 10
```

Using a similar query, 89 sources with an acceleration solution would enter GCNS, giving a total of 205 sources. These numbers would change by 13% only if we were to take a 1σ margin, and so the random errors have a weak influence on this.

Conversely, one may count sources that should no longer belong to GCNS according to their new parallax:

```
SELECT NSS.source_id, GS.phot_g_mean_mag,
NSS.parallax,
NSS.parallax_error, GS.parallax as gs_parallax,
GS.parallax_error as gs_parallax_error
FROM user_dr3int6.nss_two_body_orbit NSS,
user_dr3int6.gaia_source gs,
external.gaiaedr3_gcns_main_1 GCNS
WHERE NSS.source_id = GS.source_id AND
NSS.source_id = GCNS.source_id AND
NSS.parallax < 10,
```

[5] This holds when neglecting complications with e.g. periods of around 1 yr due to crosstalk with parallax- or scan-angle-dependent effects (Holl et al. 2023a).

amounting to 415 sources for orbital solutions plus 413 sources for acceleration solutions, giving a total of 828 sources.

Having 4723+4523 = 9246 astrometric NSS sources which are in the GCNS, these 828 sources represent 9% of the orbital plus acceleration solutions which may no longer be in the GCNS while 2% may now enter. This means that any study of the NSS completeness within the GCNS should use the NSS parallax rather than the one from the main catalogue.

One may also note that the balance between the number of NSSs that would be rejected from GCNS and the number that would enter illustrates one long-since recognised adverse effect of the random errors (Eddington 1913; Trumpler & Weaver 1953). The parallaxes of NSS sources managed as single stars in DR3 have a significant error, which should now be much reduced in the NSS tables; this, added to the asymmetric distribution of the parallaxes, means that binary sources preferentially entered GCNS that should not belong to it. As the DR3 NSS catalogue contains only a small fraction of the actual unresolved astrometric binaries, using the GCNS to compute a binarity fraction may produce a small positive bias.

As a clarification of the GCNS content using the NSS parallax is outside the scope of this article, we keep the GCNS for reference in what follows. We show the fraction of NSS sources among $G < 19$ GCNS sources as a function of parallax for all solution types (see right panel of Fig. 9). In these figures and the following, we add the `AstroSpectroSB1` counts both to orbital solution counts and SB counts, in addition to counting them independently and, for the comparison to be fair, we restricted the ratios to the typical magnitude ranges used respectively for astrometric, spectroscopic, and eclipsing binaries.

What first appears is the conspicuous increase in the fraction of SBs up to 100 pc. One reason for this may be the transition from the $G_{\rm RVS} < 12$ population of dwarfs to giants, as can be seen in Fig. 4d, with the latter having a better intrinsic RV precision at a given apparent magnitude (Katz et al. 2023), and thus a larger binary detection probability; however, a difference in the binary fraction between dwarfs and giants cannot be excluded. Second, contrary to what might have been expected, the total fraction of orbital and acceleration solutions, about 3%, appears roughly constant with distance in the GCNS, despite all the complex filtering that has been applied. For comparison, the fraction of NSSs among DR3 sources (left panel of Fig. 9) shows a drop in astrometric solutions with distance beyond 100 pc, while the fraction of spectroscopic binaries (SB+`nss_non_linearspectro`) does not vary as sharply. From this comparison, we retain the fact that even if the absolute value of the astrometric binary fraction is difficult to extrapolate after all the filtering done, the fact that it appears roughly uniform with distance in a first approximation in the GCNS sample means that this sample could be useful for studying the properties of the astrometric binaries.

Consequently, the fraction of NSSs among GCNS may provide useful insights, and Fig. 10 represents this ratio versus G apparent and absolute magnitude of the pair, respectively. The absolute magnitude `mg_gspphot` originates from the General Stellar Parametrizer from Photometry (GSP-Phot), which computed the astrophysical parameters of stars from the low-resolution BP/RP spectra and is available in the `astrophysical_parameters` table.

4. Caveats

Many validations have been performed and described in the catalogue documentation (Pourbaix et al. 2022), accompanying papers (Halbwachs et al. 2023; Holl et al. 2023b; Gosset et al., in prep.; Siopis, in prep.), and the independent validation of all catalogues (Babusiaux et al. 2023). Elsewhere in this article, we also check the distribution tails of some parameters which allowed us to discover undesired aspects and we indicate ways to circumvent them. Here, we describe two supplementary tests that draw attention to some properties of the catalogue, the first analysing the distribution of orbital parameters, the second comparing the results to binaries detected externally.

4.1. Distributions and biases of astrometric orbit parameters

Under the assumption that the orbits of binary systems are randomly oriented, we can infer the expected distributions in the geometric elements of the corresponding astrometric orbits, that is, the inclination i, the argument of periastron ω, and the longitude of the ascending node Ω[6]. In an ideal experiment, we expect to recover uniform distributions in $\cos i$, Ω, and ω. Here, we inspect the observed distributions of these parameters in DR3.

4.1.1. Observed distributions of geometric elements in DR3 solutions

Figure 11 shows the distributions of $\cos i$, Ω, and ω for the solution types `Orbital` and `AstroSpectroSB1`. To mitigate effects related to incomplete orbit coverage, we selected solutions with orbital periods shorter than 1000 days, which roughly corresponds to the DR3 time span.

For `Orbital` solutions, there is a strong modulation in $\cos i$. Although the expected suppression of edge-on orbits is present, the observed distribution deviates significantly from the empirical model defined in Sect. 3.2. For progressively face-on configurations with increasing $|\cos i|$ there is an excess of solutions compared to the model. Beyond the modes $|\cos i| \gtrsim 0.85$, the number of detected almost-face-on orbits drops sharply and far below the expected level. We also observe a smooth modulation of the Ω distribution[7] with a single minimum at $\Omega = 90°$ and a bimodal modulation of the ω distribution with minima at $\omega = 90°$ and 270°.

For `AstroSpectroSB1` solutions resulting from the combined analysis of *Gaia* astrometry and RVs, the $\cos i$ distribution shows good agreement with the empirical model for edge-on and intermediate configurations without regions of excess detections. However, there is also a clear lack of face-on orbits compared to the empirical expectation. This is influenced by the decreasing orbital RV signature towards face-on orbits. As `AstroSpectroSB1` solutions require independent detections in both astrometry and RV, the lack of face-on orbits can be expected. The modulation in Ω is similar to `Orbital` solutions but weaker[8] and there is no apparent modulation in the ω distribution.

[6] These Campbell elements were computed from the Thiele-Innes coefficients (A, B, F, G) – which are part of the archive table – using standard formulae (e.g. Halbwachs et al. 2023); software tools are available at `https://www.cosmos.esa.int/web/gaia/dr3-software-tools`

[7] The ascending node extracted from the Thiele-Innes coefficients of astrometric orbits is constrained to ±180°. By convention, the value between 0 and 180° is chosen.

[8] The Ω parameter is only constrained by the astrometry data. When deriving it from the `AstroSpectroSB1` Thiele-Innes coefficients we neglected the additional RV information that would have allowed us to compute it unambiguously.

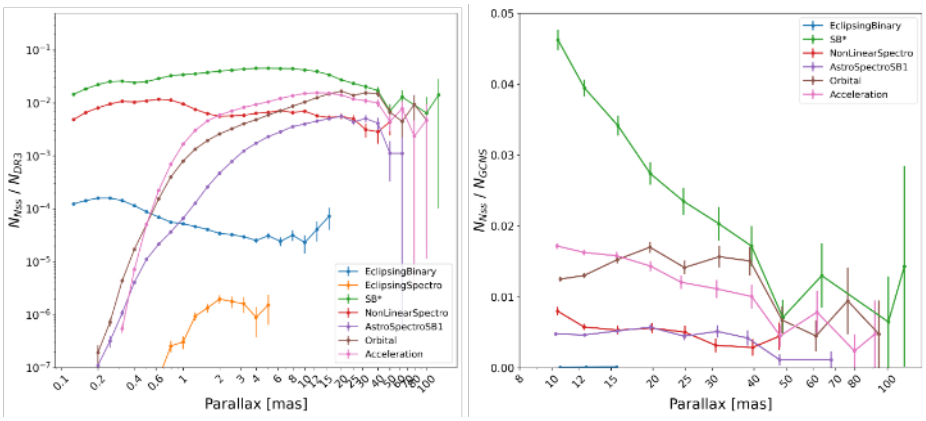

Fig. 9. Fraction of NSS solutions among EDR3 sources vs. parallax (*left*) and fraction of NSS sources in GCNS (*right*). In both figures, we added `AstroSpectroSB1` counts to `Orbital` counts and to `SB*`=`SB1`+`SB2` counts in addition to counting them individually, and we restrict the ratios to $G_{\rm RVS} < 12$ sources only for `SB*` and `NonLinearSpectro`, to $G < 19$ for `Orbital` and `Acceleration` solutions, and to $G < 20$ for eclipsing binaries.

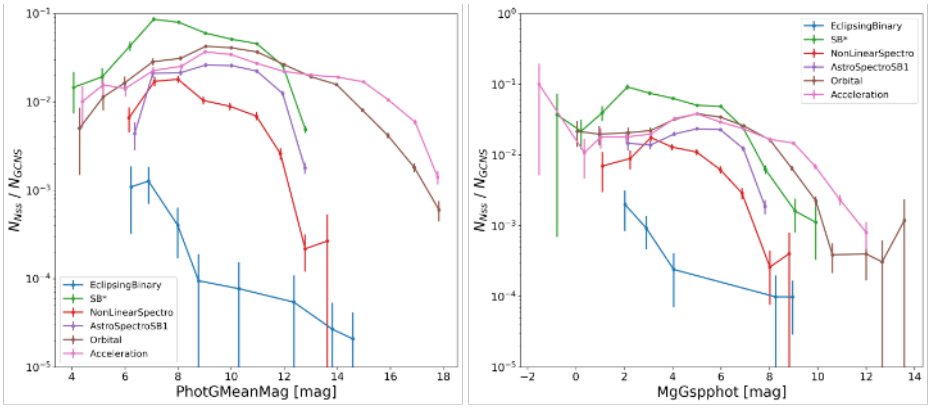

Fig. 10. Fraction of NSS solutions among GCNS sources vs. G apparent magnitude (*left*) and vs. G GSP-Phot absolute magnitude (*right*). The same constraints as mentioned in Fig. 9 have been applied.

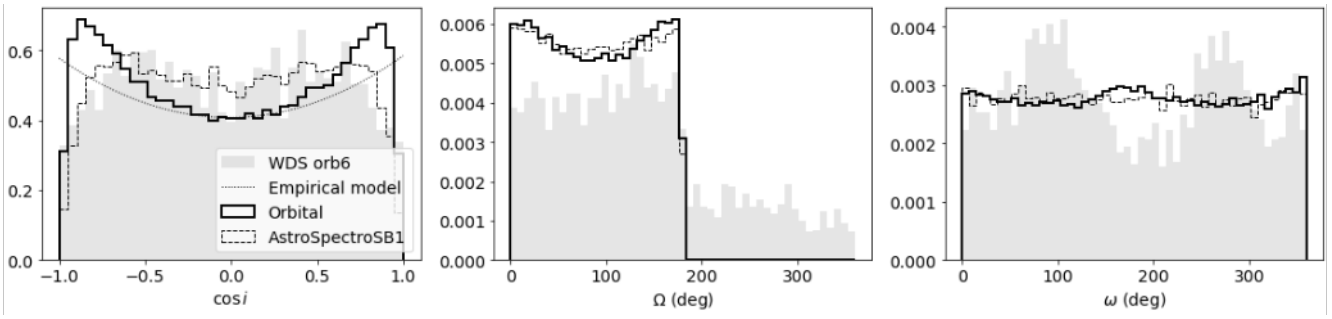

Fig. 11. Normalised distributions of $\cos i$ (*left*), Ω (*middle*), and ω (*right*) parameters. `Orbital` (solid lines, 122 989 entries) and `AstroSpectroSB1` (dashed lines, 29 770 entries) solutions with $P < 1000$ d are shown. The orb6 solutions from the literature (3405 entries, without filter on period) are shown in grey. *Left panel*: the dotted line shows the empirical model defined in Sect. 3.2, which was re-scaled on the five central histogram bins. *Right panel*: we have suppressed the circular solutions with $\omega = 0$.

Figure 12 shows the $\cos i$ distributions for systems within 200 pc, where the S/N is on average higher and the astrometric-orbit detection can be expected to be more complete. This is confirmed by the `Orbital` solutions that follow the empirical model nicely across most of the inclination range. This validates our model for the inclination-dependent detection efficiency of astrometric surveys (Sect. 3.2). The Ω and ω distributions for this subset of solutions are approximately uniform. We inspected other astrometric solution types but do not discuss these here because they have fewer (<1000) entries and are therefore less suitable for distribution analyses.

4.1.2. Origins of the geometric element biases

Concentrating on the `Orbital` solutions, we identify three main deviations from the expected uniform distributions in the low-S/N regime, which comprise most solutions and therefore dominate the overall distributions in Fig. 11: (a) a pronounced suppression of face-on orbits; (b) a smooth modulation of the Ω distribution with a single minimum; and (c) a bimodal modulation of the ω distribution.

In Appendix D.1 we identify the origin of features (a) and (b) in the linear fit of the Thiele-Innes coefficients to noisy data and reproduce these biases qualitatively in simulations. The noise bias in the recovered inclination shifts solutions away from face-on configurations leading to the observed excess at intermediate inclinations[9]. A modulation akin to feature (c) can also be caused by noise biases, albeit with a 90° phase shift. In Appendix D.2,

[9] As the survey is not volume-limited, *Gaia*'s sensitivity variation in principle leads to an expected excess of face-on orbit detections. We believe that such effects are secondary in the context of DR3.

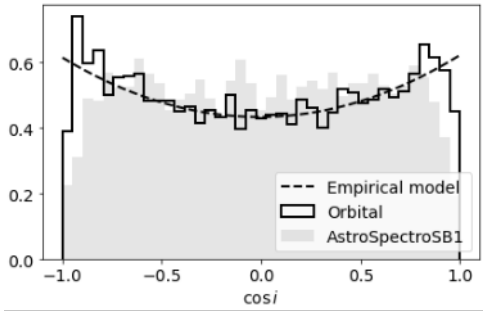

Fig. 12. Normalised distributions of $\cos i$ within 200 pc for `Orbital` (solid line, 9106 entries) and `AstroSpectroSB1` (grey-filled, 5735 entries) solutions with $P < 1000$ d and $\varpi > 5$ mas. The dashed line shows the empirical model defined in Sect. 3.2.

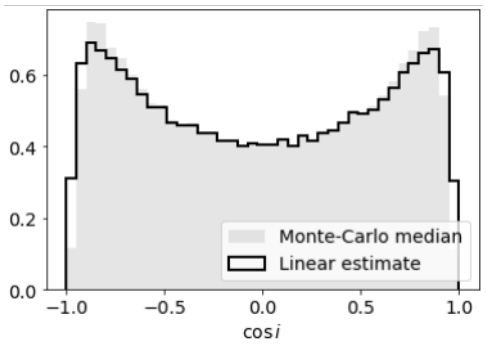

Fig. 13. Normalised distributions of $\cos i$ for non-circular `Orbital` solutions with $P < 1000$ d (121 207 entries). The linearised and Monte Carlo estimates are shown as a solid line and a filled grey area, respectively.

we show that feature (c) is instead explained by the application of a semi-major axis significance threshold when selecting the solutions to be published.

4.1.3. Geometric elements from Monte Carlo resampled Thiele-Innes coefficients

Instead of using the linearised formulae (e.g. Halbwachs et al. 2023) for converting A, B, F, and G values and uncertainties to a_0, i, Ω, and ω, one can use Monte Carlo resampling which accounts more accurately for the parameter correlations (Appendix D.3). As an example of the potential effects that this may have, we computed an alternative estimate of the orbital inclination for individual solutions as the median of the resampled Monte Carlo distribution. The difference between linearised and Monte Carlo estimates on the inclination distribution is shown in Fig. 13, where we see that the apparent depletion of face-on orbits is more pronounced when applying the resampling. We note that the resampled distributions of a_0, i, Ω, and ω are seldom Gaussian and the median value is not always a good representation. Whether it is advisable to use the linearised estimate or Monte Carlo resampling depends on the particular problem and individual orbital solution.

4.1.4. Comparison with known astrometric orbits

Figure 11 also shows the distributions of geometric elements compiled in the "Sixth Catalog of Orbits of Visual Binary Stars" (orb6, Hartkopf et al. 2001)[10]. The orb6 inclination distribution is bimodal with modes at $|\cos i| \simeq 0.5$, which could be interpreted as the same signature of lacking face-on orbits as for *Gaia* `Orbital` but setting in earlier. The comparison with the simulated inclination biases in Fig. D.1 would then lead to the interpretation that the average S/N is higher for the *Gaia* orbits than for the orb6 solutions. However, we caution that the orb6 dataset is of heterogeneous nature and such comparisons have to be made more carefully by accounting for differences in period range, significance, and other factors.

The orb6 Ω distribution does not seem to exhibit the minimum at $\Omega = 90°$ seen for *Gaia* `Orbital`. In contrast, the orb6 ω distribution shows clear modes at $\omega = 90°$ and 270°, that is, shifted by 90° relative to *Gaia* `Orbital`. Our simulations in Fig. D.3 reproduce the peak location for orb6 orbits but not for `Orbital` solutions.

It is clear that the increase in astrometric orbit solutions by a factor of more than 40 delivered by *Gaia* DR3 compared to orb6 will facilitate a multitude of population-level studies and push forward our understanding of stellar binary systems.

4.1.5. Recommendations

The observed features in the distributions of i, ω, and Ω are the result of variations in the detection sensitivity of the survey, of selection effects, and of biases that are introduced in the astrometric non-single star processing. Their presence is not specific to *Gaia* and astrometric orbits in the literature show similar features. The geometric elements of DR3 orbits are encoded in the Thiele-Innes coefficients and different conversion methods can be applied depending on the use case and individual solution. Both the distribution features and the conversion aspects have to be considered in scientific analyses of *Gaia* DR3 orbital parameters and their distributions.

4.2. Proper motion anomaly of HIPPARCOS stars of the NSS sample

4.2.1. Comparison sample

In this section, we compare the properties of the HIPPARCOS stars based on the proper-motion anomaly (PMa) approach (Kervella et al. 2022; see also Brandt 2021) and the NSS analysis. The PMa approach is described in detail by Kervella et al. (2019a). The general principle is to look for a difference in proper motion (PM) between the long-term PM computed from the HIPPARCOS (epoch 1991.25; van Leeuwen 2007a, see also ESA 1997) and *Gaia* DR3 (2016.0; Gaia Collaboration 2021a) astrometric (α, δ) positions on the one hand and the individual short-term PM vector from the *Gaia* DR3 catalogue on the other. For a single star, the long-term PM is identical to the short-term PM measured by *Gaia*, as its space velocity is constant with time. For a binary star, the short-term PM includes in addition the tangential component of the orbital velocity of its photocentre. As the latter is changing with time over the orbital period of the system, a deviation appears between the short-term and long-term PMs of the star which is due to the curvature of its sky trajectory. The PMa, namely, the difference between the short-term and long-term PM, is therefore an efficient and sensitive indicator of non-single stars.

[10] We retrieved `http://www.astro.gsu.edu/wds/orb6/orb6orbits.sql` on 2022-02-11 and did not remove orbits with two independent solutions or apply any other filters.

In order to compare the NSS catalogue with the PMa approach, we cross-matched the NSS catalogue with the PMa catalogue[11] of Kervella et al. (2022), which covers 116 343 HIPPARCOS stars. This resulted in a list of 2767 common targets with astrometric NSS `Acceleration7` or `Acceleration9` solutions and 5416 stars with `Orbital`, `AstroSpectroSB1`, or `OrbitalTargetedSearch*` orbital solutions. In addition, 4385 HIPPARCOS targets are listed in the NSS tables with `EclipsingBinary` (photometric), `SB1` or `SB2` (radial velocity) solutions. Overall, 12 568 HIPPARCOS/PMa stars have an entry in the NSS catalogue, that is, 10.8% of the HIPPARCOS/PMa catalogue.

4.2.2. Completeness of the NSS sample for HIPPARCOS stars

The *Gaia* stars that are present in the NSS catalogue were selected based on criteria on parameters from their single-star solutions tailored to identify the most probable binaries. For the astrometric solutions based on astrometry, this includes the presence of a `ruwe` higher than 1.4 in their single-star solution. As pointed out by Belokurov et al. (2020) and Stassun & Torres (2021), this criterion is efficient at identifying the stars that host partially resolved companions. Furthermore, based on the PMa analysis, the binary fraction was found to remain high for `ruwe` values lower than 1.4 by Kervella et al. (2022) with for example 30% of the stars with `ruwe` ≈ 1.2 exhibiting a PMa $S/N > 3$ (their Fig. 11). As a consequence, the degree of completeness of the star sample present in the NSS is likely relatively low because of its selection threshold on the `ruwe` value. To estimate the completeness of the NSS for the HIPPARCOS stars, we first applied to the PMa catalogue the same selection criteria as the NSS input sources (Sect. 2.2.1) – except the condition `ruwe` > 1.4 – resulting in a subsample of 92 240 stars (79.3%). Within this subsample, 28 111 stars are high-probability astrometric binaries as their PMa $S/N > 3$. Restricting the count to the NSS stars that have an astrometric solution (`Acceleration7`, `Acceleration9`, `Orbital`, `AstroSpectroSB1` or `OrbitalTargetedSearch*`), we obtain a completeness level of the NSS catalogue relative to the PMa catalogue of 8183/28 111 = 29.1%.

However, this high-level estimate based on global target numbers does not directly reflect the actual efficiency of the NSS reduction in detecting that a star is a binary or not compared to the PMa technique. To estimate this efficiency, we consider the same initial sample – following the NSS selection criteria including `ruwe` > 1.4 – and we derive the fraction of stars with an NSS solution within this common sample. The results are listed in Table 2. Overall, the astrometric solutions provided in the NSS catalogue represent 41% of the potential binaries present in the NSS reference sample, compared to 92% for the PMa catalogue.

In summary, because of the stringent selection of the solutions for the NSS, the catalogue comprises approximately 40% of the binaries from the HIPPARCOS–*Gaia* PMa catalogue that were potentially detectable from *Gaia* astrometry alone.

4.2.3. Statistics of the proper motion anomaly of NSS targets

The PMa is an efficient tracer of the presence of a massive orbiting companion, but its sensitivity is limited by two factors. Firstly, the time baseline between HIPPARCOS and *Gaia* (24.75 years), although long, significantly reduces the PMa signature of companions with orbital periods longer than approximately three times the HIPPARCOS–*Gaia* time, that is, 75 yr. Secondly, the fact that the *Gaia* DR3 proper motions are the result of an averaging over a time window of 34 months strongly smears out the signature of companions with orbital periods shorter than approximately 4 yr. In summary, the PMa technique is most sensitive for companions with orbital periods of between ≈4 and 75 yr. On the other hand, the capacity to determine orbital solutions directly from *Gaia* astrometry (or radial velocity) time series is significantly higher for binaries with periods of shorter than the *Gaia* DR3 measurement window. The longer periods remain detectable, mostly up to about twice the measurement window. However, the astrometric displacement of long-period binaries is generally detected only as an acceleration, that is, without a period determination.

[11] Available through the CDS/VizieR service as catalogue `J/A+A/657/A7/tablea1`.

Table 2. Comparison of the PMa and NSS astrometric detection rate on the common HIPPARCOS star sample.

	Number	Fraction
Objects eligible to NSS & PMa	14 748	100.0%
PMa $S/N < 3$ and absent from NSS	2254	15.3%
PMa $S/N > 3$ and absent from NSS	7320	49.6%
PMa $S/N < 3$ and present in NSS	950	6.4%
PMa $S/N > 3$ and present in NSS	4224	28.6%
Total non-single stars (PMa or NSS)	12 494	100.0%
Non-single stars detected from PMa	11 544	92.4%
Non-single stars present in NSS	5174	41.4%

Figure 14 shows the histograms of the number of NSSs with different kinds of solutions as a function of their PMa S/N. The five histograms that are colour coded in blue correspond to NSS solutions that include the *Gaia* DR3 astrometry either exclusively (`Acceleration7`, `Acceleration9`, `Orbital`) or in conjunction with spectroscopic radial velocities (`AstroSpectroSB1`) or previously known substellar orbital parameters (`OrbitalTargetedSearch*`). The eclipsing binary stars (`EclipsingBinary`; green colour) are characterised from the *Gaia* photometric data, and the spectroscopic binaries (`SB1`, `SB2`; yellow colour) rely on the spectroscopic radial velocities measured by the *Gaia* RVS (Cropper et al. 2018; Katz et al. 2019).

4.2.4. Orbital periods and sensitivity

Almost all the HIPPARCOS targets with an `Acceleration7` or `Acceleration9` solution show a significant PMa signal. This behaviour is expected for two reasons: (1) The NSS astrometric solutions were selected among the *Gaia* sources with a `ruwe` larger than 1.4. This favours partially resolved binary stars, which often have orbital periods within the sensitivity range of the PMa technique. (2) For orbital periods longer than the *Gaia* measurement window, the PMa and the acceleration are physically similar quantities, both related to the curvature of the sky trajectory of the star.

The NSS catalogue stars with `Orbital` or `AstroSpectroSB1` solutions generally have shorter orbital periods than the *Gaia* DR3 time window. Because of the time smearing of the *Gaia* EDR3 proper motions, this usually prevents the production of a clear signature in PMa. Nevertheless, approximately two-thirds of the stars of these NSS classes exhibit a significant PMa signal with $S/N > 3$ (Fig. 14). As shown in Fig. 15, the

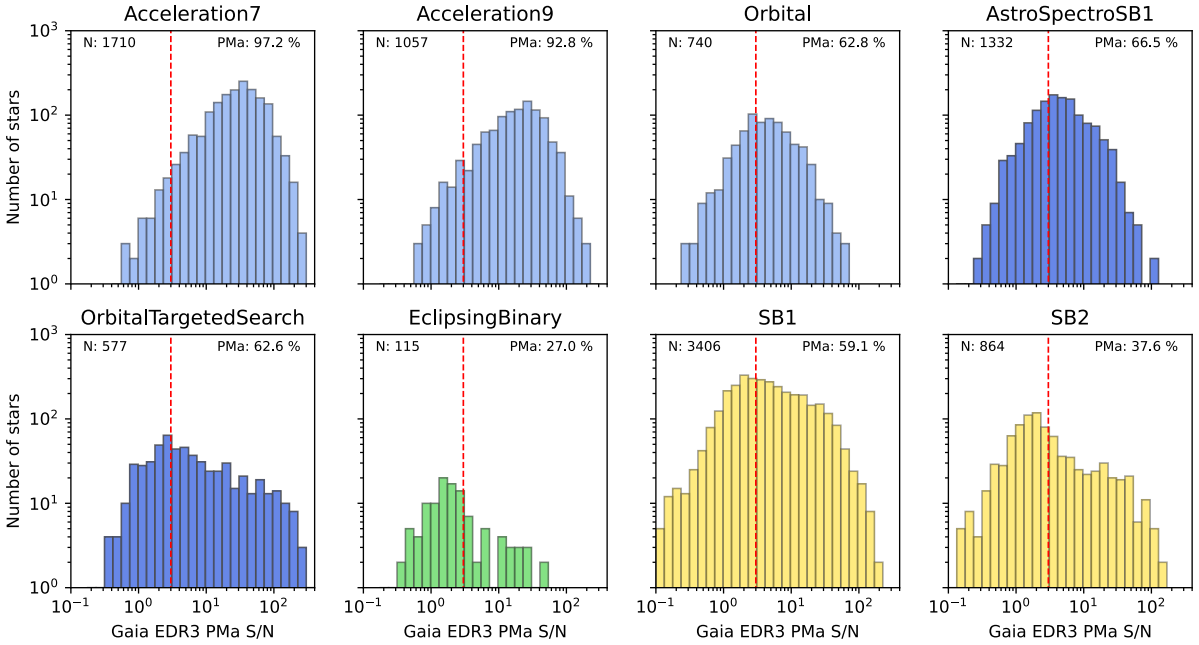

Fig. 14. Histogram of the number of NSS stars with different solution types that are present in the HIPPARCOS catalogue, as a function of the S/N of their *Gaia* DR3 proper motion anomaly from Kervella et al. (2022). The total number of targets N and the fraction of stars with a PMa S/N larger than 3 is displayed in each panel.

longer *Gaia* EDR3 time window compared to the DR2 results in a decrease of the PMa S/N for the binaries whose orbital period is shorter than ≈1000 days. This is caused by the stronger time smearing of the astrometric signal by the integration window in the *Gaia* DR3 compared to that of the DR2, which is not compensated by the increase in measurement accuracy in the *Gaia* DR3. The systems with shorter orbital periods than the *Gaia* integration window exhibit a median PMa $S/N \approx 3$. This indicates that despite the smearing, statistically, the mean *Gaia* PM vector still contains a significant signature of the binarity. The vast majority of *Gaia* NSS targets with orbital periods longer than the *Gaia* time window (both for the DR2 and EDR3) exhibit a significant PMa $S/N > 3$.

4.2.5. Long-term HIPPARCOS–*Gaia* proper motion

Here we compare the long-term proper motion deduced from the difference in position between the HIPPARCOS (1991.25) and *Gaia* DR3 (2016.0) epochs by Kervella et al. (2022; hereafter μ_{HG}) with the short-term proper motion as determined in the NSS catalogue (μ_{NSS}). Figure 16 shows the observed differences $\Delta\mu = \mu_{NSS} - \mu_{HG}$ between these two quantities for the HIPPARCOS catalogue stars with either accelerations (`Acceleration7`, `Acceleration9`) or orbital (`Orbital`, `AstroSpectroSB1`) solutions in the NSS. There is a significantly larger divergence of the long-term proper motions between the stars with NSS accelerations only for which $\sigma(\Delta\mu) \approx 2.6\,\mathrm{mas\,a^{-1}}$ than for the stars with an orbital solution for which $\sigma(\Delta\mu) \approx 0.1\,\mathrm{mas\,a^{-1}}$. The relatively poor agreement for the NSS acceleration stars may be explained by the fact that the measurement of the curvature of the sky trajectory is significantly easier with the longer HIPPARCOS–*Gaia* temporal baseline. For the full NSS orbital solutions, the agreement between the HIPPARCOS–*Gaia* PM and the NSS PM is remarkably good, demonstrating that the orbital fit procedure does not introduce systematic biases on the estimation of the mean PM value.

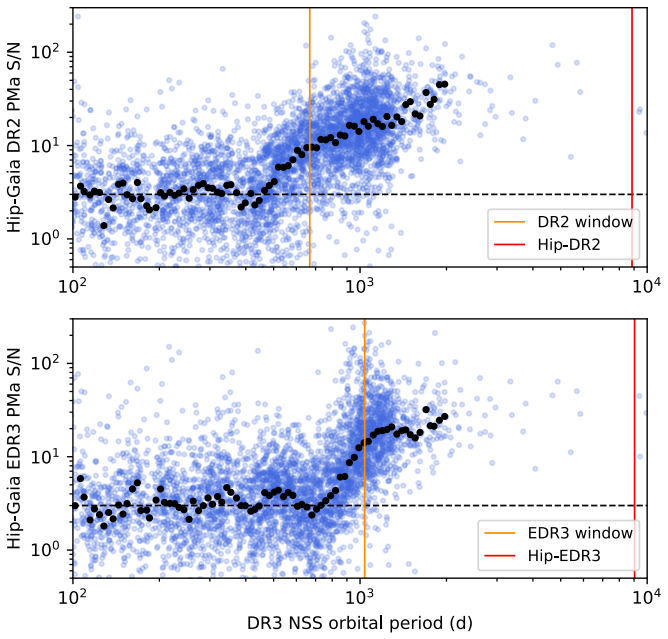

Fig. 15. Proper motion anomaly S/N as a function of the NSS catalogue orbital period for the DR2 PMa (*top panel*) from Kervella et al. (2019a); and the EDR3 PMa (*bottom panel*) from Kervella et al. (2022). The horizontal dashed line indicates the PMa $S/N = 3$ significance limit.

5. Catalogue of masses

As the `nss_two_body_orbit` table only gives access to the orbital parameters, it was found desirable to provide an estimate of the masses, the flux ratio, or the lower and upper limits of these, wherever possible. Here, we describe the construction and content of the table `binary_masses` which is made available in the *Gaia* archive.

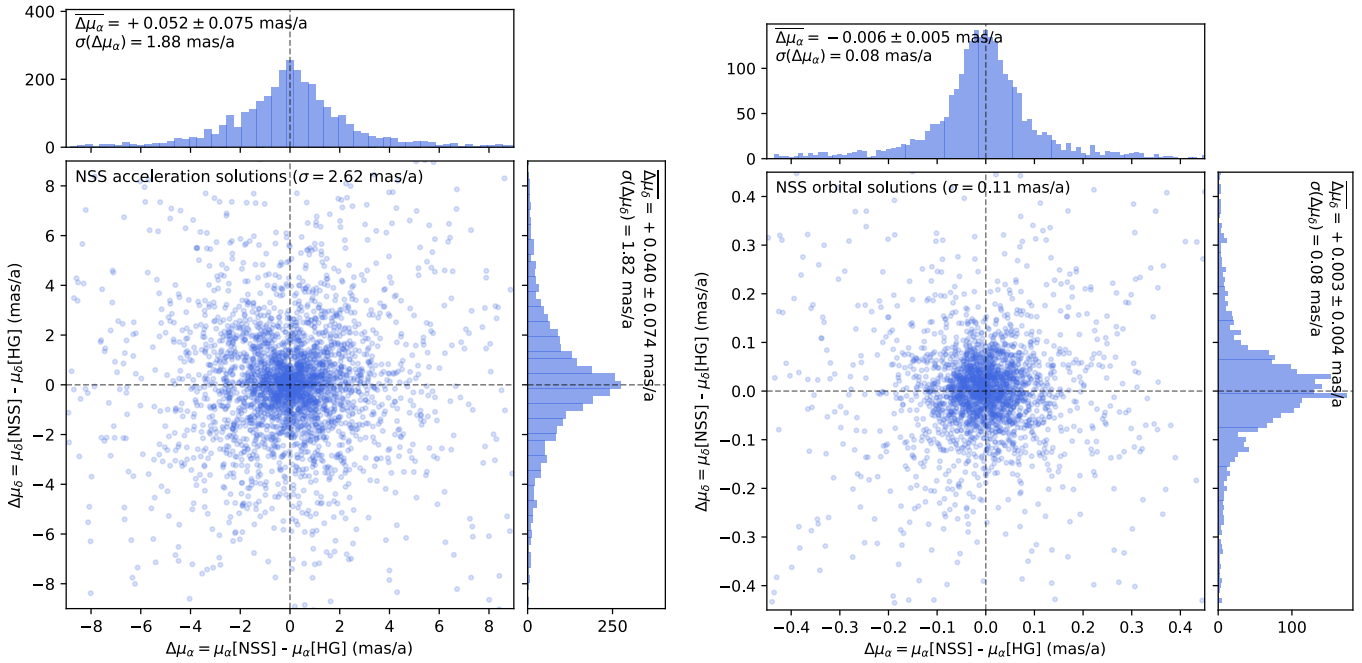

Fig. 16. Comparison of the long-term proper motions determined from the HIPPARCOS and *Gaia* DR3 positions $\mu_{\rm HG}$ with the *Gaia* DR3 proper motions $\mu_{\rm NSS}$ for NSS stars with acceleration solutions (*left panel*) and orbital solutions (*right panel*). We highlight the different scales.

5.1. Computation of the masses

The astrometric binaries give access to an astrometric mass function which depends on the flux ratio (F_2/F_1) of the components:

$$(\mathcal{M}_1 + \mathcal{M}_2)\left(\frac{\mathcal{M}_2}{\mathcal{M}_1 + \mathcal{M}_2} - \frac{F_2/F_1}{1 + F_2/F_1}\right)^3 = \frac{(a_0/\varpi)^3}{(P/365.25)^2}, \quad (2)$$

while the spectroscopic binaries provide a spectroscopic mass function which also depends on the inclination:

$$f(\mathcal{M}) = \frac{\mathcal{M}_2^3 \sin^3 i}{(\mathcal{M}_1 + \mathcal{M}_2)^2} = 1.0385 \times 10^{-7} K_1^3 (1 - e^2)^{3/2} P, \quad (3)$$

with P the period in days and K_1 the semi-amplitude of the primary in km s^{-1}. For `AstroSpectroSB1`, we have access to the Thiele Innes coefficients instead of K_1 which leads to the equivalent formula:

$$\frac{\mathcal{M}_2^3 \sin^3 i}{(\mathcal{M}_1 + \mathcal{M}_2)^2} = \frac{(C^2 + H^2)^{3/2}}{(P/365.25)^2}. \quad (4)$$

The inclination can be provided by an astrometric solution or an eclipsing one. Without the inclination, Eq. (3) only leads to minimum mass function information. When a `SB2` solution is provided, we have access to the mass ratio $\mathcal{M}_2/\mathcal{M}_1 = K_1/K_2$. When a system has a `SB2` solution and either an astrometric solution or an eclipsing one, the primary mass can be derived directly from the binary orbital parameters.

Two estimates of $\mathcal{M}_1$ are provided in the *Gaia* DR3 by the `FLAME` module (Creevey et al. 2023): `mass_flame`, based on GSP-Phot parameters and available in the `astrophysical_parameters` table, and `mass_flame_spec`, based on GSP-Spec parameters and available in the `astrophysical_parameters_supp` table. However the `FLAME` masses have three main limitations for our NSS sample: they are based on the parallax from the main catalogue while we now have a more accurate estimate for all astrometric solutions; these also assume a null flux ratio, which we know is not appropriate for a significant fraction of the NSS solutions, in particular the SB2 ones; these are not available for stars with masses smaller than 0.5 $\mathcal{M}_\odot$. We therefore implemented a specific code to derive the mass of the primary that allows us to manipulate the luminosity ratio and is described in detail in Appendix E. These masses are only derived for stars on the main sequence – as estimations for evolved stars are degenerate (e.g. Creevey et al. 2023) – and are referred to hereafter as 'IsocLum' masses. For white dwarfs, we simply assumed a fixed mass of 0.65 ± 0.16 $\mathcal{M}_\odot$ (Giammichele et al. 2012).

The uncertainties on the masses and flux ratios obtained are derived using a Monte Carlo simulation of 1000 points. We take into account the covariance matrix of the orbital parameters. For a_0 as well as for the `AstroSpectroSB1` spectroscopic part $a_1 = \sqrt{C^2 + H^2}/\sin i$, we use a Gaussian distribution with a local linearisation error estimation as Monte Carlo techniques are not adapted to the Thiele Innes coefficients (see Babusiaux et al. 2023). Only sources with a `significance` > 5 are present in NSS solutions, meaning that a Gaussian distribution of the semi-major axis errors is a reasonable assumption. The uncertainties for the SB2 and eclipsing solutions have been re-scaled according to their goodness-of-fit. For the IsocLum masses, we use the full mass distribution function as we have it available. We then compute the 16th and 84th quantiles (corresponding to ±1σ) of the derived parameter distributions to estimate the lower and upper values, respectively. The direct values are provided whenever applicable for `m1`, `m2`, and `fluxratio`.

When combining two NSS solutions, we only use those with periods and eccentricities compatible within 5σ, assuming an uncertainty of 0.1 on the eccentricity for sources with a fixed eccentricity. A weighted mean of the periods and eccentricities of both solutions is then used in Eqs. (2) and (3). For the combination of astrometric and spectroscopic solutions, the primary mass is tested for different flux ratios in steps of 0.01; for each of these, the secondary mass is then derived from Eq. (3) and the flux ratio from Eq. (2), and only solutions that are consistent with

Table 3. Content summary of the catalogue of masses.

Combination method	Number	$\mathcal{M}_1$	$\mathcal{M}_2$	F_2/F_1
Orbital+SB2	23	✓	✓	✓
EclipsingSpectro(SB2)	3	✓	✓	
Eclipsing+SB2	53	✓	✓	
AstroSpectroSB1+M1	17 578		✓	✓
Orbital+SB1+M1	1513		✓	✓
EclipsingSpectro+M1	71		✓	
Eclipsing+SB1+M1	155		✓	
SB2+M1	3856		✓	
Orbital+M1	111 792		Lower/upper	
SB1+M1	60 271		Lower	

Notes. The full table is available in the *Gaia* table `binary_masses`.

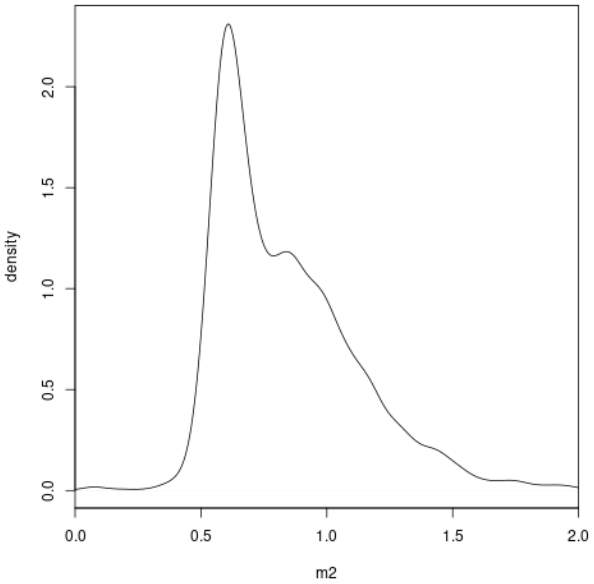

Fig. 17. Distribution of the secondary mass of astrometric solutions with `fluxratio_upper` = 0 in Table 3.

the tested flux ratio are kept. When no solution is consistent, the closest one is used.

For `Orbital` solutions, only upper and lower values can be derived as the flux ratio is not known. Different flux ratios are tested in steps of 0.01. The lower (respectively upper) secondary mass value is computed from the mass distribution obtained with the lower flux ratio (respectively upper). The solution with `fluxratio` = 0 is always kept, as soon as the primary star magnitude is compatible with the isochrones. For the other flux ratios tested, the secondary mass derived is accepted if consistent with a secondary star on the main sequence. In practice, the Monte Carlo masses of the secondary lead to a range of possible absolute magnitudes from the isochrones, which, for the flux ratio tested, are converted into an absolute magnitude of the system which is accepted when it is at less than 3 sigma from the absolute magnitude of the system measured by *Gaia*. In some cases, this means that no flux ratio is kept. These can be either pre-main sequence stars, in which case our masses are invalid, or triple systems with a primary, which needs a flux ratio, but a secondary mass not consistent with it. To handle this second option, the minimum flux ratio compatible with a primary on the main sequence is used to derive the primary mass but the secondary mass is derived with $F_2/F_1 = 0$. These cases can be isolated with a '`fluxratio_upper is NULL`' query. No limit is tested on the maximum flux ratio for white dwarfs. For SB1 solutions, the lowest valid flux ratio is used to derive the primary mass and the lower secondary mass value is derived on the distribution assuming $\sin i = 1$. For eclipsing binaries, the flux ratio is fixed by `g_luminosity_ratio`.

The catalogue of masses we derive is available in the *Gaia* Archive table `binary_masses`. A summary of the number of dynamical masses and flux ratio is presented in Table 3. We selected only sources with a S/N higher than 5 on the astrometric semi-major axis and on the spectroscopic primary semi-amplitude, as well as a S/N of higher than 2 for the eclipsing binary and astrometric $\sin i$ and the spectroscopic secondary semi-amplitude. For `AstroSpectroSB1` solutions, we verify that both the S/N for the spectroscopic part, computed as $\sqrt{C^2 + H^2}$, and for a_1 are higher than 5. If not, the `AstroSpectroSB1` is treated as an `Orbital` solution only. `OrbitalAlternative` solutions with $\log_{10}$(`parallax/parallax_error`) < 3.7−1.1 $\log_{10}$(`period`) have been excluded. There are 76 sources duplicated, having both an astrometric solution and either an eclipsing binary (6) or a `SB2` solution (70), with the astrometric solution period being larger than the other one by more than 10 sigma. For sources with both an `SB1` and an `Orbital` solution, only the `Orbital` solution has been kept.

A particularly interesting subset of this table are the astrometric solutions with `fluxratio_upper` = 0. There is only one star (*Gaia* DR3 4288765058313593856) for which the secondary mass is sufficiently small (0.57 $\mathcal{M}_\odot$) compared to the primary (1.26 $\mathcal{M}_\odot$), meaning that `fluxratio_upper` = 0 is compatible with the secondary star being on the main sequence. The others are systems for which the secondary mass solutions for flux ratio > 0 did not have a mass compatible with any of the flux ratios tested. The secondary mass distribution of these is presented in Fig. 17. There are 12 stars with a secondary mass smaller than the minimum mass handled by the isochrones of 0.1 $\mathcal{M}_\odot$ and a low mass for the primary as well. Three other stars with low-mass secondaries could be either triple or pre-main sequence stars for which the primary mass is not correct. The most predominant peak is that of the white dwarfs at $\mathcal{M}_2 = 0.61\,\mathcal{M}_\odot$ which has a standard deviation of 0.07 $\mathcal{M}_\odot$. We note that some white dwarf companions should actually have a flux ratio > 0, such as *Gaia* DR3 6416572288572864512 which is an `AstroSpectroSB1` with a significant flux ratio; the primary mass has been estimated as that of a metal-poor star because of its location on the left of the main sequence; if it had been solved as an astrometric solution only, it would have had `fluxratio_upper` = 0 and an underestimated secondary mass. A long tail of high-mass secondaries is also present. These could be compact objects, but also triple stars for which the single primary star hypothesis was not valid (see Sect. 7.1), primary stars that started to evolve, or metal-poor giants for which the primary mass is not correct.

5.2. Masses using external data

In this section, we illustrate how other mass estimates can be obtained thanks to various kinds of combinations with external data.

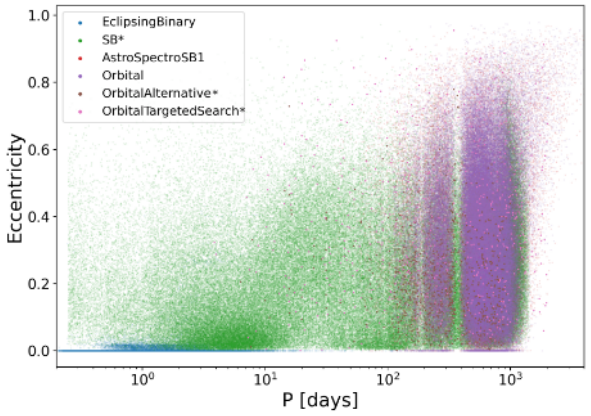

Fig. 18. Eccentricity vs. period for most solutions with orbits.

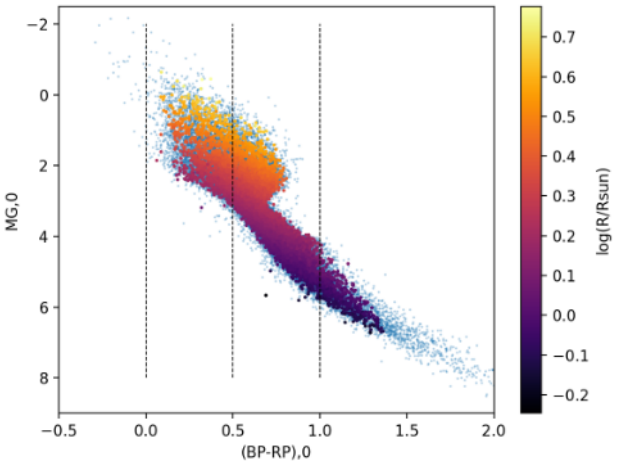

Fig. 19. Location in the dereddened HRD of the ($G_{BP,0}$−$G_{RP,0}$) bins used in Fig. 20. Small blue dots correspond to the SB1s not selected by our selection criteria. The radius is the `FLAME` estimate.

5.2.1. External SB2

Combining astrometric orbits with spectroscopic ones from large surveys is not a recently developed procedure, and was done for example with HIPPARCOS (Arenou et al. 2000). Once the inclination is known from the astrometric orbit, the semi-amplitudes from the spectroscopic orbit allow the masses and luminosities of the two components to be determined simultaneously. Recently, APOGEE DR17 data were used to detect 8105 SB2 or higher order systems (Kounkel et al. 2021). Once the required number of epochs become available, individual masses and magnitude differences will be obtained for the sources with an NSS `Orbital*` solution. Here, we simply take the example of *Gaia* DR3 702393458327135360, with $K_1 = 19.53 \pm 0.95\,\mathrm{km\,s^{-1}}$ and $K_2 = 21 \pm 1.1\,\mathrm{km\,s^{-1}}$. The masses are found to be $\mathcal{M}_1 = 1.14 \pm 0.38\,\mathcal{M}_\odot$ and $\mathcal{M}_2 = 1.06 \pm 0.35\,\mathcal{M}_\odot$ with a 0.567 ± 0.071 flux ratio.

5.2.2. Occultations

Occultations represent a method to constrain the sum of masses of binaries thanks to the measurement of their separation at a given epoch. We illustrate this with *Gaia* DR3 3162827836766605440 (HIP 36189) which is a $V \approx 6.5^m$ K giant that was discovered to be double thanks to an occultation by (704) Interamnia on 23 March 2003. Its acceleration had been detected in Kervella et al. (2019a, 2022), and Brandt (2021). Satō et al. (2014) indicate a $\rho = 12 \pm 3$ mas separation while a more precise indication is given by Herald et al. (2020): $\rho = 13.0 \pm 0.7$ mas with position angle $\theta = 231.9 \pm 4.0°$. Satō et al. (2014) evaluate the magnitude difference between components to about 1.5, to which we attribute a 0.2 mag uncertainty, accounting in particular for the observation in a band different from the G band. This source received an `Orbital` solution with a 2.6 yr period. From the combined information, the masses of the components are found to be $\mathcal{M}_1 = 3.9 \pm 2.2\,\mathcal{M}_\odot$ and $\mathcal{M}_2 = 3.5 \pm 1.6\,\mathcal{M}_\odot$.

5.2.3. One SB1 Cepheid

Although it is known that many Cepheids are in binary systems (e.g. Kervella et al. 2019b), not many orbits are present in the NSS catalogue. On the spectroscopic orbit side, the *Gaia* DR3 data processing does not yet include the simultaneous handling of orbit and Cepheid pulsations, meaning that it is only the latter that were detected. Consequently, these solutions were filtered out from the catalogue to avoid any confusion. On the astrometric orbit side, one Cepheid received an `Orbital` solution.

Gaia DR3 4703611143398494 72 = RX Cam is known to be a G2Ib+A0V spectroscopic binary from Imbert (1996). The *Gaia* solution has a period of 999 ± 104 d and an eccentricity of 0.514 ± 0.049, consistent at 1σ with respectively 1113.8 ± 0.5 d and 0.459 ± 0.007 from Groenewegen (2013). The inclination is $i = 113.^\circ5 \pm 1.^\circ7$. We may safely assume that there is no flux contribution (in the G band) from the companion, as confirmed by the difference between the semi-major axis of the primary and that of the photocentre $a_1 - a_0 = 0.04 \pm 0.12$ au. Using the semi-amplitude $K_1 = 14.27 \pm 0.11\,\mathrm{km\,s^{-1}}$ from Groenewegen (2013) and the estimation of the primary mass from Kervella et al. (2019b), $\mathcal{M}_1 = 5.40 \pm 0.81\,\mathcal{M}_\odot$, we obtain $\mathcal{M}_2 = 2.87 \pm 0.34\,\mathcal{M}_\odot$.

6. Special binaries in the HRD

In this section, we select several illustrative cases where the NSS catalogue can provide useful insights into populations of the HRD. For further reference, Fig. 18 presents the period–eccentricity diagram for the NSS solutions with orbits.

6.1. Spectroscopic binaries along the main sequence

In this section, we present and briefly discuss the eccentricity–period (hereafter e−P) diagrams of SB1s along the main sequence, defined as $-7.5 + 10\,(G_{BP,0}-G_{RP,0}) < M_{G,0}$ (Fig. 19), with the photometry being de-reddened in the same way as for the mass derivation (see Appendix E). Stars along the main sequence are divided according to ($G_{BP,0}$−$G_{RP,0}$) bins, as indicated in Fig. 19.

The e−P diagrams along the main sequence are displayed in Fig. 20. Because of the aliasing problems faced by the SB1 processing (see Sect. 2.3), these diagrams are shown for different filtering based on the significance of the RV semi-amplitude, namely K_1/σ_{K_1} larger than 10, 20, or 40 (from left to right). This filtering removes both high-eccentricity short-period orbits and long-period orbits. The former filtering is designed to remove possibly spurious solutions, while the disappearance of the long-period solutions is a side effect due to the fact that long periods have on average smaller K_1 and thus smaller significances as well. Nevertheless, this filtering has the consequence of revealing the shape expected for any e−P diagram, namely short-period orbits being

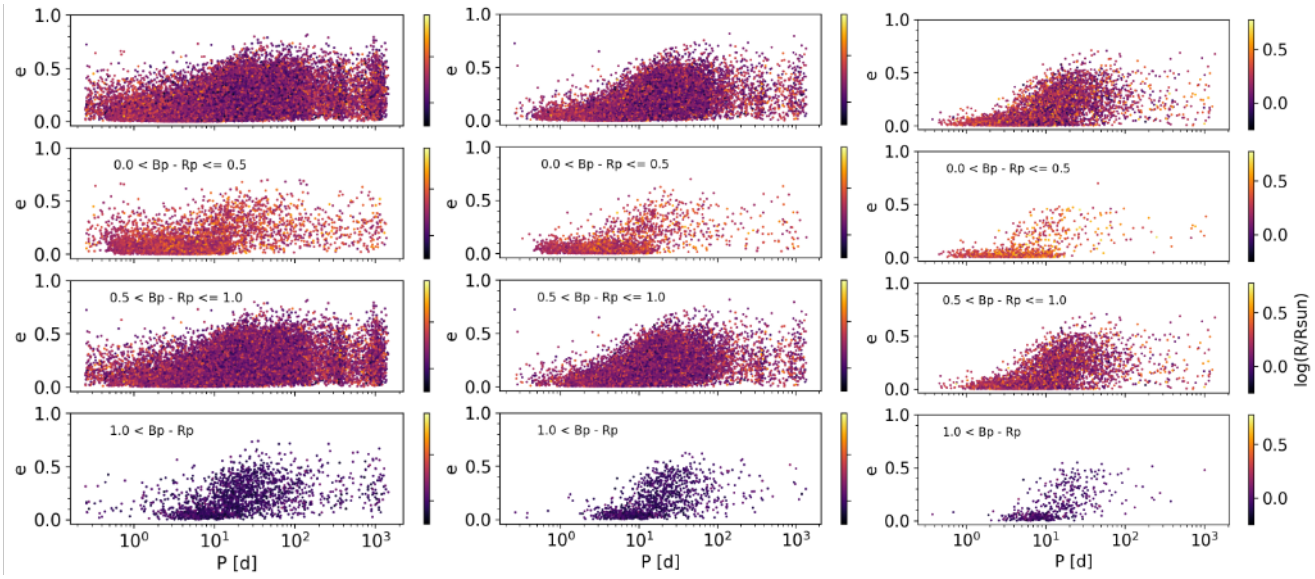

Fig. 20. The $e-P$ diagram for SB1s along the main sequence, filtered according to significance factors larger than 10, 20, or 40 (*from left to right*), and for different ($G_{BP,0}-G_{RP,0}$) spans (*top to bottom*). Filtering on the significance removes potentially spurious high-eccentricity solutions at small periods with the side effect of removing long period solutions. The drop in the number of systems at $P = 0.5$ yr due to insufficient sampling at this specific period. The color codes for the `FLAME` radius estimate.

almost exclusively circular below a given 'transition' period (see e.g. Mazeh 2008, for a detailed discussion).

The $e-P$ diagrams along the main sequence, when ordered according to bins of increasing $G_{BP,0}-G_{RP,0}$ and properly filtered on a significance larger than 40 (right panels of Fig. 20), reveal that this transition period does not vary strongly between the various $G_{BP,0}-G_{RP,0}$ bins, which is contrary to the situation prevailing along the giant branch, as discussed below (Sect. 6.2). Mazeh (2008) reviewed the processes shaping $e-P$ diagrams, and concluded that the constancy of the transition period along the main sequence would naturally result if the circularisation occurred during the pre-main sequence phase when the stars were large, following a suggestion by Zahn & Bouchet (1989) for F, G, and K stars from 0.25 to 1.25 $\mathcal{M}_\odot$. The transition period of these stars indeed remains constant along the main sequence at about 8 d as, although the transition period observed in Fig. 20 appears slightly shorter.

Mazeh (2008) also argues that short-period binaries (i.e. below the transition period) with non-circular orbits could result from a third distant companion pumping eccentricity into the binary orbit. However, at present, because of the confusion caused by possible period aliasing among short-period SB1 systems, this possibility cannot be tested with confidence.

6.2. Binaries along the RGB and AGB

The goal of this section is to show that the transition period between circular and non-circular orbits increases with radius and luminosity along the red giant branch (RGB) and asymptotic giant branch (AGB). To select stars on these branches, it is more efficient to use the 2MASS colour–magnitude diagram ($J-K$, M_K) rather than the usual *Gaia* colour–magnitude diagram. We used the 2MASS cross-match with EDR3 available within the data archive[12] and used the following criteria to select giants:

$$J-K > 0 \text{ and } M_K < 0, \tag{5}$$

as illustrated below.

6.2.1. Period–radius diagram

The existence of a circularisation threshold period in the $e-P$ diagram was demonstrated by Pourbaix et al. (2004) in their Fig. 4 (see also Zahn & Bouchet 1989; Verbunt & Phinney 1995; Mazeh 2008; Jorissen et al. 2009; Escorza et al. 2019). Its analytic form in a period–radius diagram may be easily obtained from the simple expression of the Roche radius R_R around the star of mass $\mathcal{M}_1$ with a companion of mass $\mathcal{M}_2$ (Paczyński 1971):

$$R_{R,1} = a\left(0.38 + 0.2\log\frac{\mathcal{M}_1}{\mathcal{M}_2}\right). \tag{6}$$

Substituting Kepler's third law, and assuming that the period threshold (expressed in days) corresponds to the situation where the star radius is equal to the Roche radius, one finds after some algebra:

$$\begin{aligned}\log(P_d^{\rm thresh}/365.25) &= (3/2)\ \log(R_1/216\,R_\odot)\\ &\quad -(1/2)\ \log(\mathcal{M}_1+\mathcal{M}_2)\\ &\quad -(3/2)\ \log\left(0.38+0.2\ \log\frac{\mathcal{M}_1}{\mathcal{M}_2}\right) \qquad (7)\\ &\equiv (3/2)\ \log(R_1/216\,R_\odot) \qquad (8)\\ &\quad -(3/2)\ c_1,\end{aligned}$$

where c_1 only depends on the masses. These thresholds are represented in Fig. 21 as dashed lines (corresponding to different choices for the pair $\mathcal{M}_1, \mathcal{M}_2$). Figure 21 presents all SB1 solutions falling along the RGB and AGB as defined above based on the ($J-K$, M_K) colour–magnitude diagram; however, it reveals that there are many SB1 solutions involving giant stars that do not fulfill the condition $P \geq P^{\rm thresh}$ expressed by Eqs. (7) and (8), especially when the significance K_1/σ_{K_1} is smaller than 40.

[12] https://gea.esac.esa.int/archive/documentation/GEDR3/Gaia_archive/chap_datamodel/sec_dm_crossmatches/ssec_dm_tmass_psc_xsc_best_neighbour.html

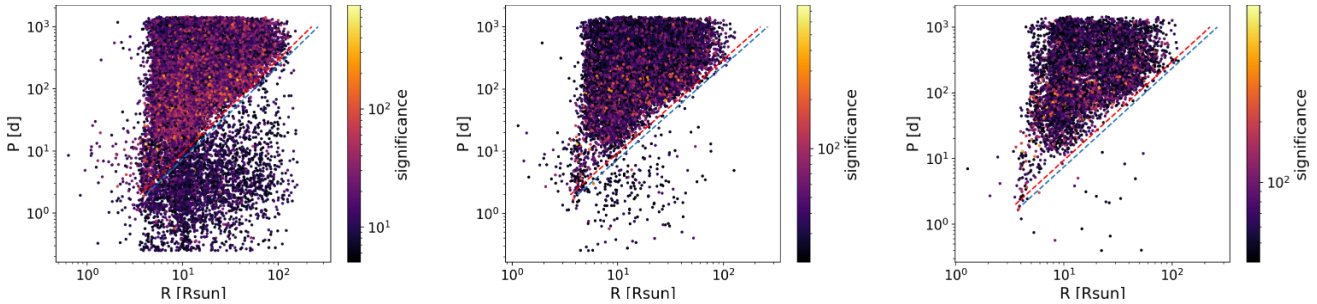

Fig. 21. Period–radius diagram for all SB1 solutions falling along the RGB and AGB according to criterion (5), and with a radius available from `radius_flame`. The dashed lines correspond to the threshold periods expressed by Eq. (7) for $\mathcal{M}_1 = 1.3\,\mathcal{M}_\odot$ and $\mathcal{M}_2 = 1.0\,\mathcal{M}_\odot$ (red dashed line) and $\mathcal{M}_1 = 1.3\,\mathcal{M}_\odot$ and $\mathcal{M}_2 = 0.2\,\mathcal{M}_\odot$ (cyan dashed line). *Left (a)*: unfiltered, 44 706 SB1 solutions (among which 3056 unphysical, that is, below the cyan dashed line). *Middle (b)*: filtered by significance $K_1/\sigma_{K_1} > 20$, 27 404 solutions are rejected and 17 302 are kept (among which 214 unphysical). *Right (c)*: filtered by significance >40, 37 850 solutions are rejected and 6856 are kept (among which 21 unphysical).

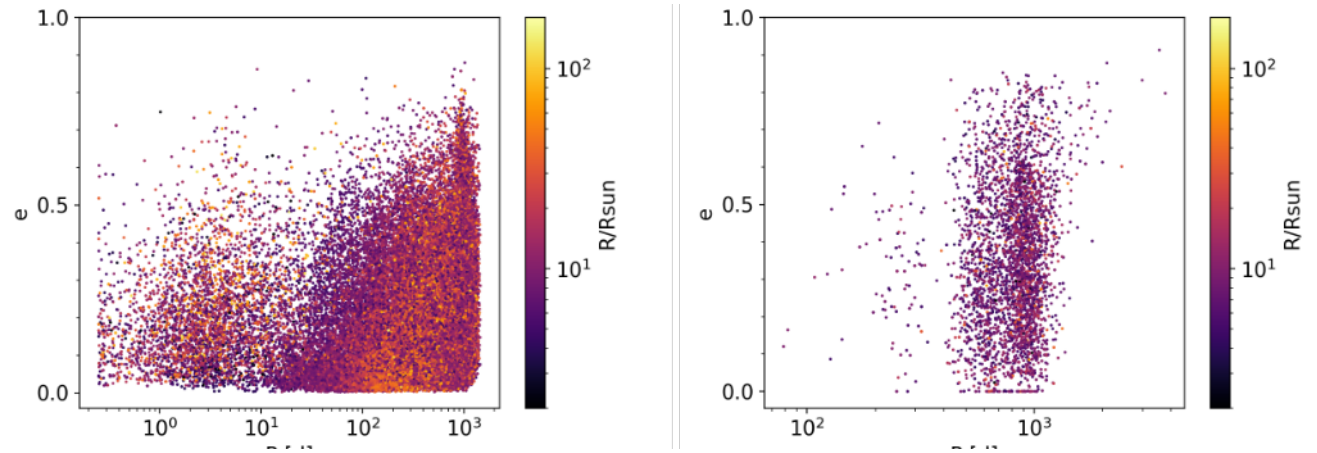

Fig. 22. $e-P$ diagrams for all SB1 (*left panel*: unfiltered; adequately filtered $e-P$ diagrams for SB1 with an RGB and AGB primary are presented in Fig. 24) and astrometric (*right panel*) solutions falling along the RGB and AGB according to criterion (5), and with a radius available from `radius_flame`. We highlight the restricted period scale of the astrometric binaries as compared to the SB1, and the lack of binaries with periods close to 1 yr among astrometric binaries.

6.2.2. $e-P$, $P-f(\mathcal{M})$ diagrams

Figure 22 presents the $e-P$ diagram for the same set of SB1 solutions (left panel) as shown in Fig. 21a, as compared to astrometric binaries along the RGB and AGB (right panel). The difference between the period range covered by SB1 and astrometric solutions is striking. As most astrometric orbits have periods longer than about 200 d, they clearly satisfy the criterion expressed by the dashed line in Fig. 21 and do not overfill their Roche lobe, contrary to the short-period SB1 solutions.

Figure 23a is similar to Fig. 22 but replacing eccentricities by mass functions, and revealing again two populations of SB1 solutions, the short-period ones being characterised by very small mass functions $f(\mathcal{M})$. The origin of this population of short-period SB1 solutions among RGB and AGB stars clearly requires clarification. In the following, we show that they are associated with poorly defined solutions. It appears indeed that almost all unphysical SB1 solutions may be eliminated by using the same purely observational criterion as used in Sect. 6.1, and based on the value of the significance of the SB1 solution (available in Table `nss_two_body_orbit` from the *Gaia* archive), that is, K_1/σ_{K_1}, K_1 being the semi-amplitude of the first component.

Almost no outlier remains in the $P-R$ (Fig. 21c) and $P-f(\mathcal{M})$ (Fig. 23c) diagrams when that significance exceeds 40; a few outliers remain when it exceeds 20 but many more solutions are kept, as may be seen from Table 4. This latter table shows that the gradual disappearance of unphysical SB1 solutions as the significance increases corresponds to a real filtering out of unphysical solutions, because the fraction of remaining unphysical solutions truly diminishes as the significance increases (passing from 6.8% in the absence of any filtering to 0.3% when the significance threshold is set at 40; see Table 4). However, the drawback of a filtering on significance is that it tends to also filter out solutions with long orbital periods, as those have on average smaller K_1 values (this was also very clear from Fig. 20). Alternatively, if one does not want to loose the long-period orbits as a result of filtering on significance, filtering is also possible with the physical condition $P \geq P^{\rm thresh}$ – Eq. (7) – with appropriate mass values; however Fig. 21 reveals that for systems with periods above 10 d, the boundary between physical and unphysical systems does not depend sensitively on the choice of $\mathcal{M}_1$, $\mathcal{M}_2$.

Now that the sample of RGB and AGB stars has been adequately cleaned of its unphysical orbits, it is possible to investigate the properties of the $e-P$ diagram for giant stars. Figure 24 presents those for bins of increasing radius (as taken from the

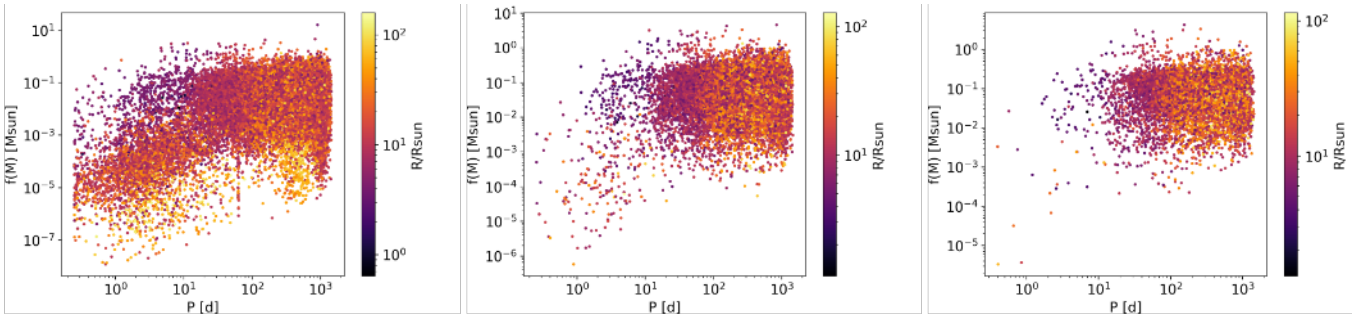

Fig. 23. $P-f(\mathcal{M})$ diagrams for SB1 along the RGB/AGB. *Left* (*a*): unfiltered. The yellow tail extending down to $f(\mathcal{M}) \sim 10^{-5}\,\mathcal{M}_\odot$ at periods in the range 300−800 d corresponds to pseudo-orbits associated to long-period pulsators (see Sect. 6.2.3). *Middle* (*b*): filtered by significance >20 and >40 (*right* (*c*)).

Table 4. Sizes of the SB1 sample involving RGB and AGB primaries for different filtering on the significance threshold.

Significance	Rejected	Accepted	Unphys.	(%)
All	0	44 706	3056	(6.8%)
>20	27 404	17 302	214	(1.2%)
>40	37 850	6856	21	(0.3%)

Notes. The numbers of rejected and kept sources are given as a function of the significance K_1/σ_{K_1}. The column labelled 'unphys.' lists the number of sources which would have a Roche filling factor of greater than unity (or $P < P^{\rm thresh}$ in Eq. (7), thus falling below the cyan dashed line on Fig. 21). The column labelled 'fraction' gives the ratio 'unphys.'/'accepted' (expressed as %).

corresponding `radius_flame` field), as indicated in the figure labels. As expected from the dashed lines in Fig. 21, the minimum period increases with increasing radius. In the following discussion, we adopt $\mathcal{M}_1 = 1.3\,\mathcal{M}_\odot$ and $\mathcal{M}_2 = 1.0\,\mathcal{M}_\odot$ (corresponding to the red dashed line in Fig. 21, and $c_1 = -0.274$ in Eq. (8)), as these values nicely match the observed trend. The above value for c_1 combined with the upper bound of the radius range adopted in each panel of Fig. 24 defines the lower bound on the orbital period $P_{\rm min}$ for $e = 0$. It appears that the upper envelope of the data cloud observed in each panel of Fig. 24 is well fitted by the empirical relation $P = P_{\rm min}(1-e)^3$ represented by the solid black lines in Fig. 24, as already found by Pourbaix et al. (2004) in their analysis of the Ninth Catalogue of Spectroscopic Binary Orbits (their Fig. 5). Despite the fact that this curve matches the uppermost data points in almost all panels rather well, it must be stressed that there seems to be no physical justification for this specific analytical form. However, a closer look at each of these panels reveals an interesting substructure. At the shortest periods, each panel is dominated by a large number of (nearly) circular orbits caused by circularisation operating in those systems where the giant stars with their convective envelope are close to filling their Roche lobe. As shown above, $P_{\rm min}$ in each panel actually refers to systems where the giants with the shortest radius in the considered range fill their Roche lobe (see e.g. Verbunt & Phinney 1995; Mazeh 2008, for details).

6.2.3. A search for genuine SB1s among giants

The identification of SB1 among highly evolved giants and long-period variables (LPVs) is made difficult by the envelope pulsation (Alvarez et al. 2001; Hinkle et al. 2002; Jorissen 2004; Jorissen et al. 2019). Other methods have therefore previously been used (Jorissen & Frankowski 2008; Sahai et al. 2008; Decin et al. 2020; Ortiz & Guerrero 2021) to identify true binaries. *Gaia* on the other hand, with its survey combining radial velocity and photometry data, offers exquisite prospects to disentangle pulsational from orbital RV variations. In that respect, the bottom panel of Fig. 24 offers an interesting benchmark sample of giant stars with $R > 80\,R_\odot$ with a SB1-like signature in their RVs. There are 40 such giants if the significance threshold

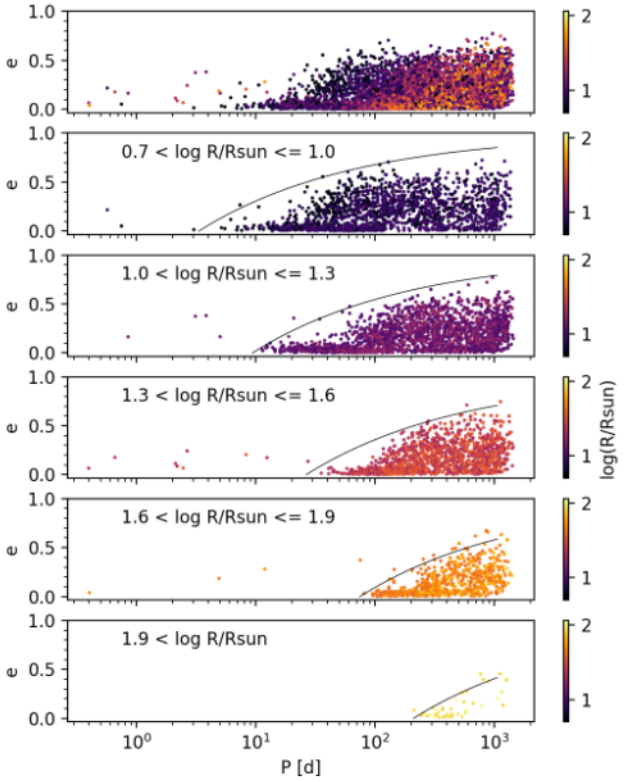

Fig. 24. $e-P$ diagram for SB1 systems along the giant branch, filtered according to significance factors larger than 40, for various radius spans. *Top panel*: is the full sample. The solid black lines correspond to the loci where $P(1-e)^3$ is constant (see text). The sample sizes are, *from top to bottom*, 1960, 2358, 1643, 737, and 40. The location in the HRD of the SB1 systems with $0.7 < \log(R/R_\odot) \leq 1.0$ (*second panel from top*), $1.3 < \log(R/R_\odot) \leq 1.6$ (*fourth panel from top*) and $1.9 < \log(R/R_\odot)$ (*bottom panel*) is shown in Fig. 25.

Table 5. Source id and basic parameters for SB1 and acceleration solutions with significance larger than 40 for giants with $R > 80\,R_\odot$ (*bottom panel* of Fig. 24).

Source_id	R ($R_\odot$)	Significance	ΔG (mag)	G (mag)	e	$P_{\rm nss}$ (d)	$P_{\rm lpv}$ (d)	Alt. id	Rem.
SB1									
1825471125022885760	111	62	0.0006	11.74	0.19	475			
1963830094814564992	98	71	0.0025	9.48	0.06	340	183		Ell. var.
1972501599433801856	104	82	0.0026	11.21	0.04	286	143		Ell. var.
1993611806061037824	91	45	0.0011	11.66	0.05	255			Ell. var.?
1996190371286907904	81	51	0.0005	11.65	0.10	368			
2022016864326072832	81	53	0.0002	12.25	0.13	1182			
203396083342181248	79	55	0.0004	11.48	0.28	851			
2072346498024572672	83	49	0.0006	11.67	0.40	805			
2153213619706962304	85	97	0.0008	8.92	0.02	395		BD+56 2152 [(a)]	
2179330422489474304	100	63	0.0007	11.27	0.40	1261			
2189793031540178560	109	46	0.0010	11.33	0.09	582			
2190661233409369728	81	179	0.0008	9.75	0.26	538			
2198983058969830272	91	148	0.0019	10.03	0.06	573			Ell. var.?
2203704946009240576	85	78	0.0007	11.05	0.15	663			
3023454391367052928	79	52	0.0023	11.52	0.02	513			Ell. var.?
3385138711262550144	107	48	0.0012	10.34	0.20	1063			
3441375569926160768	82	72	0.0004	8.58	0.45	1115		BD+26 935 [(b)]	
4309778580925549312	86	78	0.0024	11.75	0.08	212	106		Ell. var.
4538064682637397504	81	55	0.0005	10.47	0.02	440			
4657870428934536960 [(c)]	87	58	0.0107	11.71	0.09	366	375	V688 Cas [(d)]	Mira
4683286670959027200	82	109	0.0008	11.31	0.01	374			
4479122750503280512	100	64	0.0009	10.89	0.11	307			
519141188227099776	89	46	0.0004	11.04	0.02	420			
5340777165298298880	83	53	0.0014	11.57	0.02	273			Ell. var.?
5354875859285271936	89	50	0.0003	9.56	0.45	764			
5355550645876247808	95	49	0.0014	12.44	0.04	399			Ell. var.?
5405499126983935872	83	40	0.0018	13.07	0.03	243	122		Ell. var.
5406434021101010176	102	40	0.0010	10.28	0.01	319			
5604143357268838400	80	53	0.0007	9.97	0.04	327			
5697523299266655104	81	108	0.0007	10.69	0.28	585			
5697806664034217216	98	65	0.0009	10.59	0.37	889			
5796104824632537600	86	104	0.0018	10.12	0.17	556			
5806597567164955776	88	55	0.0018	10.94	0.09	433			Ell. var.?
5835609040544745344	108	40	0.0007	11.46	0.09	315			
5847195453486047616	105	56	0.0009	9.74	0.13	782			
5854104780242997504	92	49	0.0003	12.22	0.04	495			
5878621900997292800	92	114	0.0004	12.20	0.02	534			
5888442292197648000	115	40	0.0018	12.08	0.04	305			Ell. var.?
6012575363926683648	84	44	0.0014	12.43	0.04	439			Ell. var.?
6056302632126821760	84	43	0.0005	12.05	0.39	1261			
6057537697261452288	86	86	0.0004	11.02	0.01	343			
992206959423861888	94	43	0.0022	9.72	0.27	1152			
Acceleration solution									
6665511449204071424	92	28	0.0013	4.19	–	–	–	HD 190421	

Notes. The radius R is from FLAME. The column labelled ΔG gives a proxy of the variability in the G band (see text). The column labelled $P_{\rm lpv}$ lists the period obtained by the photometric analysis (table vari_long_period_variable), whereas $P_{\rm nss}$ corresponds to the period of the SB1 orbit. [(a)]Limb-darkened diameter of 0.393 mas from Cruzalèbes et al. (2019), or $96\,R_\odot$ with parallax $\varpi = 0.44$ mas. [(b)]Limb-darkened diameter of 0.672 mas from Cruzalèbes et al. (2019), or $106\,R_\odot$ with parallax $\varpi = 0.68$ mas. [(c)]Since $P_{\rm nss} = P_{\rm lpv}$, the RV variations are due to atmospheric pulsations rather than to orbital motion. [(d)]Also CGCS 396 in the General Catalog of Cool Galactic Carbon Stars, flagged as Mira variable.

is set at 40[13]. No Orbital or AstroSpectroSB1 solutions are present in *Gaia* DR3 among those giants with large radii (compare with Fig. 22). Table 5 lists their main properties, while their location in the HRD is shown as the yellow crosses in Fig. 25.

As mentioned above, as there is the risk that some of these SB1 solutions may be caused by envelope pulsation mistaken as SB1s, a proxy for the amplitude of the photometric variation

[13] Interested readers may set the significance threshold at 20 instead to get more SB1-like solutions (namely approximately 100), especially with long periods for the reasons explained in the text, or use instead the physical filtering $P \geq P^{\rm thresh}$.

in the G band has been listed as well, namely `phot_g_mean_mag_error` from the `gaia_source` table. We also performed a cross-match with table `vari_long_ period_variable` and the photometric period, whenever available, has been listed in the column $P_{\rm lpv}$. It appears that only one LPV is present in this list (a carbon star also known as V688 Cas), as confirmed by its largest ΔG value in Table 5. As $P_{\rm lpv} = P_{\rm nss}$ for this star, the RV variations are not due to orbital motion but to envelope pulsations. Many other such cases will be discussed below (Table 6). There are only four other stars appearing in the `vari_long_period_variable` table in common with Table 5, and these four have the unexpected property that $P_{\rm lpv} = 0.5\,P_{\rm nss}$, with a moderate ΔG value (on the order of 0.001–0.003 mag). We argue below, in relation to Table 6, that these are ellipsoidal variables and therefore true binaries, where the giant primary is close to filling its Roche lobe. Based on the fact that these ellipsoidal variables identified in table `vari_long_period_variable` have small eccentricities ($e < 0.1$), we suspect that Table 5 contains many more such ellipsoidal variables, namely those with $e < 0.1$ and $\Delta G > 0.001$ mag, flagged as 'Ell. var.?' in the last column of Table 5.

The longest period found in Table 5 is 1261 d, a value that is nicely in line with the *Gaia* DR3 time span, but short with respect to the periods expected among evolved giants (consider for instance the 17 yr period found by Jorissen et al. 2019 for the carbon Mira V Hya). Such long periods are not detectable at the current stage of the *Gaia* mission, either as SB1 or as astrometric binaries. Nevertheless, one may look for acceleration solutions (there is only one solution from table `nss_acceleration_astro` matching the giant criteria defining Table 5; regarding acceleration solutions, we refer to Sect. 4.2.3 and Wielen et al. 1999; Makarov & Kaplan 2005; Frankowski et al. 2007; Kervella et al. 2019a, 2022; Brandt 2021) or for solutions with a trend in the RVs (121 solutions for giants found in `nss_non_linearspectro`, not listed here).

6.2.4. Combining photometry and spectroscopy to diagnose RV variations in giants: pulsation, ellipsoidal variables, and rotational modulation

With the initial aim being to further investigate how many LPVs may be mistaken as SB1, we searched for targets in common between `SB1` from the NSS `nss_two_body_orbit` table and LPVs as provided by the variability study in the `vari_long_period_variable` table. However, this cross-match revealed some surprises. The following query yields 1869 entries, as shown in Fig. 26:

```
SELECT * from gaiadr3.nss_two_body_orbit TBO,
gaiadr3.vari_long_period_variable LPV where
LPV.source_id = TBO.source_id and
LPV.frequency is
not null
```

The three panels differ in terms of the level of filtering applied on the `SB1` significance parameter, as defined above: >5 (default, top panel), >20 (middle panel), and >40 (bottom panel). Striking are the two straight sequences observed in all three panels. The upper sequence corresponds to $P_{\rm lpv}/P_{\rm nss} = 0.5$ (as expected for ellipsoidal variables), whereas the lower sequence corresponds to $P_{\rm lpv}/P_{\rm nss} = 1$ (as expected for pulsating stars or rotational modulation in a synchronised system). The lower sequence is further made of two distinct clumps, one at short periods ($P_{\rm lpv} \lesssim 100$ d; which could be due to starspot modulation on a spin–orbit synchronised star) and the other at long periods ($200 \lesssim P_{\rm lpv} \lesssim 1000$ d; LPVs). These are discussed in turn below.

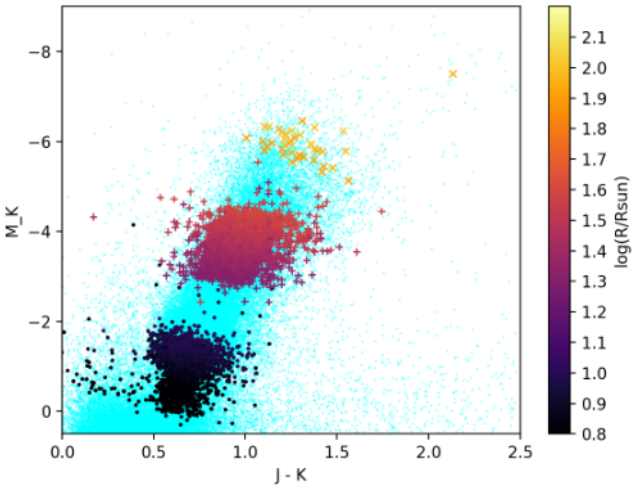

Fig. 25. Location in the HRD of three among the samples displayed in Fig. 24, namely $0.7 < \log(R/R_\odot) \leq 1.0$ (dots), $1.3 < \log(R/R_\odot) \leq 1.6$ (plus symbols), and $1.9 < \log(R/R_\odot)$ (crosses). Small cyan dots correspond to the `SB1` not selected by our selection criteria. See Table 5 for a full discussion of the properties of the yellow crosses.

Ellipsoidal variables

Besides the obvious property of their light-to-RV period ratio equal to 0.5, the ellipsoidal–variable sequence is further confirmed from its following properties: (i) small eccentricities ($e \lesssim 0.1$), (ii) large filling factors ($R_1/R_{\rm R,1} \gtrsim 0.65$ from Eq. (6), adopting radii from `FLAME` and the same typical masses as above – $\mathcal{M}_1 = 1.3\,\mathcal{M}_\odot$, $\mathcal{M}_2 = 1.0\,\mathcal{M}_\odot$) whenever available, and (iii) small G amplitudes ($0.01 \leq \Delta G \leq 0.1$ mag; see Table 6). We confirmed the ellipsoidal nature of these stars by a comparison between the light and RV curves. As expected, the maximum light indeed occurs at the quadratures, when the RV is maximum or minimum. In Fig. 24, at any given radius range, these ellipsoidal variables are located in the nearly circular tail of each panel. The full list of ellipsoidal variables is not provided here as the reader may easily obtain it from the ADQL query mentioned at the beginning of this section and filtering on $P_{\rm lpv}/P_{\rm nss}$ around 0.5 (370 stars in the inclusive range 0.45–0.55, most of them having significances in excess of 20). The first part of Table 6 nevertheless lists a few randomly selected examples. Figure 27 shows the position in the 2MASS infrared colour–magnitude diagram ($M_K, J-K$) of the 370 stars with $0.45 \leq P_{\rm lpv}/P_{\rm nss} \leq 0.55$ (dots). These ellipsoidal variables are located from the tip of the RGB[14] to 3 mag below. We note that some among these stars might be young pre-main sequence stars. *Gaia* DR3 2162167694508896128 = V1540 Cyg listed in Table 6 is one such case (on Fig. 27, it is located at $M_K = -6.2$ and $J-K = 1.41$).

Long-period variables

The transition between dots and crosses in Fig. 27 corresponds to the transition across the RGB tip. Above the RGB tip, most

[14] $M_{\rm K,RGB-tip} = -6.49$ as derived from Lebzelter et al. (2019) who find $K_{\rm RGB-tip} = 12$ for the LMC and considering its distance modulus 18.49± 0.09 mag (de Grijs et al. 2017).

Table 6. A few illustrative examples of ellipsoidal variables ($P_{\rm lpv}/P_{\rm nss}$ = 0.5) mistaken as LPVs in the `vari_long_period_variable` table, LPVs with a pseudo `SB1` orbit ($P_{\rm lpv}/P_{\rm nss}$ = 1, $\Delta G > 0.1$ mag, $P_{\rm lpv} > 180$ d) in the `nss_two_body_orbit` table and short-period ($P_{\rm lpv} < 100$ d) 'LPVs' with $P_{\rm lpv}/P_{\rm nss}$ = 1.

Gaia DR3 id	$P_{\rm lpv}$ (d)	$P_{\rm nss}$ (d)	$P_{\rm lpv}/P_{\rm nss}$	Signif.	K_1 (km s^{-1})	e	ΔG (mag)	$f(\mathcal{M})$ ($\mathcal{M}_\odot$)	$R/R_{\rm R}$	R ($R_\odot$)	Alt. id.
Ellipsoidal variables											
2162167694508896128 [a]	143 ± 16	289 ± 0.2	0.49	249	34.6	0.02	0.02	1.25			V1540 Cyg
5871624883899265280	48 ± 3	95.1 ± 0.1	0.50	101	22.0	0.03	0.06	0.105	0.92	43	
449088171382718848	178 ± 50	356.7 ± 1.5	0.50	73	12.8	0.01	0.09	0.078			
528840770565380352	54 ± 4	108.6 ± 0.1	0.50	89	31.7	0.01	0.06	0.358	0.88	45	
1837292073273265920	58 ± 3	116.1 ± 0.1	0.50	86	34.1	0.04	0.06	0.477			
4305358093199399168	42 ± 2	84.8 ± 0.1	0.50	89	27.3	0.03	0.05	0.178	0.74	32	
6653811713476525440	63 ± 4	125.4 ± 0.6	0.50	19	11.4	0.16	0.09	0.019	0.67	37	
5933194270923372288	10 ± 17	219.4 ± 0.4	0.50	63	21.8	0.05	0.05	0.234			
5998937575770407936	69 ± 9	137.1 ± 0.2	0.50	50	21.7	0.05	0.07	0.144	0.75	45	
...											
Large-amplitude ($\Delta G > 0.1$ mag) LPVs (Mira or SRa,b) with a pseudo SB1 orbit											
3029929312263388416	330 ± 39	329 ± 3	1.002	27	5.6	0.27	0.27	0.0057			
5861476288517412096	351 ± 29	350 ± 4	1.003	22	7.4	0.14	0.25	0.0144			
4498570706006456320	196 ± 11	196 ± 2	1.003	9	3.5	0.36	0.21	0.0008			
6635121600650977280	310 ± 42	309 ± 2	1.003	18	2.9	0.29	0.13	0.0007	–	75	
185224454669173120	491 ± 114	490 ± 9	1.003	9	2.6	0.34	0.29	0.0009			
5522324157261027968	321 ± 37	320 ± 3	1.003	23	5.3	0.29	0.26	0.0046			
5428546471231540608	455 ± 60	454 ± 6	1.003	7	2.8	0.33	0.24	0.0009			
463720476424410624	353 ± 24	352 ± 3	1.004	34	7.5	0.17	0.26	0.0152			
6358622017131465728	181 ± 7	180 ± 2	1.004	10	6.4	0.17	0.54	0.0048	–	138	
2180493018598279296	178 ± 25	177 ± 2	1.004	11	2.3	0.09	0.12	0.0002			
5318375436185802368	262 ± 32	260 ± 2	1.005	12	2.6	0.33	0.18	0.0004	–	97	
5522970154700635392	412 ± 37	410 ± 8	1.005	14	5.6	0.04	0.49	0.0076	–	84	
1989628623330891904	419 ± 18	[25 ± 0.02] [b]	[16.7] [b]	21	6.6	0.39	0.60	0.0006			
...											
Genuine binaries among LPVs ($P_{\rm lpv}/P_{\rm nss} \neq 0.5$ or 1; see text)											
5341773936978279296	220 ± 61	1252 ± 113	0.18	21	7.7	0.25	0.05	0.053	0.28	73	
5597415372601747456	168 ± 54	656 ± 6	0.26	43	11.1	0.36	0.03	0.076	0.33	56	
5414646307794529792	196 ± 42	753 ± 17	0.26	34	6.2	0.4	0.04	0.015			
5875470387113872768	214 ± 28	746 ± 16	0.29	29	15.3	0.35	0.09	0.229			
5347893273248921984	279 ± 75	913 ± 40	0.31	29	13.3	0.07	0.03	0.220			
5404683839108805248	215 ± 28	662 ± 1	0.32	89	11.5	0.33	0.11	0.088			
5796098502440628864	244 ± 28	701 ± 4	0.35	59	6.5	0.08	0.08	0.020			
1642955252784454144	374 ± 93	503 ± 6	0.74	33	6.0	0.28	0.04	0.010			
304717076269774336	310 ± 102	158 ± 1	1.96	22	12.0	0.01	0.10	0.028	0.39	26	
6661657003818388480	197 ± 92	42 ± 0.1	4.66	32	26.3	0.11	0.04	0.078	0.25	7	
187075684355571200	161 ± 64	26 ± 0.01	6.29	73	22.9	0.08	0.05	0.031			
5473442554645523712	209 ± 57	32 ± 0.02	6.62	30	36.4	0.06	0.04	0.157			
Short-period ($P_{\rm lpv} < 100$ d) light and RV variations with $P_{\rm lpv}/P_{\rm nss}$ = 1: starspot modulation or short-period pulsators?											
5498026500770376576	46 ± 1	45.7 ± 0.12	1.003	10	4.9	0.41	0.14	0.0005			
4498425604828703104	50 ± 4	49.57 ± 0.05	1.003	39	32.5	0.03	0.08	0.18	0.24	7	
3047643956417931264	64 ± 1	63.4 ± 0.1	1.003	13 [c]	8.3	0.21	0.24	0.0036	1.44 [c]	52	
5637220068643463424	57 ± 4	57.17 ± 0.03	1.004	66	33.6	0.04	0.06	0.2240	0.27	9	
5883587875302126592	78 ± 4	78.1 ± 0.1	1.004	16	12.2	0.42	0.08	0.0128			
5701792904776417152	43 ± 4	43.38 ± 0.01	1.005	85	25.2	0.02	0.06	0.0718			
1869696952997313664	53 ± 3	52.56 ± 0.05	1.005	31	22.5	0.13	0.08	0.0610	0.26	8	
5235000057883364864	43 ± 1	42.5 ± 0.03	1.005	30	35.6	0.10	0.06	0.1969	0.40	11	
...											

Notes. Radii are the `FLAME` DR3 estimates (see text for how the filling factors $R/R_{\rm R}$ were estimated). The column labelled ΔG lists the field `mad_mag_g_fov` (median absolute deviation) from table `vari_summary`. [a]This is likely a young star, of 'Orion-variable' type (V1540 Cyg). [b]This star has been added to the Mira category, despite its $P_{\rm lpv}/P_{\rm nss}$ ratio vastly different from 1, to illustrate some kind of aliasing problems (see text). [c]Doubtful case: significance is only 13 and filling factor is 1.4.

stars from the `vari_long_period_variable` table belong to the sequence $P_{\rm lpv}/P_{\rm nss}$ around unity. They correspond to LPVs with a RV variation caused by the envelope pulsation. Although displayed in Fig. 27, the full source list is not given here as they are easily obtained in a way similar to that discussed above for ellipsoidal variables. Table 6 nevertheless lists a few randomly selected examples.

These LPVs are easily identified by their velocity semi-amplitudes, which are smaller than 10 km s^{-1} (in that sense, they differ markedly from the ellipsoidal variables which generally have much larger K_1 values) and periods in excess of 180 d, as expected for Mira pulsations (Alvarez et al. 2001; Hinkle et al. 2002). Therefore, given these relatively small values of K_1, the significance of the SB1-like solution (namely K_1/σ_{K_1}) may in several cases be smaller than 20, but the identity of the NSS and LPV periods is in itself an indication of the reliability of the RV model. We note that the filling factor has no meaning in this stellar category, because there is no true orbit

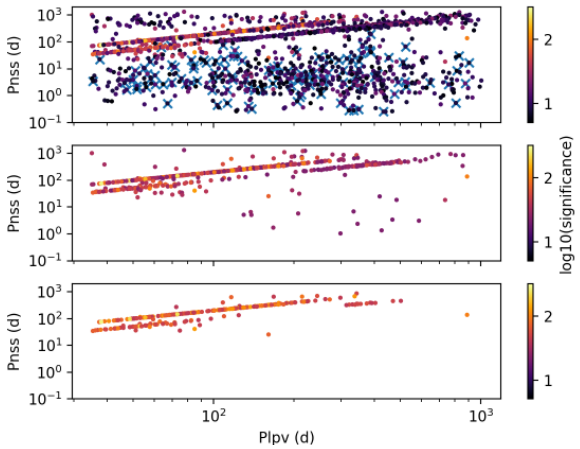

Fig. 26. Orbital period from the `nss_two_body_orbit` table vs. photometric period from the `vari_long_period_variable` table. *Top, middle*, and *bottom panels*: correspond to different filtering based on the SB1 significance (respectively, larger than 5, 20, and 40). The two sequences observed in all panels correspond to $P_{\rm lpv}/P_{\rm nss} = 0.5$ (ellipsoidal variables; *upper sequence*) and $P_{\rm lpv}/P_{\rm nss} = 1$ (LPVs or rotational modulation in a synchronised system; *lower sequence*). *Top panel*: crosses denote NSS solutions for which the Roche-lobe filling factor is above unity, and therefore unphysical. The filtering with significance larger than 40 makes these latter disappear almost entirely.

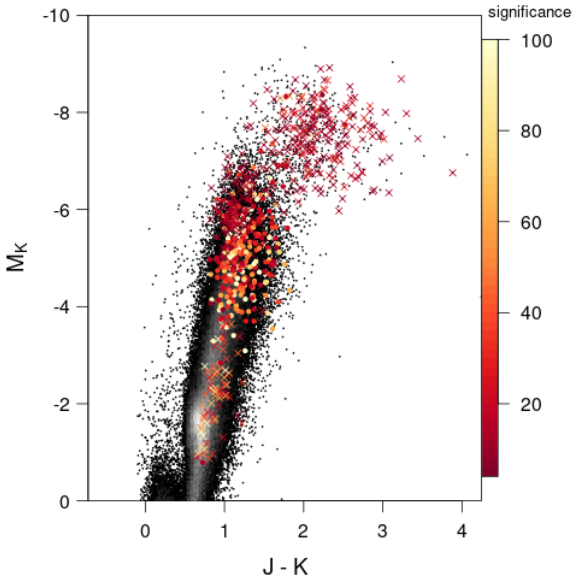

Fig. 27. Location in the infrared colour–magnitude diagram of the stars with $0.45 \leq P_{\rm lpv}/P_{\rm nss} \leq 0.55$ (ellipsoidal variables) from Table 6 (dots). Stars with $0.95 \leq P_{\rm lpv}/P_{\rm nss} \leq 1.05$ are represented by crosses; they appear in two different locations, among LPVs with low orbital significance on one hand, and among less luminous giants with much larger orbital significance on the other, perhaps suggesting starspot modulation or short-period pulsators. The dots ($P_{\rm lpv}/P_{\rm nss} \sim 0.5$) fall in between these two groups, as they are located just below the tip of the RGB.

associated. Mira variables are recognised as well by their large amplitude in G (>0.1 mag). The pseudo-eccentricities found by Hinkle et al. (2002) for Miras and semi-regular variables were clustered around 0.35, with a few cases below 0.1 as well. Here, the pseudo eccentricities range all the way from 0.09 to 0.48 (Table 6). Furthermore, it has been checked that the maximum RV is reached at phase 0.8 while maximum light is reached at phase 1.0, a phase lag that is expected for Mira pulsators. Furthermore, as for these stars, K_1 is relatively small and P is long, the pseudo mass functions are consequently smaller than $10^{-3}\,\mathcal{M}_\odot$, with some values as small as $10^{-5}\,\mathcal{M}_\odot$, in agreement with the findings of Hinkle et al. (2002; their Table 2) for Mira and semi-regular variables. LPVs with low values of $f(\mathcal{M})$ are most clearly seen in Fig. 23a as the yellow tail extending down to $f(\mathcal{M}) \sim 10^{-5}\,\mathcal{M}_\odot$ for periods between 300 d and 1000 d. On that same figure, many more stars with large radii ($R > 100\,R_\odot$) are found at shorter periods, but those are spurious 'SB1-like' solutions because their periods do not even match the LPV one, and when available, their filling factors are above unity, which is non-physical. Furthermore, their mass functions are quite small (down to $10^{-8}\,\mathcal{M}_\odot$). Therefore, these targets cannot be genuine binaries.

Genuine binaries

Genuine binaries among Miras are expected to have much longer orbital periods than currently detectable by *Gaia* DR3 (as is the case for instance for the carbon Mira V Hya quoted above). Intriguingly, several genuine SB1s have nevertheless been found among stars with $P_{\rm lpv} > 150$ d, a property generally associated to LPVs. In the fourth part of Table 6 are listed 14 SB1s selected among the 1189 SB1 solutions in common with the `vari_long_period_variable` table. These SB1s have a significance of greater than 20, the 1σ confidence range of $P_{\rm lpv}/P_{\rm nss}$ falling outside the ranges 0.45−0.55 and 0.9−1.1 (to avoid SB1-like variations caused by pulsations), $P_{\rm nss} > 20$ d and $P_{\rm lpv} > 150$ d. Their RV curves were checked visually and showed no peculiarity that would make the solution dubious. This visual inspection nevertheless revealed that some kind of aliasing problems remain with the NSS SB1 periods. The star *Gaia* DR3 1989628623330891904 was originally considered as a possible genuine binary among LPVs because $P_{\rm lpv} = 419$ d as compared to $P_{\rm nss} = 25$ d. However, the visual inspection of the RV curve revealed that the LPV period of 419 d is clearly present in the RV curve, although it was not selected by the period-selection algorithm, which gave no warning about a possible problem with that solution (significance = 21, period confidence = 1.000, ruwe = 1.09) except for the goodness-of-fit of 2.5. Therefore, that star has been added to the 'Large-amplitude LPVs' section.

Rotational modulation on a spin–orbit synchronised star

The final category of interest in Table 6 contains targets with short periods (i.e. $P \lesssim 100$ d) on the $P_{\rm lpv}/P_{\rm nss} = 1$ sequence. These are listed in the fourth part of Table 6 and identified in the HRD of Fig. 27 as the crosses at the bottom of the giant branch. Contrary to the situation prevailing for ellipsoidal variables and long-period variables discussed above, the phase lag between velocity and light curves now appears to be anything between 0 and π. For this reason, their light variation could be due to starspot modulation on a spin–orbit synchronised primary star (e.g. Mazeh 2008). A less likely possibility is that they could be small-amplitude pulsating stars.

6.3. Identifying EL CVn systems in Gaia data

EL CVn systems are short-period eclipsing binaries (EBs) consisting of an A/F-type main sequence (MS) primary and a low-mass pre-helium white dwarf (pre-He-WD) secondary. These systems are a result of mass transfer from the evolved pre-He-WD progenitor to the currently observed primary star (e.g.

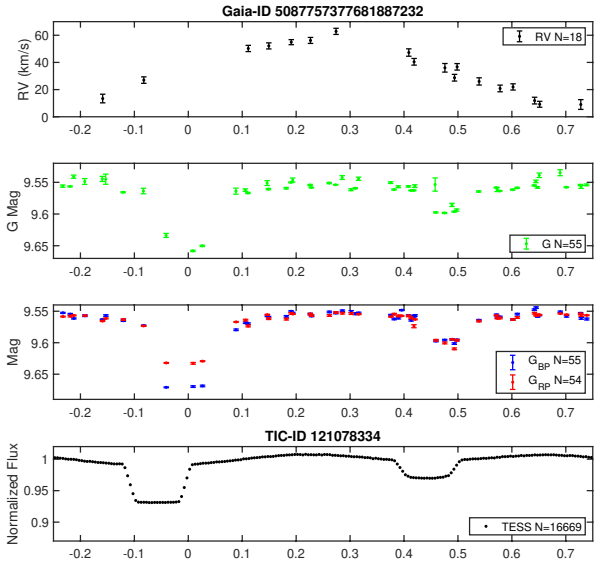

Fig. 28. Folded *Gaia G*, G_{BP}, and G_{RP} photometry and RV data of HD 23692, together with TESS binned data. *Top panel*: *Gaia* RV data, while the *second panel* presents the *Gaia G* data. *Third panel*: *Gaia* G_{BP} and G_{RP} data, with medians shifted to the *Gaia G* median for clarity, and the *bottom panel* presents the TESS data binned to 200 phase bins. All plots were folded using the *Gaia* EB-model period and deeper-eclipse epoch as the folding period and phase zero, respectively. We note that the primary eclipse is at phase 0.5, while the secondary eclipse is at phase *zero*. Observed TESS-eclipse phase drift is due to the more than 1300 day delay from the last *Gaia* point to the first TESS point. For clarity, the three *bottom panels* use the same *y*-axis scale.

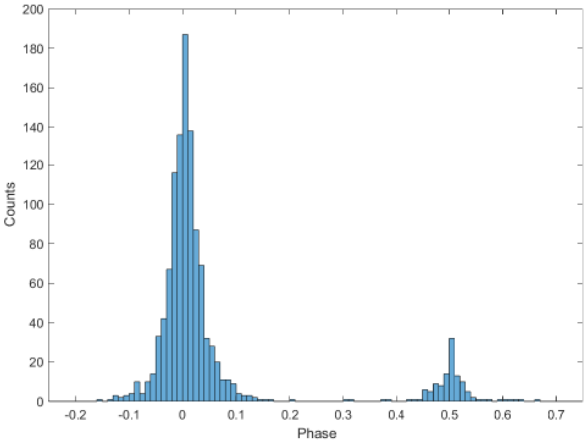

Fig. 29. Histogram of phase difference between SB-model primary eclipse and EB-model deeper eclipse, for 1174 stars in our sample.

van Kerkwijk et al. 2010; Maxted et al. 2011; Rappaport et al. 2015). EL CVn systems are at a rare stage of binary evolution in which the young pre-He-WD is bloated, with a radius of up to $\sim 0.5\,R_{\odot}$, and hotter than the more luminous A/F-type primary. As a result, such systems, harbouring a low-mass WD precursor, are detectable even in ground-based photometric surveys. EL CVns with smaller and cooler He-WD secondaries can be detected in space photometry (Faigler et al. 2015). Consequently, 10, 18, and 36 such systems were discovered in the *Kepler* (van Kerkwijk et al. 2010; Carter et al. 2011; Breton et al. 2012; Rappaport et al. 2015; Faigler et al. 2015), WASP (Maxted et al. 2011, 2014), and PTF (van Roestel et al. 2018) photometric surveys, respectively.

The detection of these systems in photometric surveys is based on identifying an eclipsing-binary folded light curve with a 'boxy' deeper eclipse (steep ingress and egress and a flat bottom) and a shallower eclipse with a limb-darkening curved bottom. In an EL CVn, the deeper boxy eclipse is actually the secondary eclipse (total eclipse of the pre-He-WD secondary by the MS primary), while the shallower eclipse is the primary eclipse (pre-He-WD transit of the primary star). This is because the pre-He-WD secondary is hotter than the primary. Such photometric detections usually require confirmation through follow-up spectroscopic RV observations, which enable the light-curve primary and secondary eclipses to be identified from the RV-curve phase.

However, the *Gaia* data enable direct detection of EL CVn systems by combining the *Gaia* photometry and RV data. Figure 28 shows the folded *Gaia G*, G_{BP}, and G_{RP} photometry and RV data, for a known EL CVn-type system (HD 23692, Maxted et al. 2014), together with the Transiting Exoplanet Survey Satellite (TESS; Ricker et al. 2015) binned data. Detrending of the TESS data was done using cosine detrending following Faigler et al. (2015). The *Gaia* EB-model period and deeper-eclipse epoch were used as the folding period and phase zero, respectively. The RV plot enables us to identify the phase-zero eclipse as the secondary eclipse, and the 0.5-phase eclipse as the primary eclipse. The figure shows that for this system, indeed the secondary eclipse is box-shaped and deeper than the primary one, the main signatures of an EL CVn system. In addition, we see that the G_{BP} secondary eclipse is much deeper than the G_{RP} one, an additional indication for the high temperature of the secondary.

6.3.1. Sorting through the *Gaia* data

To build the initial sample, from which we can identify EL CVn systems, we selected from the *Gaia* DR3 data systems with (a) an eclipsing-binary solution from *Gaia* photometry, (b) a SB solution (`SB1`, `SB1C`, `SB2` or `SB2C`) derived from the *Gaia* RV data, (c) an orbital-frequency difference between the EB and SB solutions of smaller than $\frac{1}{100}\,\mathrm{d}^{-1}$, and (d) an orbital period of shorter than 2 d. The maximum orbital-frequency difference was selected as significantly larger than the inverse of the data time span ($\sim$1000 d), a rough estimate for the orbital-frequency uncertainty lower limit. Limiting the orbital-frequency difference to $\frac{1}{500}\,\mathrm{d}^{-1}$ yielded the same sample. An orbital-period limit of 2 days was chosen because most discovered EL CVn systems are below it (see Fig. 5 of van Roestel et al. 2018). These criteria resulted in an initial sample of 1174 systems.

Next, we calculated the phase difference between the SB-model primary eclipse and the EB-model deeper eclipse for all stars in our sample. For a common binary, for example one consisting of two MS stars, we expect this phase difference to be zero. However, for an EL CVn, in which the secondary is hotter and therefore the secondary eclipse is deeper, we expect the phase difference to be $\sim$0.5, assuming a small eccentricity. Figure 29 shows the phase-difference histogram of our initial sample, with a main peak at phase zero, and a much smaller peak at phase 0.5, as expected.

Based on this, we selected an initial list of EL CVn candidates with: (a) an eclipse phase difference in the 0.4–0.6 range, (b) an eccentricity smaller than 0.3, and (c) an EB-model eclipse-depth difference with a S/N of higher than 5. The eclipse-depth difference S/N was required because our method relies on reliably identifying a secondary eclipse that is significantly deeper than the primary one. These criteria yielded 16 systems. Finally, we visually inspected the *Gaia* photometry and RV data and