Binary asteroid candidates in *Gaia* DR3 astrometry[*]

L. Liberato[1,2], P. Tanga[1], D. Mary[1], K. Minker[1], B. Carry[1], F. Spoto[3], P. Bartczak[5,8], B. Sicardy[4],
D. Oszkiewicz[5], and J. Desmars[6,7]

[1] Université Côte d'Azur, Observatoire de la Côte d'Azur, CNRS, Laboratoire Lagrange, Bd de l'Observatoire, CS 34229,
06304 Nice Cedex 4, France
e-mail: luana.liberato@oca.eu
[2] São Paulo State University, Grupo de Dinâmica Orbital e Planetologia, CEP 12516-410, Guaratinguetá, SP, Brazil
[3] Harvard-Smithsonian Center for Astrophysics, 60 Garden St., MS 15, Cambridge, MA 02138, USA
[4] LESIA, Paris Observatory, PSL University, CNRS, Sorbonne University, Univ. Paris Diderot, Sorbonne Paris Cité,
5 place Jules Janssen, 92195 Meudon, France
[5] Astronomical Observatory Institute, Faculty of Physics, Adam Mickiewicz University, Słoneczna 36, 60-286 Poznań, Poland
[6] Polytechnic Institute of Advanced Sciences-IPSA, 63 Boulevard de Brandebourg, 94200 Ivry-sur-Seine, France
[7] Institute of Celestial Mechanics and Ephemeris Calculation (IMCCE), Paris Observatory, PSL Research University, CNRS,
Sorbonne University, UPMC Univ Paris 06, Univ. Lille, 77 Av. Denfert-Rochereau, 75014 Paris, France
[8] Instituto Universitario de Física Aplicada a las Ciencias y las Tecnologías (IUFACyT), Universidad de Alicante, Ctra. San Vicente
del Raspeig s/n, 03690 San Vicente del Raspeig, Spain

ABSTRACT

Context. Asteroids with companions constitute an excellent sample for studying the collisional and dynamical evolution of minor planets. The currently known binary population were discovered by different complementary techniques that produce, for the moment, a strongly biased distribution, especially in a range of intermediate asteroid sizes ($\approx$2–100 km) where both mutual photometric events and high-resolution adaptive optic imaging are poorly efficient.
Aims. A totally independent technique of binary asteroid discovery, based on astrometry, can help to reveal new binary systems and populate a range of sizes and separations that remain nearly unexplored.
Methods. In this work, we describe a dedicated period detection method and its results for the *Gaia* DR3 data set. This method looks for the presence of a periodic signature in the orbit post-fit residuals.
Results. After conservative filtering and validation based on statistical and physical criteria, we are able to present a first sample of astrometric binary candidates, to be confirmed by other observation techniques such as photometric light curves and stellar occultations.

Key words. methods: data analysis – methods: statistical – astronomical databases: miscellaneous – astrometry – minor planets, asteroids: general

1. Introduction

Binary asteroids have attracted the attention of the scientific community due to their interesting properties and the significant impact they have on our understanding of the Solar System. Unlike single asteroids, binary systems offer unique insights into many fundamental processes, including the formation and evolution of planetary bodies, collision dynamics, and gravitational interactions.

Numerical simulations suggest the existence of nearly equal-sized binaries as a byproduct of catastrophic collisions (Durda et al. 2004) but they remain essentially unknown. On the other hand, multiple craters on Mars demonstrate that asteroid binaries with properties not represented by the known sample should exist, or have existed (Vavilov et al. 2022). More generally, the formation of asteroid companions by fragment ejection and in-orbit re-accumulation (Walsh & Richardson 2006; Walsh et al. 2008; Ćuk 2007; Pravec et al. 2010; Jacobson et al. 2013; Madeira et al. 2023) seems to be common among near-Earth

asteroids (and probably small main belt objects) but the collisional evolution is also expected to play a role (Doressoundiram et al. 1997; Michel et al. 2001; Durda et al. 2007; Jutzi 2019), with an undefined boundary between the two creation mechanisms. However, some open questions on the origin of binaries and the evolution of asteroids cannot be answered without a greater number of known objects and a better knowledge of the sample.

The presence of asteroid companions has been revealed by different techniques, such as high-resolution imaging from ground-based and space-based telescopes, photometry, radar ranging, and stellar occultations. When taken together, these different approaches are complementary and cover a wide range of separations and size ratios for the system components. However, the known population of binary asteroids is still strongly biased.

While imaging favours large separations of the primary and the secondary, and bright primaries in the case of adaptive optics, photometric studies, based on detecting mutual events, strongly favour compact systems (Merline et al. 2002; Pravec et al. 2006; Richardson & Walsh 2006). Radar techniques are strongly limited in range and are efficient for near-Earth asteroids (Ostro et al. 2002).

---

[*] Full Table 3 is available at the CDS via anonymous ftp to cdsarc.cds.unistra.fr (130.79.128.5) or via https://cdsarc.cds.unistra.fr/viz-bin/cat/J/A+A/688/A50

Some satellites have also been discovered by space probes and studied during close encounters. While they constitute a precious sample of detailed information, in statistical terms their contribution to the knowledge of the global population of binaries is limited. Astrometric detection of binaries alone cannot solve all these problems, but can very usefully fill an unexplored space of exploration (Pravec & Scheirich 2012). The astrometric signature of the presence of asteroid satellites is expected to show up in the residuals of the asteroid motion (on the sky or, equivalently, concerning the heliocentric orbit) due to the difference between the position of the photocentre (the one usually provided by the observations) and the centre of mass of the system. This difference, which in principle can also be due to the irregular shape of a single object, is enhanced by the presence of a satellite.

Anomalous residuals in the astrometry, linked to the presence of a satellite, have been revealed in the case of (90482) Orcus (Ortiz et al. 2011). For trans-Neptunian objects, the effect is amplified by the large separation of the components, particularly in cases of high size ratio (secondary/primary).

The astrometric signature is generally much harder to detect for main belt asteroids since it requires much more precise astrometry to reveal a perturbed motion due to a companion that is not spatially resolved. A rather solid detection, the first using *Gaia* astrometry, was obtained for (4337) Arecibo (Tanga et al. 2023). Also in this case, however, the presence of a satellite was already known. A systematic search of astrometry for new binary systems is a much more complex task. Only the exploration of a large data set of homogeneous measurements of extremely high quality can potentially be successful. The data produced by the *Gaia* mission of ESA are the main sources available today for this search. A first attempt on the very limited data set (in the time range and the number of objects) provided by *Gaia* DR2 was unsuccessful (Segev et al. 2023).

In the following we present a blind search of astrometric binaries in the *Gaia* DR3 asteroid astrometry. Our exploration represents a first step only where we focus on the detection of periodic signals (wobble, in the following) in the residuals obtained from orbit adjustment to the data. For the moment this approach limits our investigation to sets of astrometric positions collected in short time spans. As a consequence, short-period systems are favoured. The result of our search is also filtered on the basis of rather strict criteria. These choices produce results that are certainly not comprehensive of what can be obtained from *Gaia* data. Nonetheless, it provides a first consolidated set of candidate detections that can be fruitfully explored by other techniques for confirmation.

The article is organised as follows. In Sect. 2 we recall the main properties of the data set that we explore and the goal of the search. Section 3 presents the method of period search. The selection, filtering, and validation of the candidates found, by statistical and by physical criteria, are presented in Sect. 4. Section 5 presents the content of the final list of binary candidates. We put our findings in a broader context in Sect. 6, and resume the work, with future perspectives, in Sect. 7.

2. Astrometry and binary signatures: Data set and model

The asteroid astrometry in *Gaia* DR3 for a large number of asteroids provides an outstanding opportunity for a search of astrometric signatures produced by possible satellites. The properties of *Gaia* astrometry are illustrated in detail by Tanga et al. (2023).

Here we recall a few properties specifically relevant to our search.

The first property is the time distribution of observations. The two *Gaia* telescopes scan the sky continuously with a six-hour rotation period. The two fields of view (FOV) sweep the same area of the sky 106 minutes apart. The spin axis precesses in such a way that a large field overlap is granted every subsequent rotation. The combination of this rotation–precession with the revolution around the Sun results in a sparse series of detections (also called transits in the FOV) for any given source, fixed or moving. However, when the motion on the scanning path in the sky combines favourably with the position and displacement of a target source, a short consecutive series of detections over a few satellite rotations is possible.

Over the coverage of *Gaia* DR3 (34 months), it is rather common to obtain sequences of a few consecutive observations spanning a few rotations (12h, four detections) for each object. A few occurrences of longer sequences with several rotations up to 10–12 detections (five or six revolutions) can also happen, although less frequently.

As is discussed in the following section, the length of time windows of consecutive data and the presence of constant periods in the data sampling control, respectively, the longest detectable period in our analysis, and the possible emergence of spurious frequencies. At the same time, irregular sampling allows the exploration of a large range of frequencies (Eyer & Bartholdi 1999; VanderPlas 2018).

In addition, the nearly mono-dimensional astrometry provided for each transit brings accurate information only in the direction of the scan (along-scan, AL). The only measurable residual from orbital fitting is the projection of the total residual onto the AL direction, whose direction is known from the attitude data of the satellite. The direction of the total post-orbital fit residual in the sky, however, is not directly measurable from *Gaia* astrometry alone at a given epoch (or in an interval of time in which the geometry of the system concerning the scan direction does not evolve). Constraints on the total wobble in the astrometry, and the possible primary-secondary separation, are the result of applying some physical constraints to the measured wobble projection and period.

For this first search for astrometric binaries, we use the simplest first-order model for the wobbling motion, assuming systems composed of uniformly illuminated spheres, of equal bulk densities and albedos. The other important assumption is that the components are not resolved so that the wobble is generated by the difference in position between the barycentre of the system (to which the orbit is adjusted) and the photocentre. Given the size of the *Gaia* pixel (60 mas) in the AL direction, this should be the case for a wide range of binaries in a reasonable range of sizes.

Figure 1 illustrates the principles discussed, where at the top of the plot it is shown a representation of an ideal case where *Gaia* is observing a binary asteroid perpendicularly to the plane of rotation of the objects around the barycentre, the along-scan direction (AL) of the *Gaia* satellite is on the *x*-axis and the across-scan direction is on the y-axis. In principle, at any given epoch, the projection of the instantaneous wobble amplitude α for the satellite position on its orbit, projected along AL, is provided by the residual of *Gaia* astrometry to the orbital fit. The blue and green dashed lines represent the projection of the positions of the photocentre and the barycentre of the binary, respectively. The difference between these projections represents the AL residuals, as shown at the bottom of Fig. 1.
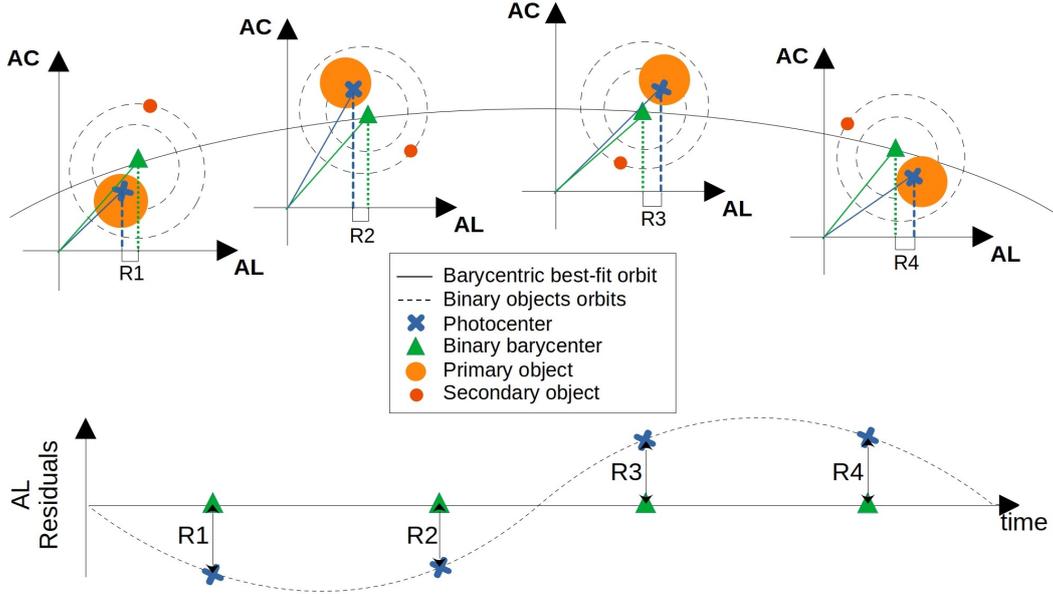
Fig. 1. Representation of a simplified case where the *Gaia* satellite would be observing a binary system perpendicularly to the plane of rotation of the bodies, along with the projections of the positions of the barycentre and the photocentre on the AL direction and the representation of the orbital fit residuals' projection in the along-scan direction. AL direction is represented horizontally for simplicity, but in reality it can take any orientation, slowly rotating by some degrees on a scale of several hours.

With the assumptions of uniform illumination and unresolved components, it is possible to derive the wobble amplitude as the difference between the position of the barycentre and the position of the photocentre. This is a function of the mass ratio q only (Hestroffer et al. 2010), scaled for the separation a:

$$\alpha = a \left| \frac{1}{1 + q^{-1}} - \frac{1}{1 + q^{-\frac{2}{3}}} \right|, \quad q \in \,]0, 1] \tag{1}$$

This function (Fig. 2) has a maximum at $q_{\max} \approx 0.154$ (corresponding to a size ratio $k = \sqrt[3]{q_{\max}} = 0.536$) and tends to zero at the extremes of the range of $q \in \,]0, 1]$, that is, for a vanishingly small satellite or equal-sized components. During a single transit in the Astrometric Field of *Gaia* (about 40 seconds), up to nine consecutive positions are obtained. We fit all *Gaia* DR3 astrometry for each asteroid by a version of the ORBFIT software[1], optimised to exploit the full accuracy of *Gaia*, following the same procedure as in Gaia Collaboration (2018). From the orbital fit, for each astrometric point, a residual is obtained in the AL direction.

Given that the observation geometry over that time span is nearly constant (and a relative displacement of a satellite negligible) we average the residuals from all transits from one observation into a single transit. Their standard deviation also provides uncertainties on the measured residual itself. The average AL residuals per transit, and their uncertainty, are the data on which we base our analysis. To prevent future confusion, throughout the work we refer to the residuals from the orbital fitting of the average transit observations in the along-scan direction simply as AL residuals.

3. Wobbling detection

For this first data exploration, our search targets intervals of time in which several consecutive astrometric measurements are

---

[1] See http://adams.dm.unipi.it/orbfit
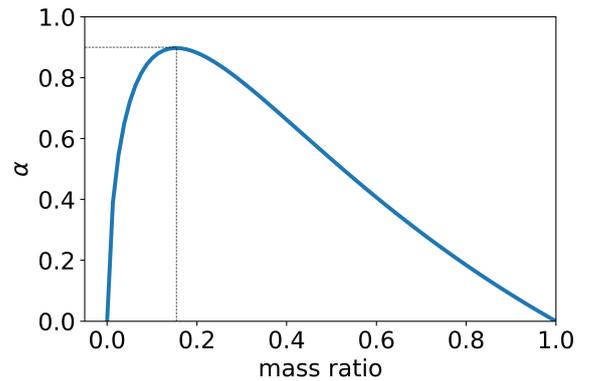
Fig. 2. Plot of the amplitude of the photocentre wobbling as a function of the binary mass ratio, as described in Eq. (1). In this example we use a separation $a = 10$ mas, which means that we would observe a maximum photocentre wobbling of 0.9 mas for a mass ratio $q \approx 0.1538$.

available, sufficiently short to neglect any change of orientation of the asteroid with respect to *Gaia*. On the observations available over that interval, a period search is performed.

3.1. Sample selection

To easily identify the longest sequences of consecutive astrometric measurements, we perform a convolution of the observation epochs with a rectangular function and find the peaks of large data density, as shown in Fig. 3. In the left plot, we show the density of AL residuals over time for asteroid (3457) Arnenordheim, and the peak close to 450 days represents the 2.75-day window of time with 17 observations shown in the right plot. The rectangular function has a unity value over an interval of ten days and zero elsewhere.

The considered threshold of ten days is a compromise between the length of the time series and the resulting number of candidates, as longer time series allow for better frequency
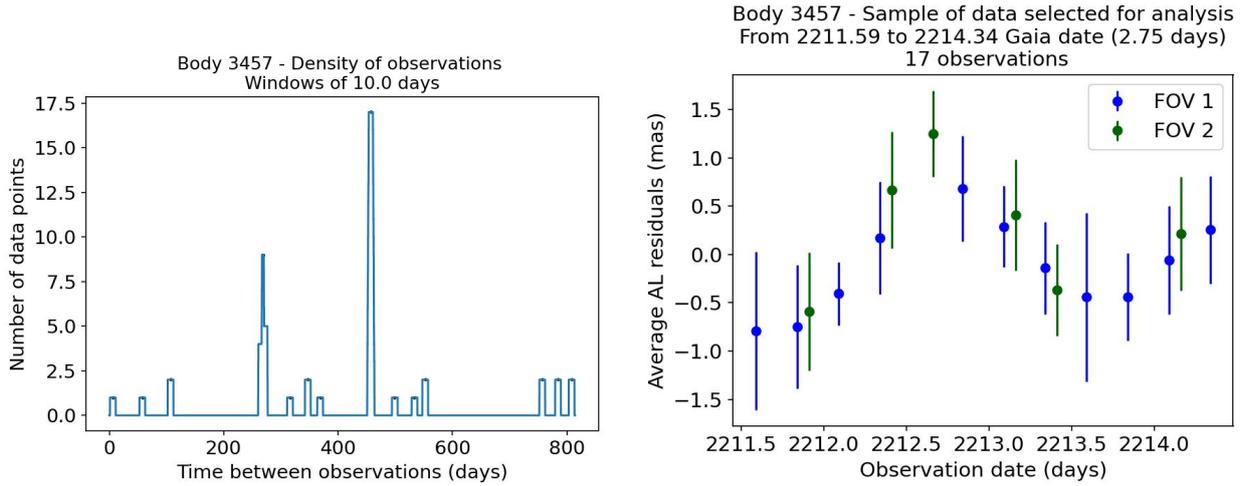
Fig. 3. Plots of the density of observations (number in ten consecutive days) for asteroid (3457) Arnenordheim and selected sample of data. *Left*: resulting convolution of the observation time span in days vs the number of observations in each ten-day window. *Right*: sample of data at the peak of the 10-day window of search combining 17 observations. The *x*-axis is the observation date and the y-axis is the average AL residual from the orbital fit. Each dot represents one averaged set of transits observed in the FOV 1 in blue and FOV 2 in green.

analysis but reduce the number of targets. It should be noted that this selection method allows for gaps in the data sequence, where the observations are absent due to some technical reasons (missed observation, rejection for low quality). Over the whole DR3 data set of 156 801 asteroids, our selection extracts 30 030 objects with sequences satisfying the above-mentioned criteria.

Next, we identify the observations in each one of the windows and perform an outlier check, removing the points with nominal values or standard deviations larger than 3σ of the dispersion of AL residuals in the window of search.

3.2. Period search

The Generalised Lomb–Scargle Periodogram (GLSP) (Zechmeister & Kürster 2009; VanderPlas & Ivezic 2015; VanderPlas 2018) is a statistical technique for analysing irregularly sampled time series data and identifying periodic signals in the data. The GLSP method is a generalisation of the classical Lomb–Scargle Periodogram (LSP) (Lomb 1976; Scargle 1982).

As described in VanderPlas (2018), the GLSP corresponds to a weighted least-squares problem where a sinusoidal signal of unknown amplitude and phase plus an unknown constant are jointly estimated at each test frequency. Formally, using the notation from VanderPlas (2018), the GLSP $P(f)$ can be defined from the following equations:

$$y_{\text{model}}(t; f) \equiv y_0(f) + A_{\text{f}} \sin(2\pi f t + \phi_f),\tag{2}$$

$$\chi^2(f) \equiv \sum_{n=1}^{N} \left(\frac{y_n - y_{\text{model}}(t_n; f)}{\sigma_n} \right)^2,\tag{3}$$

$$P(f) \equiv \frac{\widehat{\chi_0^2} - \widehat{\chi^2}(f)}{\widehat{\chi_0^2}}.\tag{4}$$

Here, y_{model} Eq. (2) is the assumed true model: a sinusoid (unknown frequency f, amplitude A_{f}, and phase ϕ_f) plus a constant offset term y_0 that may depend on the considered frequency. If the N data points of a time series y, $y(t_n) \equiv y_n$ for

$n = 1, \cdots, N$, correspond to the sum of the true model of Eq. (2) plus uncorrelated Gaussian noise of variance σ_n^2 for each of the n samples, then the weighted sum in Eq. (3) is a χ^2 random variable with N degrees of freedom.

The GLSP computes the residual sum of squares (RSS) of the weighted residuals (noted $\widehat{\chi^2}(f)$) obtained when adjusting a sinusoid with unknown amplitude and phase for each test frequency plus a constant and compares it to the RSS obtained when fitting only a constant ($\widehat{\chi_0^2}$). Precisely, $\widehat{\chi^2}(f)$ (resp., $\widehat{\chi_0^2}$) is obtained by replacing y_{model} by the estimated sinusoid plus constant (resp., the estimated constant) in Eq. (3). The GLSP $P(f)$ in Eq. (4) is thus the relative reduction in the sum of square brought by the adjusted model at considered frequency f (with some abuse of notation on f). The higher $P(f)$, the smaller $\widehat{\chi^2}(f)$, which means that the model and the data are in greater agreement.

The GLSP is probably the most common technique to detect periodic signals in time series with irregular sampling. We opted for this approach because it has proven efficiency and is computationally simple enough to allow for large-scale Monte Carlo simulations (this is an important aspect of our study, see next section). The GLSP in Eq. (4) is implemented with the ASTROPY package (Astropy Collaboration 2013, 2018, 2022), with error bars σ_n estimated as explained in the last paragraph of Sect. 2.

Given the sampling frequency and the time range of ten days for our data sequences, we limit the search to periods between 3 h and 3 times the length of the observations sample and apply our period search to all the asteroids in the 30 030 objects selected in the previous step.

The period search provides a periodogram for all the studied bodies, and they are very different depending on the time sampling of observations. The highest peak in the periodogram represents the period that best fits the sample. The left plot of Fig. 4 shows, as an example, the GLSP obtained for the asteroid (106016) 2000 SS293 from the sample of AL residuals shown in the right plot of the same figure. The red line is the GLSP from the fitting on the data of AL residuals in the current window of search. Using the frequency with the highest power from the GLSP on the data we obtain the sinusoidal fit with amplitude A and, along with the same time sampling and uncertainties, we run the GLSP to obtain the blue line that represents the profile
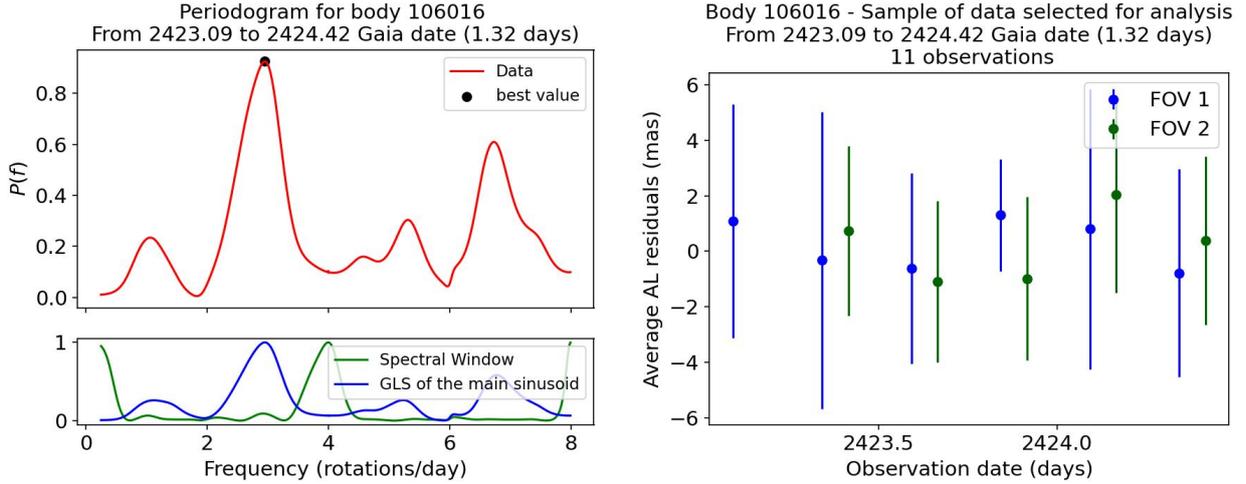
Fig. 4. Examples of periodograms obtained for asteroid (106016) 2000 SS293 (left plot), with residuals shown on the right plot. The *y*-axis shows the GLSP $P(f)$ for each frequency shown on the *x*-axis. The red line represents the periodogram obtained for the data analysed. The blue line represents the GLSP obtained for a purely sinusoidal signal whose parameters correspond to the largest peak of $P(f)$ (about 3 cycles per day here), with the same time sampling as the data and the same error bars for the weights. Finally, the green line represents the spectral window of the time sampling. The right plot shows the AL residual time series that reproduces the periodogram on the left plot.

of the periodogram only due to the sinusoidal signal. Finally, the spectral window (green line) is obtained by computing the Lomb–Scargle periodogram of a time series with constant unit amplitude[2].

It is interesting to compare the shapes of the GLSP of the data (red curve in the left panel of Fig. 4), of the main estimated sinusoid (blue curve) and spectral window (green curve). The spectral window indicates the specific frequency pattern related to the specific time sampling in a time series. Similarities in the structures of the spectral window and the GLSP are often a clear sign that some peaks are fake (sampling-induced) rather than genuine signal periodicities. Similarly, the pure sinusoid analysis (blue curve) encodes the same information but focuses on the analysis of a particular frequency. As an example, we see that the peak at almost 7 cycles per day in the data GLSP (red curve) is probably, from the blue curve, a spurious sidelobe created by the specific time sampling. This is also visible from the green curve, which shows that an alias peak appears at four rotations/days after the main peak.

From the considerations explained above, it is clear that a detection method based on the highest peak of the periodogram must evaluate the "statistical significance" of a given maximum peak's height. One way to do this is to measure the probability of obtaining a value of the largest peak as high as the one observed in the data when there is only noise in the time series. This is investigated in the next section.

4. Binary candidates selection method

4.1. Statistical selection

For each window of search of each object, we computed the GLSP as described in Sect. 3.2 and obtained the value (say, *v*) of the highest peak. A common proxy for the "statistical significance" of *v* is the *p*-value (see e.g. Efron 2010), defined as

follows. Let *V* be the random variable defined as the maximum value of the periodogram, conditional on the time sampling grid and error bars of the considered data. The *p*-value of *v* is the probability of obtaining a maximum peak value at least as large as *v* under the signal-absent hypothesis, with the same time sampling and noise error bars (denoted by $\mathcal{H}_0$). Formally this probability is

$$p(v) \equiv \Pr(V > v \mid \mathcal{H}_0). \qquad (5)$$

In practice, the *p*-value $p(v)$ of a particular time series can be evaluated by Monte Carlo simulations by estimating empirically this probability as follows. We generate 10 000 simulated time series sampled with the same time sampling as the time series under investigation. Each time series is composed of a Gaussian noise consistent with the error bars plus a random constant (an offset drawn randomly from a zero-mean Laplacian distribution, which is consistent with our data, see Appendix A). The GLSPs of the 10 000 time series are computed and we count how often their highest peak has an equal or larger power than the one from the GLSP under investigation (*v*).

A *p*-value is computed in this way for each of the selected 30 030 time series. To be selected as a possible candidate, the time series must satisfy three conditions. The first condition is that the *p*-value estimated from the MC under $\mathcal{H}_0$ must be smaller than 0.05. The second condition is a check of the "robustness" of the periodicity which we call "quality factor" below, and which also provides a method to estimate uncertainties for the amplitude and frequency associated with the main peak of the GLSP. For this second condition, we perform MC simulations under the hypothesis that there is a sinusoidal signal in the data with the frequency and amplitude found from the GLSP of the candidate time series (plus noise with the considered error bars, plus an unknown offset, and with same time-sampling; we note this hypothesis $\widehat{\mathcal{H}}_1$). We then perform another series of 5000 MC simulations of synthetic time series under $\widehat{\mathcal{H}}_1$.

A robust period detection should repeat nearly identically several times with different realisations of the synthetic data. To obtain the quality factor of the signal, denoted by *Q* below, we estimate the fraction of time series for which the GLSP highest

---

[2] Precisely, the spectral window indicates, at each frequency, the reduction of the residual sum of squares (relative to the initial sum of squares) obtained when fitting a pure sinusoid at that frequency to a constant signal sampled like the data.
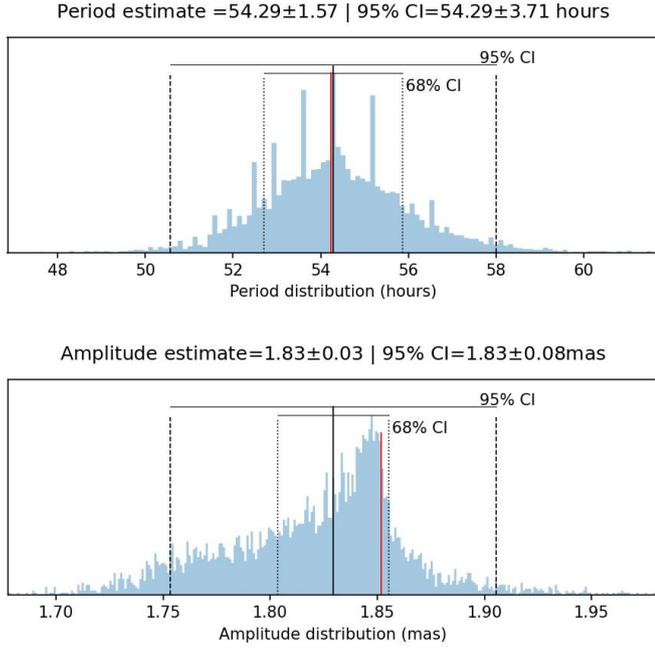
Fig. 5. Examples of uncertainty determination from MC simulations under $\widehat{\mathcal{H}_1}$. The top and bottom plots show, respectively, the distribution of periods and amplitudes obtained from the best fit in each MC simulation (in blue). The red lines show the value estimated from the first GLSP run only on *Gaia* data, and the solid black line represents the median obtained from the distribution of amplitude and period values from the MC under $\widehat{\mathcal{H}_1}$. The dotted vertical lines show the 68% distribution interval while the dashed lines represent the 95% confidence interval of the distribution. The final results are given as the first estimate from the *Gaia* data (red line), and the uncertainty is the 95% interval. For instance, in this example, the final amplitude value is 1.85 ± 0.08.

peak is found ± 0.15 cycles per day around the injected frequency. Our second condition is that Q should be larger than 0.5, meaning that the initially injected period should be retrieved more than half of the time in nominal noise and sampling conditions.

We underline that a large maximum peak caused by noise in the GLSP (i.e. a false alarm) will lead to a high value of the quality factor, because the wobbling amplitude estimated from the data will be large, so any strong but 'fake' detection will appear as robust concerning the Q analysis. In the opposite direction, a maximum peak with low amplitude will lead to a low estimated wobbling amplitude and a low value of the quality factor. Hence, while a large value of Q is not a guarantee that the detection is robust, the cross-analysis brought by the quality factor essentially provides a supplementary flag that some detections may not be robust (this will be further investigated in Sect. 4.2).

Finally, we turn to the third condition. An important aspect is that we are looking for astrometric binaries, which means that *Gaia* satellite would not have been able to observe the asteroid and the satellite separately. A wobble amplitude of 20 mas translates to a minimum separation of about 220 mas, which is above the *Gaia* resolution limit of about 200 mas. Hence, the third condition is that the wobble amplitude estimated $\widehat{A}$ must be smaller than 20 mas.

In summary, if the signal detected in the current window has a p-value $< 5\%$ as estimated from MC simulations under $\mathcal{H}_0$, a quality factor $Q > 0.5$ as estimated from MC simulations under

$\widehat{\mathcal{H}_1}$ and estimated wobble amplitude smaller than 20 mas, then the object is selected as a preliminary candidate and the values estimated in the current window are stored. The threshold of p-value $< 5\%$ is not very conservative since the expected number of false positives is about 1500 (5% of 30 030). In our case, the number of candidates that pass the triple test above is 3038. This number (larger than 1500) can be explained by two reasons. First, by a population of true binaries that lead to smaller p-values. Second, by possible modelling errors (for instance, some un-modelled, low-order polynomial trends would lead to GLSP with larger peaks at low frequencies)[3]. However, as we discuss in Sect. 4.2, the latter seems unlikely.

To obtain estimates of the period and amplitude uncertainties, we use the results from the MC simulations under $\widehat{\mathcal{H}_1}$. In the case of mean period determination, we look for frequencies that are no more than 0.25 cycles per day away from the best frequency found with the first GLSP and scan along the distribution to find the mean value, the 68% confidence interval, which corresponds to the mean $\pm 1\sigma$ for a Gaussian distribution, and the 95% confidence interval. The value estimated from the first GLSP run is selected as the wobble period and the 95% CI is the uncertainty on the period. The procedure is the same for the amplitude, but in this case, we limit the interval for amplitudes up to 2 mas away from the first amplitude estimate. Figure 5 shows an example of uncertainty determination for the period on the top plot and the amplitude on the bottom plot.

If the object under the period search has more than one window of data, as discussed in Sect. 3.1, we repeat the same procedure described in all the other windows as well. Figure 6 shows a flowchart that summarises the statistical selection processes discussed above.

4.2. Analysis of the detection performance of the statistical selection process

To evaluate the detection performance of our method, in this section, we conduct a general study based on the analysis of two cases. In the first case, there is indeed a signal in each time series analysed and we look at the effectiveness of the method at detecting and characterising it. In the second case, the time series have no signal and we examine how incorrect the method can be in falsely selecting residuals with only noise. A supplementary investigation into the detection of eccentric binary systems is discussed in Appendix B.

4.2.1. Signal detection

In this analysis, we study the signal detection performance by grouping the simulated time series into classes of signal-to-noise ratio and looking at the results in terms of detections (at the right period or not), missed detections, associated estimated p-values and Q-factors. We use two signal-to-noise ratio definitions:

"Base S/N" : $S/N_{\text{base}} \equiv A/\sigma$ and $\qquad\qquad$ (6)

"Post-fit S/N" : $\widehat{S/N} \equiv \widehat{A}/\bar{\sigma}$, $\qquad\qquad$ (7)

where A is the true amplitude of the time series and σ is the nominal average uncertainty of the data set, $\widehat{A}$ is the amplitude

---

[3] While we do have some hints that a model with low-order polynomials could be more accurate than the model in Eq. (2), this would require to compute a much more computationally expensive least-squares fitting at each frequency than the current GLSP. This improvement will be implemented in a subsequent version of our detection procedure.
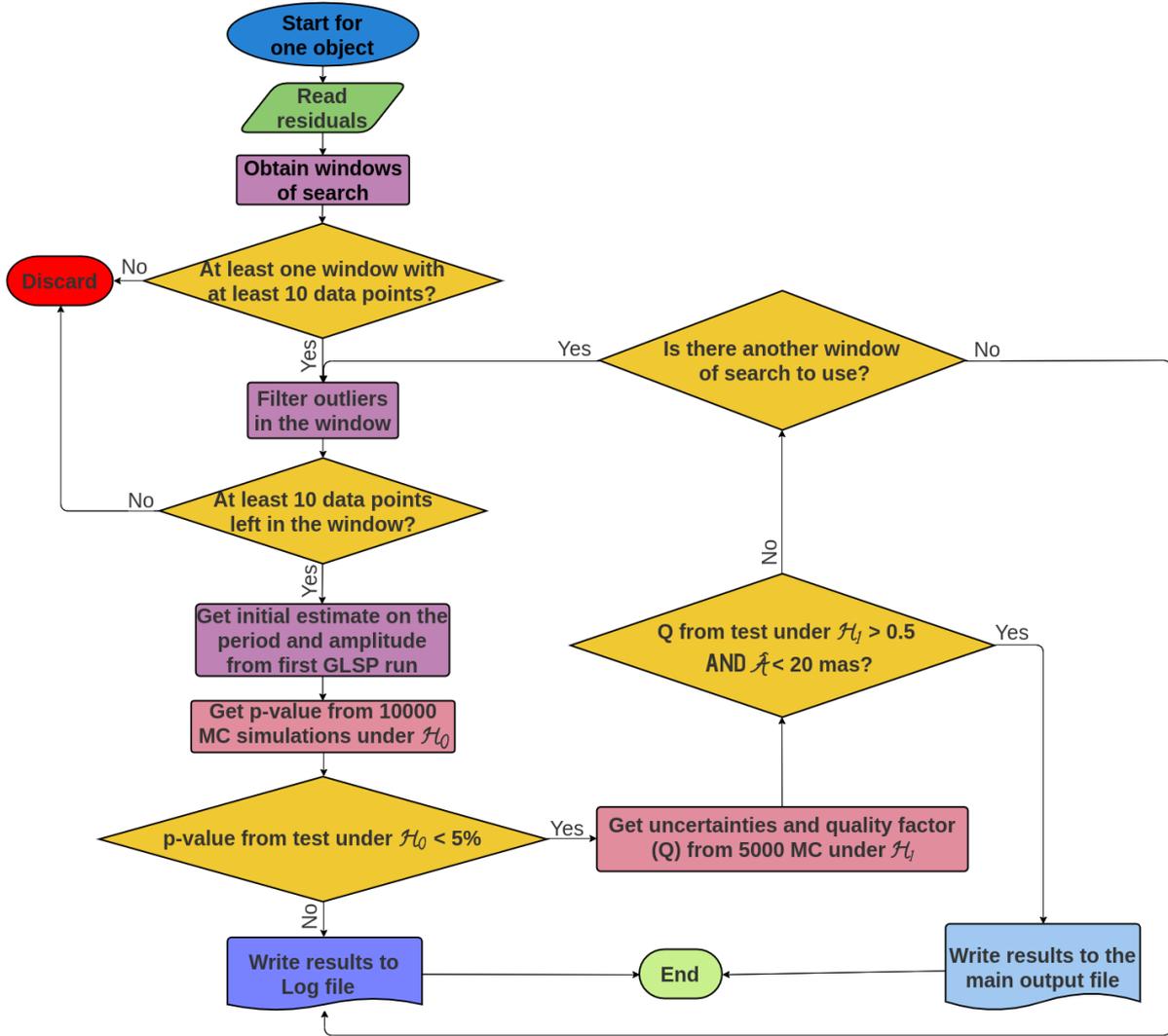
Fig. 6. Flowchart showing the period search and statistical selection processes discussed in Sects. 3 and 4. It represents part of the full detection process, which is coupled with physical characterisation tests that are discussed in the following sections.

estimate from the data time series and $\bar{\sigma}$ is the estimated average uncertainty of the data set. To obtain the performance analysis we perform the period search and selection method shown in Fig. 6 for six different S/N_{base} values, with Monte Carlo simulations based on 1000 time series of synthetic AL residual.

First, we choose a set of different values of S/N_{base} between 0.5 and 3.0. We create each set of synthetic data from a sinusoid with a known amplitude A of 1 mas, a period of 23 hours (a frequency close to but not exactly 1 cycle per day to avoid possible bias) and a random phase between 0 and 2π. The nominal average noise σ for each set of synthetic samples is estimated using Eq. (6).

For the simulations of the time series, we draw 12 observation epochs, randomly distributed in 6 days. The small number of points, the randomness in the data and the time sampling are aimed not only at representing the data from *Gaia*, which is usually modestly sampled over a few days but also at providing conservative results that cover most of the non-ideal cases.

For each time series, the actual standard deviations of the errors σ_n are obtained by a random perturbation of σ (using a zero-mean, normal distribution of std 0.8; this setting is aimed at creating some scatter in the actual S/N within each S/N

class), and we add one random Laplacian offset as described in Appendix A. Then, for each one of the 6000 synthetic time series simulated in this way, we run the statistical selection (Sect. 4.1).

Figure 7 shows the results for the (p, Q) pairs of 6000 time series simulated. We found that, as expected, for low values of S/N_{base} the parameters p and Q tend to be more scattered and present less correlation between themselves. But for higher S/N_{base} values, the couples (p, Q) tend to cluster around the "best conditions" represented by very low p-values and high Q.

Intuitively, the larger the S/N_{base} of a signal the easier it would be for the method to detect it. It is also expected that very noisy signals hamper the period fitting and provide incorrect detections more often, and our results follow that behaviour. For time series with S/N_{base} close to 0.5, the method achieves 11.3% of detections, of which 69% have incorrect periods. For the S/N_{base} class of 1.0, the detections represent 56% and about half present incorrect periods. For the S/N_{base} classes larger than 2.0, the method detects more than 93% of the signals with a low rate of incorrect detections.

It is also interesting to check the distribution of the quality factor (Q) and the p-value depending on the S/N_{base}, as presented in Fig. 8. It shows the distribution of best periods fitted
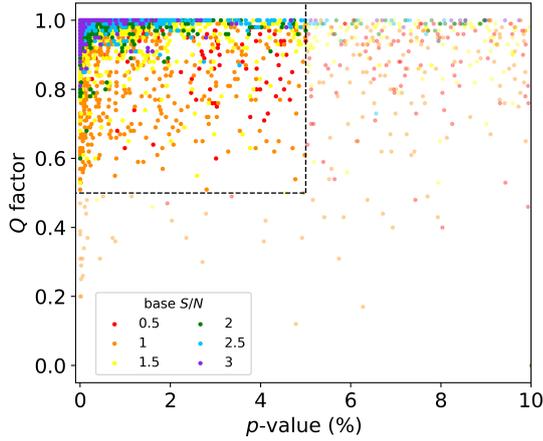
Fig. 7. Distribution of pairs (p, Q) for each of the 6000 sets of synthetic data simulated. The partially transparent dots represent the cases where the detections do not meet the requirements and were disqualified, while the remaining dots represent the detections selected. Of the 6000 simulations, 1303 have p-value>0.1. Hence, their Q factors were not calculated and the corresponding points are not shown in this plot.

on the top plots, the distribution of Q in the plots in the middle and the p-value distributions on the bottom plots for each different class of S/N_{base} studied. As discussed before, we notice that for small S/N_{base} there is a significant amount of detections with periods far from the real value, represented by the dotted black line. In addition, the corresponding distribution of Q shows that the quality of the detection tends to be mostly below 0.8, much smaller than for larger values of S/N_{base} where Q is usually closer to 1.0. Similar behaviour can also be seen in the p-value distributions where the higher the S/N_{base}, the larger the concentration of results with small p-values.

The p-value is usually the only parameter used in the selection of detections. In fact, for $S/N_{base} > 1$, we find that almost all missed detections have $Q > 0.5$, so the p-value is the parameter ruling the selection and Q has almost no influence in the selection process. However, for $S/N_{base} \leq 1$ there are many cases with p-values smaller than 5% where Q is used to filter the weak detections, as seen in Fig. 7.

Of course, when applying the detection method to real *Gaia* data, we can only estimate the signal-to-noise ratio based on the sinusoidal fit from GLSP. This is the reason why, in addition to the analysis of the detection performance based on the base S/N_{base} values, it is also interesting to study the results based on the post-fit S/N ($\widehat{S/N}$), which is indeed different from the S/N_{base}. As shown in Fig. 9 we see that the number of detections increases for the largest values of S/N_{base}, as in Table 1.

By combining all the results from the 6000 time series, we can simulate the situation where we do not know the original parameters of the time series and have only the post-fit signal-to-noise ratio. Then, we separate them in intervals of mean $\widehat{S/N} \pm 0.1$ and check the statistics, as it was made previously for the S/N_{base} classes. The results are presented in Table 2. From the comparison between the analysis based on S/N_{base} classes presented previously, and the one based on the post-fit $\widehat{S/N}$, these results indicate that there are no major differences in the statistics of detections between both approaches, especially for $\widehat{S/N} \geq 1.5$.

It is interesting to note from these results that in the scenario where there are signals in all of the time series analysed, but we do not know the S/N_{base}, $\widehat{S/N} \leq 1$ lead to p-values and Q

factors that often fail to pass the selection test. Meanwhile, for $\widehat{S/N} \geq 2$ we find more than 85% of the detections around the correct period with a rate of incorrect/correct lower than 10% and we miss less than 3% of the detections.

We conclude that a $\widehat{S/N}$ of about one appears to be a limit for the detection regime in this setting. For increasingly stronger signals, with larger $\widehat{S/N}$, we have an increasingly smaller rate of incorrect detections. The method succeeds accordingly at classifying the sample as detection, at recovering the correct period in the data, and at providing a good approximation of the real S/N.

4.2.2. Noise-only scenario

The purpose of this analysis is to understand how the method performs when there is no signal in the data and to assess its robustness in handling data with no wobble at all. For that purpose, we generated 6000 sets of simulated AL residuals by replicating the error bars and time sampling of original AL residuals from *Gaia* asteroids. These synthetic noisy time series were generated by drawing random samples with Gaussian noise perturbation corresponding to the target error bars and we added in each time series a random offset consistent with the *Gaia* time series, cf. Appendix A.

As discussed in Sect. 4.1, with a p-value threshold of 0.05 it is by definition expected that 5% of the time series are selected (on average) if none of the data present a real signal. Our results showed that in the considered noise-only scenario, the statistical selection process selected as detections about 4.8% of the time series analysed.

The left plot in Fig. 10 shows the distribution of (p, Q) pairs for the 609 time series with p-value < 0.1. The remaining 5391 time series present p-value > 0.1 and therefore their Q was not calculated. We see that the resulting dispersion in the (p, Q) plane is similar to that observed in Fig. 7 for classes of $S/N_{base} \leq 1$. This is in contrast with classes of larger S/N values for which the points tend to cluster around the top left corner.

Turning to the right panel of Fig. 10, we see that noise-only data lead to post-fit S/N values distributed around $\widehat{S/N} \approx 1$, with few cases having $\widehat{S/N} \geq 2$. As expected, this distribution is compatible with the lowest S/N classes of Fig. 9.

In a nutshell, we conclude from this two-fold performance study that, when a wobble is indeed present, our statistical selection process behaves as expected and increasingly succeeds at retrieving this wobble as its amplitude increases. In the situation when there is no wobble, the incorrect detection rate is controlled and does not exceed the prescribed value (set to 5% here).

Finally, the S/N value of one appears to be a turning point in the detection regime, post-fit S/N values larger than ≈ 2 being a serious hint of the presence of a population of real wobbles[4].

4.3. Period search results overview

In this section, we apply the detection method to the AL residuals data set. Figure 11 visually illustrates the results of the

---

[4] Indeed, the detectability of a signal depends also on the time sampling. Two wobbles with the same S/N but sampled differently will in general have different p-values and probabilities of detection. In our simulations, the time sampling of the simulations is generated in agreement with the *Gaia* sampling of the considered AL residuals. Our results are averaged over a large number of different sampling grids. Hence, the frontier of a unit S/N should not be taken at face value but reflects a global behaviour corresponding to the considered sampling conditions.
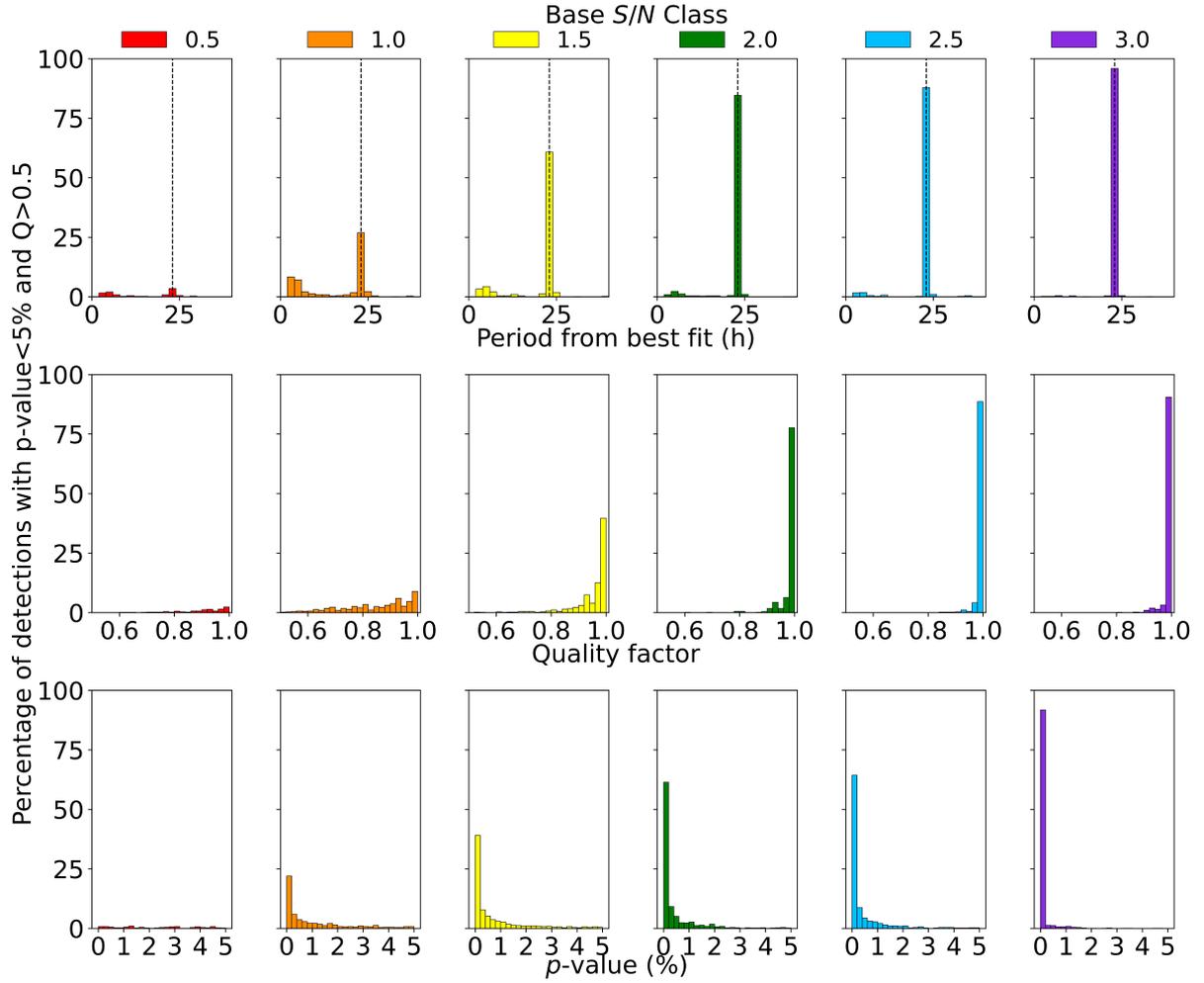
Fig. 8. Distribution of periods from best fit (top row), quality factors (centre row), and *p*-values (bottom row) as percentages of the detections selected with *p*-value < 5% and $Q > 0.5$. The different coloured bars represent the different classes of S/N_{base} values explored. The dotted line at 23 h represents the nominal (true) period of the injected signal.
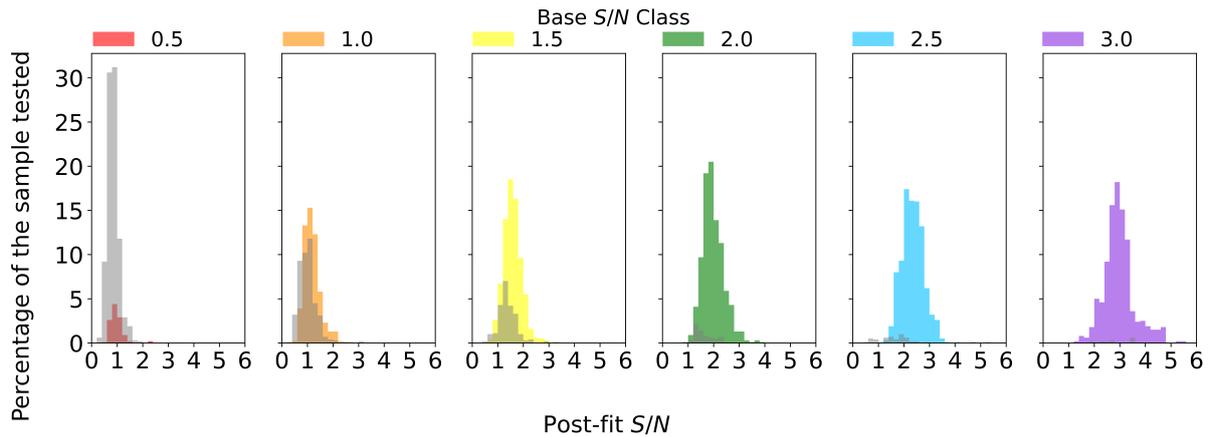
Fig. 9. Distribution of time series selected as detections (coloured bars) and discarded (grey) as a function of the post-fit $\widehat{S/N}$, for the six different S/N_{base} classes explored.

processes of selection and parameters determination for the case of an object that satisfies the requirements (at least one window with ten observations in less than 10 days, *p*-value < 0.05, $Q > 0.5$ and estimated wobble amplitude $\hat{A}$ smaller than 20 mas) and was selected as a possible candidate by the period search. In this case, for (3457) Arnenordheim, we have a window with 17 observations spread over 2.75 days (as shown in Fig. 3, right panel). The periodogram shows a large peak (Fig. 11 top left) that provides the best fit in the observations with a period of 45.95h (Fig. 11 top right). The MC analysis leads to a very low *p*-value

Table 1. Results for each class of S/N$_{base}$ studied.

S/N$_{base}$	Detections	Detections around correct period	Detections with incorrect period	Missed detections	Incorrect/correct period detection rate
0.5	11.30	3.50	7.80	19.60	69.0
1.0	56.10	26.90	29.20	13.80	52.0
1.5	76.88	60.76	16.12	11.81	21.0
2.0	93.30	84.60	8.70	5.10	9.3
2.5	95.20	87.80	7.40	3.00	7.8
3.0	98.40	95.90	2.50	1.30	2.5

Notes. The first column represents the base signal-to-noise ratio used. The second column shows the total percentage of detections in the sample. The third column shows the percentage of detections in which the best period fit is up to one hour away from the nominal period chosen. The fourth column presents the percentage of detections where the period is further than one hour from the nominal value, complementary to the previous column. The fifth column shows the percentage of missed detections because they did not pass the selection tests, but the period found is close to the nominal period. The last column shows the fraction of the detections with incorrect periods concerning the total number of detections, representing the probability of incorrect detection.

Table 2. Same as Table 1, but with the analysis based on the $\widehat{S/N}$.

$\widehat{S/N}$	Detections	Detections around correct period	Detections with incorrect period	Missed detections	Incorrect/correct period detection rate
0.5	11.61	2.90	8.71	17.10	75.0
1.0	43.25	20.68	22.56	16.75	52.2
1.5	76.56	56.64	19.92	11.62	26.0
2.0	94.58	85.85	8.73	2.83	9.2
2.5	98.15	94.14	4.01	1.54	4.1
3.0	99.15	94.87	4.27	0.00	4.3

($< 0.001\%$, Fig. 11 bottom left) and a quality factor Q of one (Fig. 11 bottom right).

The distribution of estimated periods and wobble amplitudes $\hat{A}$ for the 30 030 candidates (Fig. 12) shows that small amplitudes are preferentially discarded, as expected since the signal-to-noise ratio tends to be small, but the general distribution remains overall the same. We notice a kind of "tail" in the distribution for the large period and amplitude values, and they are mostly not selected as detections. This phenomenon might be caused by unmodelled low-frequency trends (see Sect. 4.1).

It is also noticeable in Fig. 12 a concentration of detections around 3, 4, 6, and 12 h with larger amplitudes, even though there is no over-representation of such periods. That happens due to the typical frequencies present in the *Gaia* scanning law (time interval between the two FOVs, rotation period of 6 h, precession of about 2 months, and combination of them). In some cases, the sampling of the data set has a dominant effect on the period search causing the GLSP to find the best period at, or very close to, values of 2, 4, 6, and 8 cycles per day. However, some of the detections found at such values might be real, so we leave at the next selection steps the task of cleaning the spurious detections that are not physically meaningful. In future improvements of the method, we intend to account for this effect. Finally, we note that the apparent vertical alignment of points for small values of $\hat{A}$ is due to the bin size in the determination of the mean value estimate, which is more noticeable due to the logarithmic scale of the plot.

We find that 83.8% of the windows of data analysed lead to non-detections due to a p-value larger than the 5% threshold and about 6.6% are disqualified because $Q < 0.5$ even though the p-value is within the limit. Eventually, only 9.6% of the sets of AL residuals comply with the requirements and are selected as possible binary candidates by the statistical selection, which

represents 3038 asteroids, of which 40 have estimates from two (or more) different windows of observations.

Some results from different windows are very similar, but most of them show considerably divergent values. The difference between the solutions is typically up to ~35 h in the estimated period and up to ~2 mas in amplitude as can be seen in Fig. 13. At this point, we are not able to decide which window has the most physically meaningful results. Therefore, an asteroid that has been estimated with different periods and amplitudes in different windows at this stage will each have a pair of possible solutions (period, amplitude) considered in the next stage, where the least physically consistent results will be rejected.

4.4. Selection based on physical properties

For the 3038 *Gaia* asteroids whose data passed the statistical detection test above, we move forward to check if the measured wobbling is compatible with the fundamental physical properties expected for an orbiting satellite. Under the hypothesis that the periodic signals correspond to the orbital period, we can exploit known properties measured by other surveys, such as the object size, to infer the range of size ratios and densities that could generate the observed signal. Density, in particular, should be within acceptable ranges, and compatible with the expected mineralogy for a corresponding spectral type.

We start with Kepler's third law of planetary motion

$$\frac{T^2}{a^3} = \frac{4\pi^2}{M_{total}G}, \tag{8}$$

where M_{total} is the total mass of the studied system ($M_1 + M_2$), T is the orbital period of the secondary body around the primary, which we assume to be the same as the wobbling period, and G is
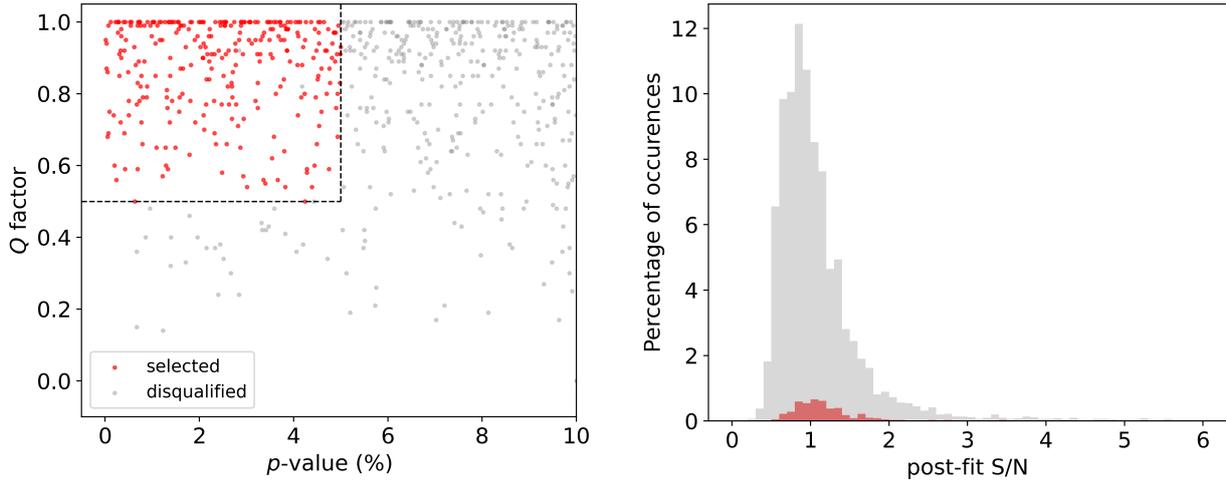
Fig. 10. Resulting distribution of parameters from the 6000 simulations in the noise-only scenario. *Left*: distribution of pairs (p, Q) for the time series simulated. 5391 of the cases have p-value > 0.1 (for these cases the Q factors were not calculated and the corresponding points are thus not shown). The grey dots represent the cases where the data was disqualified as a detection (cases for which $p > 0.05$, $Q < 0.5$, or both) and the red dots represent the detections. The two grey points in the top left square are detections discarded due to $\hat{A} > 20$ mas. *Right*: post-fit S/N distribution for the pairs in the left plot.
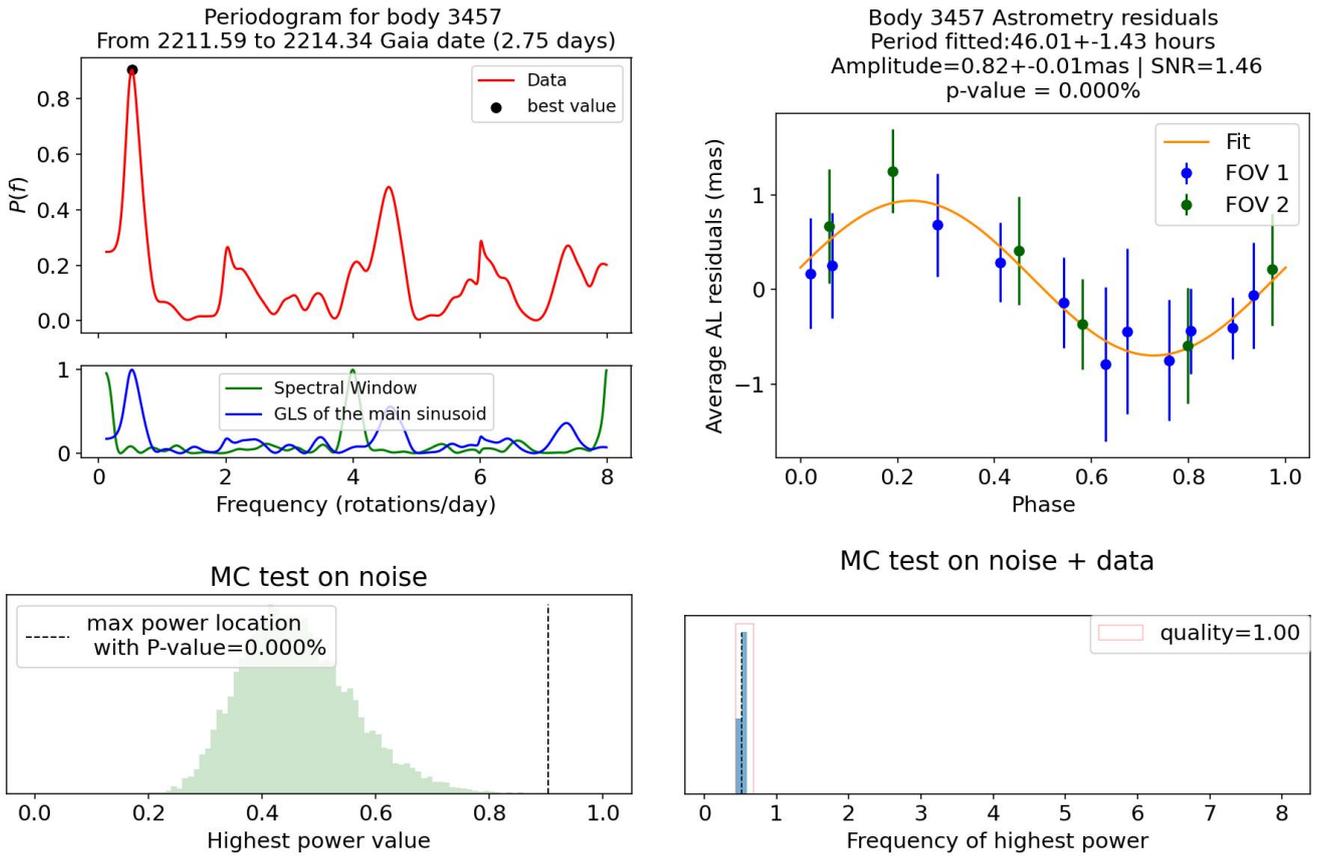
Fig. 11. Visualisation of the output of the main steps of the selection process in the period search for one candidate example ((3457) Arnenordheim). The top left plot shows the periodogram from the fitting of the 17 observations in the window. The peak represents a period of 45.95h and the best fit is shown in the phased data in the top right plot. The distribution of data points not phased is shown in the right plot in Fig. 3. The bottom left plot shows the empirical distribution of GLSP power V (maximum value of the GLSP under $\mathcal{H}_0$) and the vertical dashed line shows the position of the value of v, the power of the peak from the periodogram with the real data, with $p(v) < 0.001\%$. Finally, the bottom right plot shows the distribution of best frequencies from the 5000 MC simulations under $\widehat{\mathcal{H}}_1$). The distribution tends to accumulate around the originally estimated frequency, showing that this detection is robust (here Q is close to 1.)
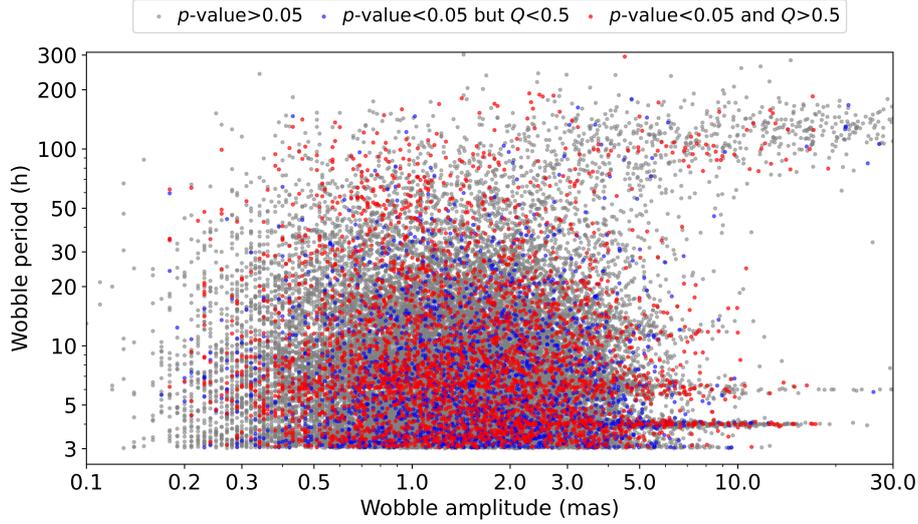
Fig. 12. Distribution of estimated wobble amplitudes and periods obtained from the period search for all objects. In grey are the time series disqualified due to a p-value > 0.05; in blue are the cases disqualified because $Q < 0.5$, but with p-value within the limit of 0.05; and in red are the cases selected as binary candidates from the period search for complying with p-value< 0.05 and $Q > 0.5$.
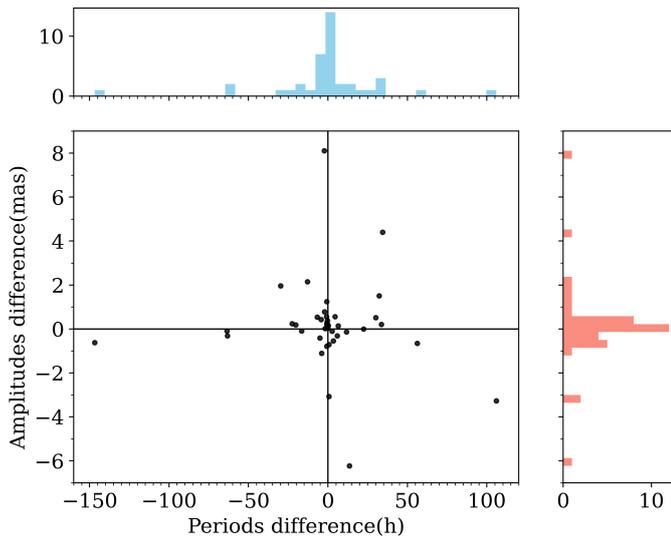
Fig. 13. Distribution of differences between the solutions for the 40 objects with multiple windows of consecutive observations.

the gravitational constant. We manipulate the equation to express it as a function of the mass ratio of the system $q = \frac{M_2}{M_1} = \frac{D_2^3}{D_1^3}$, the diameters of the primary and secondary component (D_1 and D_2, respectively), and the bulk density ρ.

Furthermore, we also consider that the currently observed diameter corresponds to a sphere of surface brightness equivalent to the sum of the binary components, $D^2 = D_1^2 + D_2^2$. Hence, the total mass of the system can be written as

$$M_{\text{total}} = \frac{\rho \pi}{6} \frac{D^3}{(1 + q^{2/3})^{3/2}} (1 + q). \tag{9}$$

Finally, substituting Eqs. (9) and (1) into Eq. (8) and with some manipulations, we obtain a relation describing the density ρ as a function of the mass ratio q:

$$\rho = \frac{24\pi\alpha^3}{GT^2D^3} \left| \frac{(1 + q^{\frac{2}{3}})^{\frac{9}{2}}(1 + q)^2}{q^2(q^{\frac{1}{3}} - 1)^3} \right|, \quad 0 < q < 1. \tag{10}$$

By this expression, the dependency of the density from the mass ratio q can be evaluated, by assuming an estimation of D, once T and α are measured from the astrometric wobbling. We retrieve D from literature using the SsODNet service[5] (Berthier et al. 2023), along with the associated uncertainty. If SsODNet does not provide the diameter for a given object, we estimate it from the absolute magnitude H by taking the mean diameter for the geometric albedos p_v spanning the interval [0.06, 0.2]. We assume a standard deviation of 15% for convenience, considering that without this assumption the errors in the diameter estimates are larger than 50%. This restrictive choice is partially compensated by the generous density interval that is allowed, as explained in the following.

While T is a direct result of our signal analysis, the measured amplitude α is only a minimum value, due to the unknown projection factor of the real wobble in the AL direction. In this respect then, having fixed all other parameters, Eq. (10) provides a *minimum* density as a function of q.

The Eq. (10), as shown in Fig. 14, is a concave function with a minimum corresponding to the maximum wobbling amplitude in Eq. (1). This minimum represents the smallest ρ of the system compatible with the observed wobbling. The uncertainties are obtained with the linear propagation of the errors in D, T and α.

For a first selection, we consider a large density range $\rho \in [0.3, 7]$ g cm^{-3}. We are interested in selecting the possible ranges of mass ratio q corresponding to densities in this interval. Therefore, when the minimum of the density function is <0.3 g cm^{-3}, we obtain two separated intervals delimited by 4 different values of mass ratio (($q_{\text{min1}}, q_{\text{max1}}$) and ($q_{\text{min2}}, q_{\text{max2}}$)). From them, and the associated values of ρ, the corresponding values of the total mass can be computed and used in Eq. (8) to derive two different intervals of possible separation of the binary (($s_{\text{max1}}, s_{\text{min1}}$) and ($s_{\text{min2}}, s_{\text{max2}}$)). This is the case of (7616) Sadako in Fig. 14.

In cases where the minimum density estimate is never smaller than the inferior limit, as for (6261) Chione, we get a single interval of q and possible separations. Eventually, if the function has a minimum value larger than the adopted upper

---

[5] https://ssp.imcce.fr/webservices/ssodnet, queried through its rocks python-interface (https://github.com/maxmahlke/rocks).
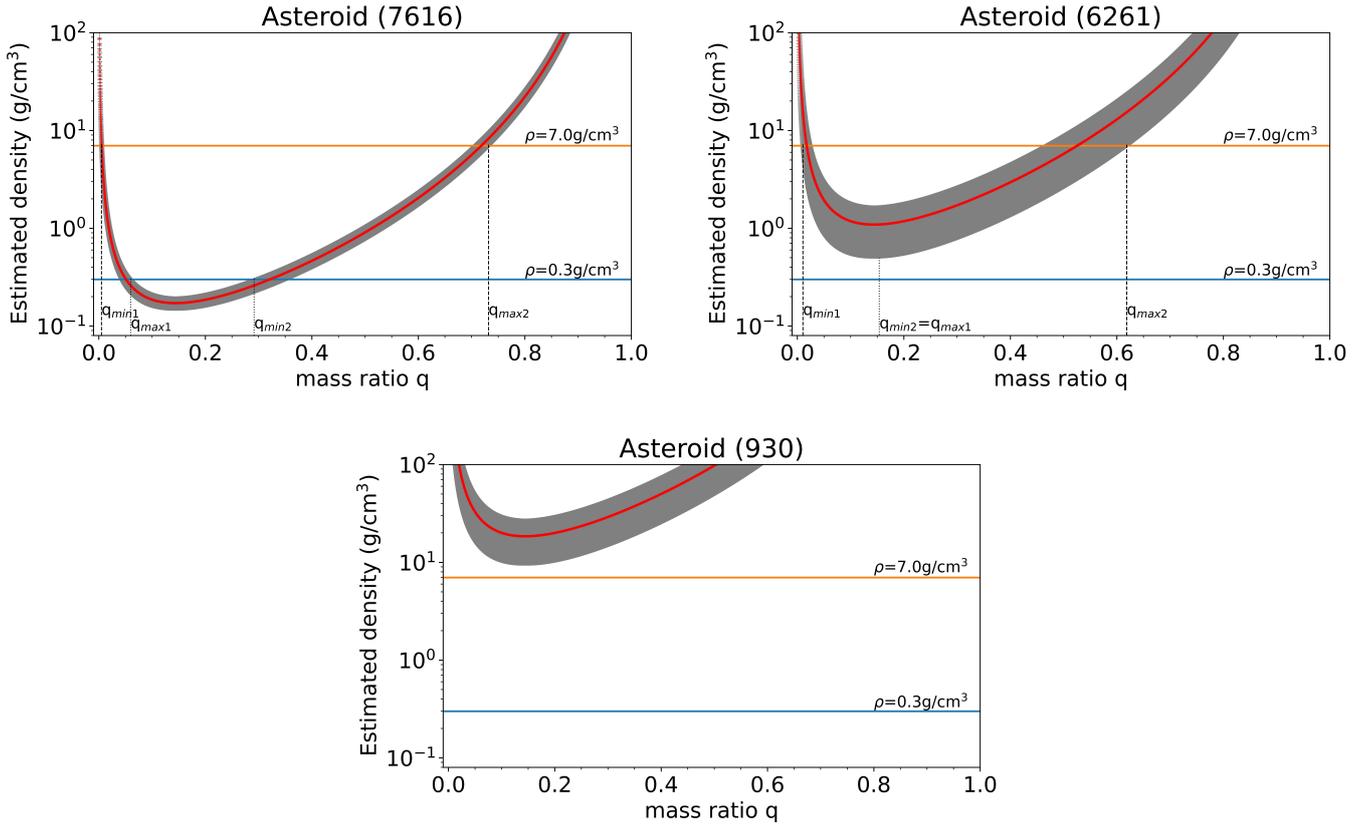
Fig. 14. Examples of the bulk density ρ as a function of the mass ratio q for two asteroids, (7616) Sadako (top left) and (6261) Chione (top right). The red line indicates the nominal value. Shown in grey is the error range on the density. The blue and orange horizontal lines represent physically reasonable limits. q_{min1}, q_{max1}, q_{min2} and q_{max2} mark the corresponding extremes of the mass ratio values. The bottom plot shows the case for (930) Westphalia, eliminated in the density filtering because the minimum estimated density is higher than the chosen range of bulk density.

threshold, the wobble signature is considered not compatible with a binary object as this would require an excessive density. By adopting the density threshold mentioned above, 934 objects are selected and their possible ranges of q and s have been computed.

A further step consists in considering the smallest separation, for each candidate, compatible with an orbiting companion. We find that for some of the objects, the range of separations partially overlaps with the Roche limit[6], computed by assuming the same density for the two components. Binary systems inside the Roche limit are unlikely to exist, and the detected wobble could instead be caused by a very irregular shape. Of course, more accurate physical data, in particular diameter measurements, could change the results for some objects and put them beyond the Roche limit. To be conservative, for this first search, we have decided to reject candidates that have separation intervals overlapping, even partially, with the Roche limit. With this selection, we remain with 362 wobble measurements for 358 objects.

As an attempt for a further verification, we considered the largest selected objects in the table having a known shape in the DAMIT[7] database (Durech et al. 2010) to compute the distance between photocentre and centre of mass, namely for the asteroids (476) Hedwig (Marciniak et al. 2023), (487) Venetia (Marciniak et al. 2018), (516) Amherstia (Durech et al. 2009; Hanuš et al. 2011), (532) Herculina (Kaasalainen et al. 2002;

Hanuš et al. 2017), (542) Susanna (Hanuš et al. 2021), (556) Phyllis (Marciniak et al. 2007), (699) Hela (Marciniak et al. 2012), (1200) Imperatrix (Ďurech et al. 2020), (1639) Bower (Ďurech et al. 2019), (2219) Mannucci (Ďurech et al. 2020), (70045) 1999 DA5 (Ďurech & Hanuš 2023). We simulated the photocentre shift due to the illuminated shape alone, at all epochs used for our period search, by using Lambert's emmition and Lommel-Seeliger's scattering laws (Lambert 1760; Pedrotti et al. 2017; Fairbairn 2005) and the corresponding illumination geometries. Then, its projection on the AL direction was used to look for correlations with the observed residuals. This test reproduces the one reported in Tanga et al. (2023) for (21) Lutetia.

We repeated the test with the multiple pole and shape solutions available for each object, also checking the possible role of a rotation phase shift concerning the nominal rotation origin. Our results show that the amplitude of the photocentre displacement could be correlated to the residuals for (487) Venetia, (1200) Imperatrix (for one of the two pole solutions available) and (2219) Mannucci (in one of the two observation windows that we consider), as shown in Fig. 15. It is also anti-correlated for (532) Herculina and (556) Phyllis. In the other cases, no correlation shows up.

This preliminary test is very sensitive to several details such as shape and pole coordinate errors, scattering law, presence of concavities, etc, and should be taken with caution. However, we notice that the order of magnitude of the wobble amplitude due to a (single) shape would be about the same as the signature

---

[6] $d_{\mathrm{Roche}} = 2.44\, R_{\mathrm{primary}}(\rho_{\mathrm{primary}}/\rho_{\mathrm{secondary}})^{1/3}$.
[7] https://astro.troja.mff.cuni.cz/projects/damit/
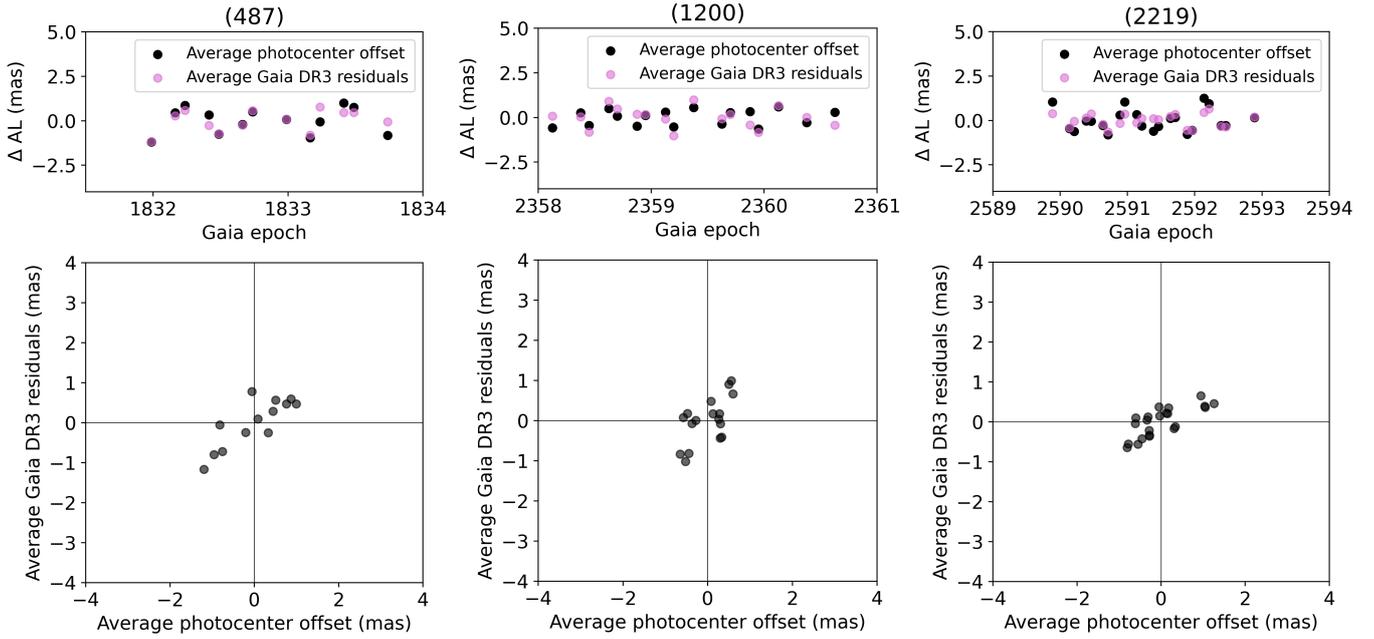
Fig. 15. Comparison between the average DR3 astrometric residuals and estimated average photocentre offset for candidates (487) Venetia, (1200) Imperatrix and (2219) Mannucci. The top plots show the variation of the data over time in the along-scan direction, while the bottom plots show the correlation between the two sets of data.

produced by a satellite for large asteroids (separations of ~200 mas or more, spatially resolved by *Gaia*). However, we notice that the amplitude of the photocentre shift is not dissimilar from our residuals. In a synchronous system (primary rotation equal to secondary orbital period) they could even add up with the same or opposite sign, depending on the phase angle of illumination. Only in this case, when the sign is opposite, this could result in the observed anti-correlation which would then be evidence supporting the satellite's presence.

All these considerations point in the direction of not excluding the largest objects remaining on our list. They certainly deserve more accurate characterisation by future observations and more extensive, complex modelling efforts when more data becomes available.

4.5. Constraints from taxonomy

The average density value for each taxonomic class can help to further constrain the physical properties of objects in our selection. We retrieved the taxonomic class for all objects in the list of preliminary candidates using SsODNet. Of the 358 candidates selected from the previous steps, 264 have their classes determined, corresponding to one of the complexes for which an average density is provided in Carry (2012).

By using the average density for the taxonomic class, we can repeat the analysis illustrated in Sect., recompute the permitted intervals of possible physical parameters, and refine the selection of the sample. We reached a total of 156 candidates. In addition, since large real wobbles tend to produce high $\widehat{S/N}$ (cf. Sect. 4.2) we flag as best candidates those with $\widehat{S/N} > 1$.

The objects of our sample on which taxonomy-derived density constraints are applied are presented in Fig. 16, which shows the estimated radius of the primary as a function of the allowed separations (in units of the radius of the primary) for the 156 objects. The purple and pink points represent the first and second intervals of possible mass ratio values. Those with dark borders

are the best candidates with $\widehat{S/N} > 1$, and the points with grey borders are the candidates with $\widehat{S/N} < 1$ that are not in the list of the best cases but still passed through all the filtering steps.

5. The final sample of candidate binary asteroids

The total outcome of 358 objects from this first search of astrometric binary candidates appears in the full version of Table 3 in CDS. It includes 6 known binaries, 67 of which (highlighted) present results with the density constrained by taxonomy and $\widehat{S/N} > 1$.

Table 4 summarises the number of candidates after each step of our selection process. We select the 30 030 asteroids in DR3 with consecutive observations in a short time. From the statistical selection, we obtained 3038 preliminary candidates that presented *p*-value< 0.05 and $Q > 0.5$. The validation on the physical ground and the selection based on known taxonomic type further restricted this list to 156 objects presenting characteristics compatible with binary systems. Most of them have an estimated $\widehat{S/N} < 1$, as the regime of the scarceness of the data and potentially low wobble amplitudes make detectability difficult. Therefore, we can consider that those with $\widehat{S/N} > 1$ among the 156, are the best characterised. The distribution of their parameters concerning the statistical thresholds of period determination is shown in Fig. 17.

In the final list of 358 candidates, there are also 31 objects whose estimated period is not very reliable. In fact, their values reach is towards the highest ones in the allowed range of the GLSP (reaching three times the length of the observation window).

An example is (1509) Esclangona, a known binary with an estimated period of ~410 ± 120 h (assuming a bulk density of $\rho = 2.0 \pm 1.0 \, g \, cm^{-3}$, Marchis et al. 2012) well beyond the periodogram range (left panel of Fig. 18). A poorly defined best period is found, in a plateau of the periodogram where a whole
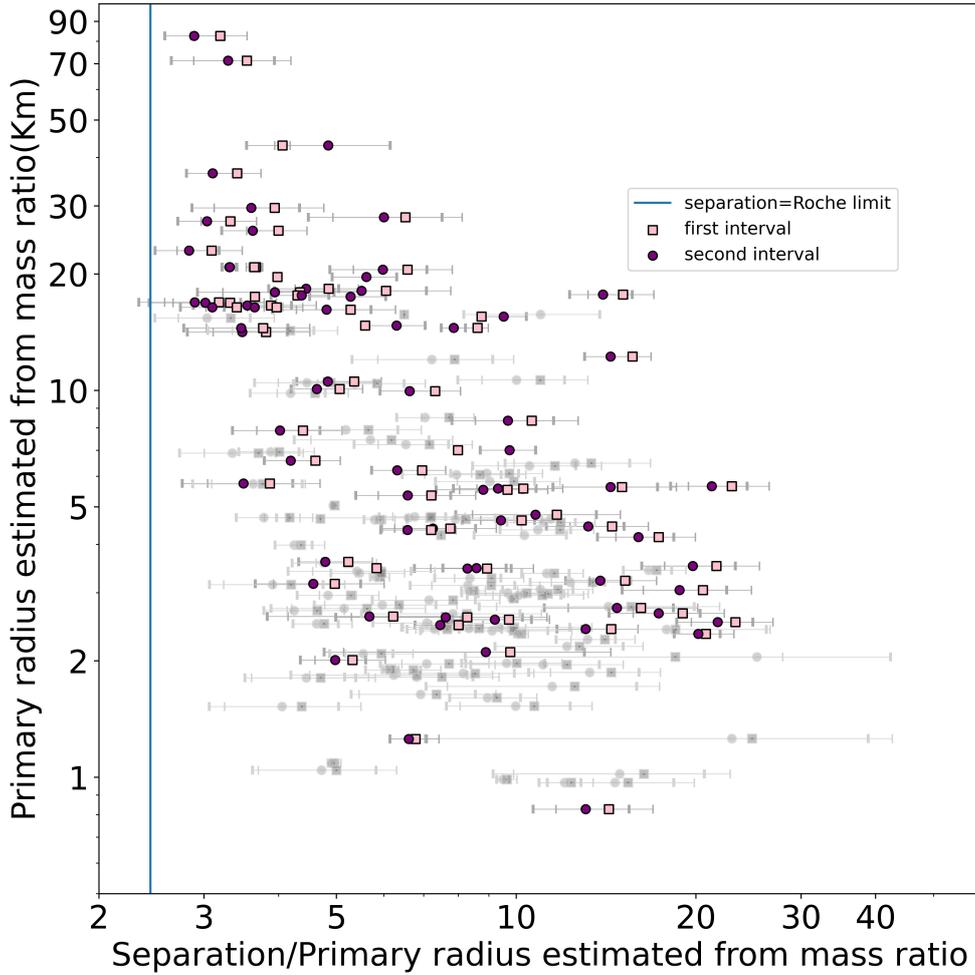
Fig. 16. Plot of the radius of the primary as a function of the separation intervals normalised with the Roche limit for fluid satellites, for the sample of 156 asteroids with known taxonomy. The squares represent the mean value of the first mass ratio interval estimated and the circles represent the second interval. The coloured symbols with dark borders are the objects from the taxonomic filtering with $\widehat{S/N} > 1$, and the grey symbols are those with $\widehat{S/N} < 1$.

range of values is equally possible. Going towards the extreme of the highest periods tested, the fitting function to the residuals (covering 84 hours right panel) tends to locally approach a straight line, resulting in a whole range of similar fitting quality.

It should be noted that the clear trend could indicate a genuine long period. Hence, the data may contain part of a periodic signal that we cannot consistently estimate, as this would require a longer data sequence. In this case, and the other ones resulting in a similar periodogram, we are only able to give a minimum boundary and estimate that the period should be larger than the length of the observation window. Candidate binaries in this category are identified by an orange background in the final list of candidates in Table 3.

Additionally, we notice that the method seems to favour the selection of objects presenting periods longer than 10 h, which is perhaps the sign that high frequencies not captured by the constant plus sinusoid model of the GLSP are present in the time series. By filtering the physically meaningful candidates we eliminate most of the cases with p-values $> 2\%$ and $Q < 0.9$ and the best candidates are mostly those with very low p-value and Q close to 1.

The comparison between Figs. 19 and 10, showing the heavier right tail of the $\widehat{S/N}$ distribution obtained from *Gaia* data, further strengthen the evidence that our candidates are indeed binary asteroids. Two of the best candidates in the list, (2219) Mannucci and (542) Susanna present matching wobble periods from two different time windows, separated in time by weeks or months. At such large intervals, the orientation of the

wobbling motion relative to the scan direction can be different (as confirmed by the different wobble amplitude), thus making these measurements independent. The fact that the two estimated wobble periods of each object match is another sign that our method of period search is rather robust.

The bodies selected as best binary candidates, including the known binary asteroids, are in large majority main-belt asteroids (Inner MB: 11, Middle MB: 26, Outer MB: 22), but 3 Phocaea, 3 Mars-Crosser, 2 Hungaria, 1 NEA Amor, and 1 Trojan are also present. All of our best candidates present separations <200 mas, below the limit at which the onboard Video Processing Unit of *Gaia* treats sources separately.

5.1. Known binaries

Our results contain 6 known binaries: (1509) Escanglona, (5817) Robertfrazer, (2871) Schober, (4337) Arecibo, (317) Roxane and (18301) Konyukhov. Such a small number of known binaries is not unexpected, since our filtering and statistical validation reduce a large sample to a very small population. For each binary, the probability of showing an astrometric signature in *Gaia* DR3 is not very high, as it requires the appropriate orientation of the wobble motion concerning the scan motion, during the periods of consecutive observations.

Table 5 contains the parameters described in the literature and obtained from observations of the six mentioned known binary systems. These six known binaries offer in principle the possibility of validating our approach, but in practice, other

Table 3. List of estimated parameters for the *Gaia* astrometric asteroid binary candidates.

ID	Wobble period(h)	Wobble amplitude (mas)	Separation (km) 1st interval	Mass ratio 1st interval	Separation (km) 2nd interval	Mass ratio 2nd interval	Taxonomy correction	$\widehat{S/N} > 1$
			KNOWN BINARIES					
317	127.56 ± 64.57	2.38 ± 1.69	(5.08 ± 4.79)e+02	(1.21 ± 1.11)e-02	(5.27 ± 5.27)e+04	(7.36 ± 2.63)e-01		*
1509	> 84.06	–	–	–	–	–	–	–
2871	20.69 ± 4.63	0.45 ± 0.07	(1.56 ± 0.18)e+01	(1.18 ± 0.25)e-02	(1.43 ± 0.16)e+01	(6.00 ± 0.38)e-01	*	
4337	35.93 ± 6.36	0.61 ± 0.07	(6.57 ± 0.66)e+01	(2.40 ± 0.40)e-03	(5.99 ± 0.62)e+01	(8.18 ± 0.17)e-01	*	
5817	> 144.11	–	–	–	–	–	–	–
18301	139.82 ± 70.02	3.10 ± 1.92	(7.61 ± 1.38)e+01	(9.78 ± 3.19)e-03	(6.95 ± 1.24)e+01	(6.35 ± 0.56)e-01	*	*
			OBJECTS WITH TWO SOLUTIONS					
542	10.02 ± 0.02 / 10.04 ± 0.04 / 14.37 ± 0.19	0.74 ± 0 / 0.56 ± 0.01 / 0.48 ± 0.01	(7.64 ± 0.20)e+01	(2.30 ± 0.10)e-03	(6.90 ± 0.19)e+01	(8.20 ± 0.05)e-01	*	*
2219	13.56 ± 0.36	0.74 ± 0.06	(5.59 ± 1.10)e+01	(2.50 ± 0.80)e-03	(5.08 ± 0.99)e+01	(8.16 ± 0.33)e-01	*	*
9573	63.47 ± 13.54 / 57.59 ± 3.56 / 10.33 ± 0.13	0.68 ± 0.05 / 0.99 ± 0.05 / 0.56 ± 0.05	(3.37 ± 0.57)e+01	(1.08 ± 0.33)e-02	(3.08 ± 0.50)e+01	(6.18 ± 0.53)e-01	*	
10766	73.57 ± 9.51	0.87 ± 0.55	(1.17 ± 0.24)e+02	(4.99 ± 1.70)e-03	(1.07 ± 0.22)e+02	(7.36 ± 0.48)e-01	*	
			OBJECTS WITH ONE SOLUTION					
53	9.08 ± 0.01	1.42 ± 0.01	(2.39 ± 0.96)e+02	(2.50 ± 1.50)e-03	(2.54 ± 1.25)e+02	(8.34 ± 0.76)e-01		*
203	23.89 ± 0.27	0.88 ± 0.01	(2.32 ± 0.03)e+02	(1.00 ± 0)e-03	(5.94 ± 3.71)e+02	(9.35 ± 0.39)e-01		*
217	12.70 ± 0.04	0.88 ± 0.02	(1.84 ± 0.82)e+02	(2.75 ± 1.75)e-03	(2.19 ± 1.25)e+02	(8.33 ± 0.88)e-01		*
238	9.85 ± 0.01	2.23 ± 0.02	(3.23 ± 1.07)e+02	(2.00 ± 1.00)e-03	(3.92 ± 1.96)e+02	(8.59 ± 0.66)e-01		*
247	12.10 ± 0.06	0.55 ± 0.02	(1.46 ± 0.05)e+02	(1.00 ± 0)e-03	(4.41 ± 2.38)e+02	(9.53 ± 0.24)e-01		*
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

Notes. The grey lines represent the objects with density constrained by taxonomy, and the magenta lines indicate the best candidates accounting for the taxonomic correction and a strong detection with $\widehat{S/N} > 1$. The objects with orange ID represent those that were selected as binary candidates but with periods larger than the length of the observation window, therefore the period given as the minimum is the length of the observation window used. This is an extract of the full table, which is available at the CDS.

Table 4. Number of binary candidates remaining after each stage of the systematic search.

	Number of candidates
Asteroids in DR3	156 801
Candidates with a reasonably long sequence of observations	30 030
Candidates' time series with detection from the statistical selection	3038
Candidates' time series with physically meaningful density estimate	934
Candidates with separation beyond the Roche limit	358
Candidates from the taxonomic filtering	156
Best candidates with $\widehat{S/N} > 1$	67

limitations appear. As mentioned previously, (1509) Esclangona cannot be exploited for the comparison, as it has a small satellite with an orbital period well beyond the reach of our data sample. The most probable periodicity given by our periodogram, 123.71±11.09 hr, is roughly between 2 and 4 times smaller and should be considered as a lower limit (Sect. 5).

In the case of (5817) Robertfrazer the data from *Gaia* DR3 presents two windows of observation that we use, but the result obtained from one of them does not pass the physical validation process and is therefore discarded. So, for the remaining window, we also find a matching size ratio ($k = D_2/D_1$), however, our period is more than 5 times larger than reported (~28.862 hr), and again close to the upper limit of the frequency range of the GLSP.

Interestingly, however, for both these incorrect matches the apparent wobble signal is strong, with a $\widehat{S/N}$ of 6.09 for Esclangona, and 2.64 for Robertfrazer, and the wobble amplitudes are significantly high (5.09 ± 0.54 mas for Esclangona and 1.61 ± 1.19 mas for Robertfrazer), especially for Esclangona. A possible alternative interpretation could be the presence of another satellite in the system, and future *Gaia* astrometry will probably help to clarify which scenario is more probable. In both cases, we estimate that the interest in selecting these objects by our procedure remains.

For (2871) Schober, the size ratio from the literature matches our results. We find solutions from two different time windows of observation, with wobble periods of 20±4 hr and 50±11 hr: the first can be considered an alias emerging from the data at about

Table 5. Parameters of the known binaries from the literature.

ID	a (km)	D_1 (km)	D_2 (km)	Size ratio k	T (h)	Misc.	Ref.
(1509)	140	–	4	–	–	D_{eff}=9 ± 1 km	Merline et al. (2003)
Esclangona	<132	8.0 ± 0.8	4.0 ± 0.7	0.5 ± 0.1	–	type Sw	Marchis et al. (2012)
(4337)	–	24.4 ± 0.6	13.0 ± 1.5	0.533 ± 0.063	–	Occultation	Gault et al. (2022)
Arecibo	49.9 ± 1.0	–	–	0.186 ± 0.022	32.9728	$\rho \approx 1\,\mathrm{g\,cm^{-3}}$	Tanga et al. (2023)
(2871) Schober	–	–	–	>0.28	42.47 ± 0.02	–	Benishek et al. (2023a)
(5817) Robertfrazer	–	–	–	>0.19	28.862 ± 0.007	–	Stephens & Warner (2020)
(317) Roxane	245	19	5	–	–	–	Merline et al. (2009)
(18301) Konyukhov	–	–	–	0.26±0.02	35.7±0.1 or twice	–	Benishek et al. (2023b)

Notes. a is the binary separation; D_1 and D_2 are the primary and secondary bodies diameters, respectively; k is the size ratio D_2/D_1; T is the orbital period of the secondary around the primary body.

double the frequency due to the irregular sampling from *Gaia*, while the second results in intervals of separation that overlap partially with the Roche limit and, therefore, it was rejected. Both windows are compatible with the value reported in the literature, obtained by photometry (Benishek et al. 2023a).

In the case of (317) Roxane, we find that our results also match well with the size ratio from the literature. We also estimated the orbital period of the secondary based on the separation, as was made for Esclangona, and we found that our period estimate is about half of the expected value. The observation window from *Gaia* DR3 that was selected to perform the search is shorter than the expected period of the secondary, which once again limits the extent of our period search and invisibilises the method to find the correct period.

For (18301) Konyukhov we find that the observed size ratio of 0.26 (≈0.017 in mass ratio, Benishek et al. 2023b) is very close to the upper limit of the first interval of mass ratio of 0.013, translated into an upper value for size ratio of approximately 0.24. Regarding the period, our Monte Carlo procedure to estimate the confidence interval on this value reveals that the distribution of the estimated periods (cf. Sect. 4.1) is nearly bimodal, with two modes on 70.1 h and 140.6 h. As a consequence, the procedure provides a very large confidence interval (±70.02h), which reflects the possibility that the true period can actually be close to either of the two values.

For (4337) Arecibo we find that our estimated period is equivalent within the uncertainty to other sources (Tanga et al. 2023; Ďurech & Hanuš 2023). This binary has two different size ratio values from the literature. Stellar occultations Gault et al. (2022) provide $k_{\mathrm{occ}} = 0.533 \pm 0.063$, while the fit to both *Gaia* astrometry and occultations Tanga et al. (2023) results in $k_{\mathrm{Tanga}} = 0.35\ k_{\mathrm{occ}}$, with indications of a probable ellipsoidal shape of the components. Using the best value of surface–equivalent diameter from SsODNet (19.746±0.324 km) and our estimates for wobble amplitude (0.61±0.07 mas) and wobble period (35.9284±6.3600 h), we obtain results close to the parameters of the system previously published (lower size ratio solution in Fig. 20). The remaining difference with respect to Tanga et al. (2023) arises from the different estimates of the wobble amplitude and period, and from the fact that we do not consider here additional constraints such as those coming from the observation of stellar occultations.

Table 6. Taxonomic classes grouped as complexes (Σ).

| Classes	Σ
S	S
Q	Q
V	V
C, Ch, B, D, P, Z	C
X	X, E
K, L, M	M

5.2. Comparison to the known population

Our candidate binaries can be compared to the existing population. As our observational biases are different from other techniques, the goal is not to reproduce the same distribution. In principle, we can discover asteroids with satellites whose dynamical properties are unusual, in particular in a domain of intermediate sizes (that we can conventionally define between 20 and 100 km) where traditional techniques have very limited capabilities of discovery. On the other hand, we can also reasonably expect that some of our candidates are similar, in physical and/or dynamical properties, to the binaries known so far.

A first comparison between the astrometric binary candidates and the known binary population is based on taxonomy. Our goal is to verify if an excess of S-type binaries is present in our sample, by an approach similar to Minker & Carry (2023). We start by grouping the taxonomic classes in large complexes, as shown in Table 6. Notably, P-type asteroids are in the C complex, not the X complex. The reference sample of asteroids was constructed based on a magnitude limit of absolute magnitude $H < 16$ and within ranges of $a \in [2, 3.5]$ au, $i \in [0°, 30°]$ and $e \in [0, 0.5]$ roughly matching the distribution of the sample of binary candidates.

The reference sample is selected by choosing 30 asteroids around each binary candidate, from a partition of the 4D (a,e,i,D) parameter space (see Minker & Carry 2023 for details).

Figure 21 shows the comparison between the reference sample and our binary candidates with $D < 20$ km, which represent most of our candidates. The S-complex taxonomic types present a fraction above 1, which means that there is an over-representation of objects of this complex in our sample of
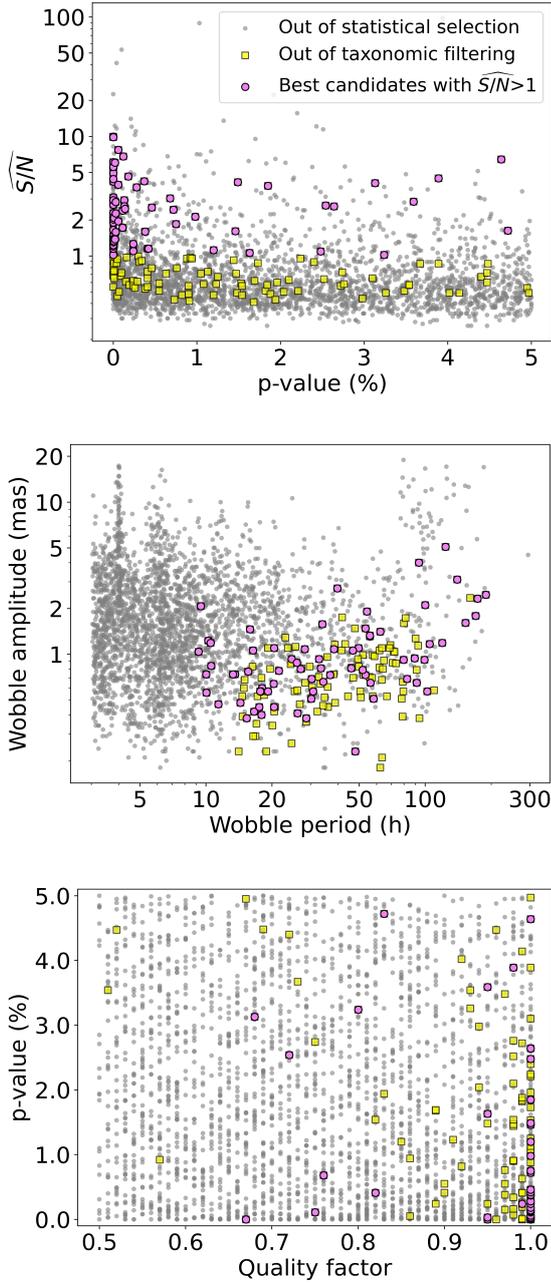
Fig. 17. Distribution of $\widehat{S/N}$ vs. *p*-value (top); estimated wobble period vs wobble amplitude (center); and *p*-value vs. quality factor (bottom) for the 3038 objects obtained from the statistical selection (in grey), the 156 objects that survived the taxonomic filtering (in yellow), and the 67 best candidates that passed all tests and have $\widehat{S/N} > 1$ (in pink).

small binary candidates. On the other hand, our small candidates show a deficit of C-complex taxonomic types in comparison with the background asteroid population.

The results from this analysis are in great agreement with Minker & Carry (2023), where it is shown that even though chondritic asteroids are the most common type in the Solar System, there is a preferential presence of S-type over C-type, among binary objects at small diameters. Thus, we can conclude that the taxonomic distribution of our candidates is comparable to the one of the previously known population.

Taking a closer look at the sub-sample with 30<D<50 km, we find no dominance of any taxonomic type (Fig. 22), suggesting a disconnect between the small- and medium-sized binary asteroid populations. However, statistical comparisons to a known population are not possible, as the only known binary asteroid in this size range is the system (243) Ida – Dactyl. Interestingly, it was discovered by a spacecraft during a fly-by (Chapman et al. 1995). This could indicate that un-conventional techniques such as the astrometric method are necessary to find similar binaries. The limited presence of large (D>50 km) asteroids in our sample of binary candidates eliminates the possibility of continuing this taxonomic analysis at larger sizes.

In order to compare the dynamical properties of our candidates to the known population, we first present (Fig. 23) the distribution of the primary rotation period as a function of the estimated primary diameter. The rectangles in the plot represent a raw approximation of the known binary asteroid groups (Pravec & Harris 2007; Pravec et al. 2010; Margot et al. 2015). Group L is for fast-rotating and large objects. Group A represents small primary asteroids with fast rotation; B, systems with small primaries but size ratios $D_2/D_1 \sim 1$ that rotate slower as the primary diameter increases.

The first striking feature of our sample is the partial overlap with the pre-existing population discovered by photometry (Group A), which clearly appears below ∼20 km. Some candidates also partially overlap group B, but their mass ratios are small in general (of the order of 0.1). At the other extreme of the size range, binary candidates from our search do not match group L (besides one). This is expected, as the wobbling signal from large objects should not be dominated by a small satellite, but by the photocentre displacements on the shape of the primary alone. Coherently with this interpretation, some potential candidates in this region are excluded due to the separation being too small relative to the Roche radius. Eventually, intermediate sizes (20–100 km), clearly underrepresented in the known sample, are present for our astrometric candidates, directly confirming the new potential of this technique.

The comparison of satellite orbits and sizes is more difficult and intrinsically ambiguous, as our analysis introduces two families of solutions, corresponding to two different possible ranges of mass ratio, as shown in Fig. 14. The estimated mass ratio intervals for the 67 selected best binary candidates, and their respective primary rotation periods, are shown in Fig. 24. The two mass ratio intervals are more or less separated depending on the minimum estimated density for the system. At this stage, we have no criteria to favour a solution over the other. However, we notice that objects of intermediate size populate preferentially extreme values of the mass ratio. This could imply that binaries of similar-sized components could exist in that population.

The analysis of the ratio of the wobble (orbital) period to primary rotation period, reveals the presence of synchronous objects and, to a minor extent, of objects in a 1:2 period ratio (Fig. 25). Among the known main belt binaries, almost all those with $D_s/D_p > 0.5$ are synchronous, as the tidal coupling is more efficient for large satellites. Applied to our sample, this evidence would imply that large mass ratios have a higher probability of showing orbit–spin coupling. Arecibo, at $D_s/D_p > 0.53$ and synchronous, clearly falls in this category. Periods on the 1:2 ratio could simply represent an alias and could be considered synchronous too. Interestingly, intermediate-size objects are in large majority in this category.
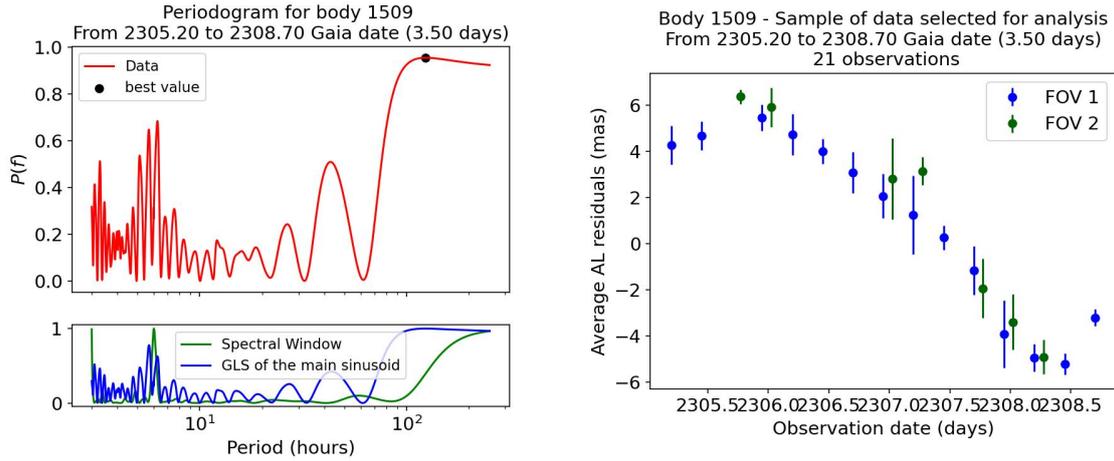
Fig. 18. (1509) Esclangona residuals distribution (*right*) and the periodogram from the fitting of the 21 data points in the window (*left*).
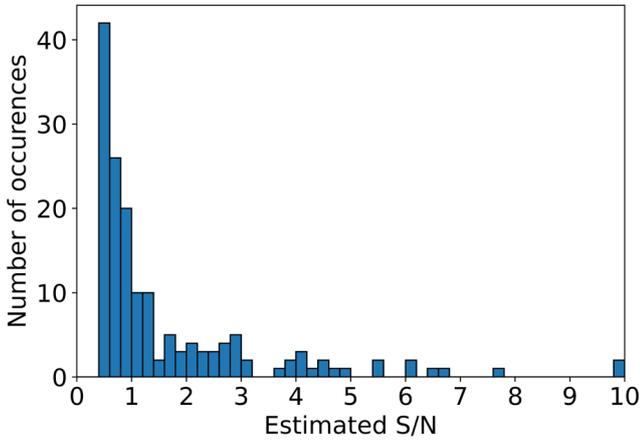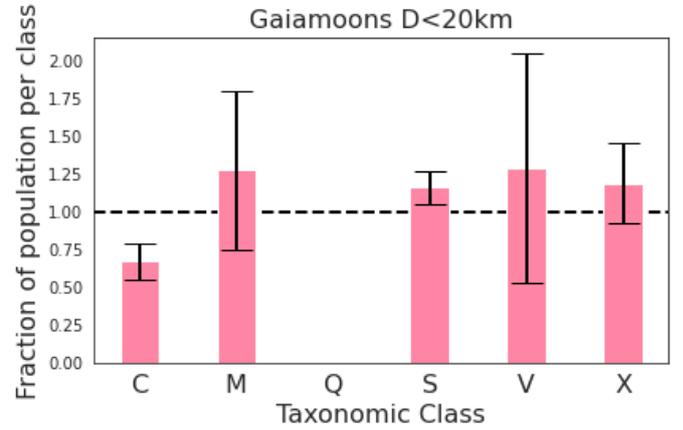

Fig. 19. Distribution of $\widehat{S/N}$ for our list of 156 binary candidates selected from the taxonomic filtering. The 67 best candidates are those with $\widehat{S/N} > 1$
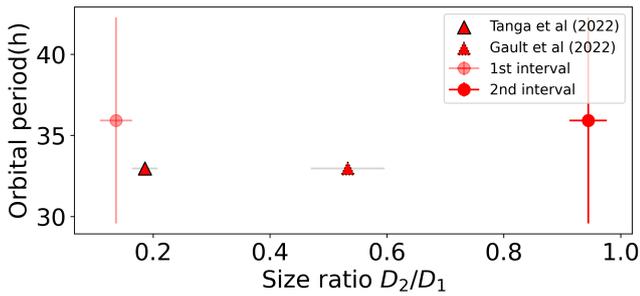

Fig. 20. Comparison between the parameters for the Arecibo system. The triangle with black edges and grey error bars represents the values from literature (Table 5), while the circles represent the first (transparent) and second intervals of size ratio estimated from our method using the same wobble amplitude as in Tanga et al. (2023).

Fig. 21. Comparison of the fraction of small objects with $D < 20$ km in each taxonomic complex between the reference sample of background asteroids and the population of our astrometric binary candidates. A taxonomic type with no column indicates that asteroids of that type are present in the reference population but not in the group of binary candidates.
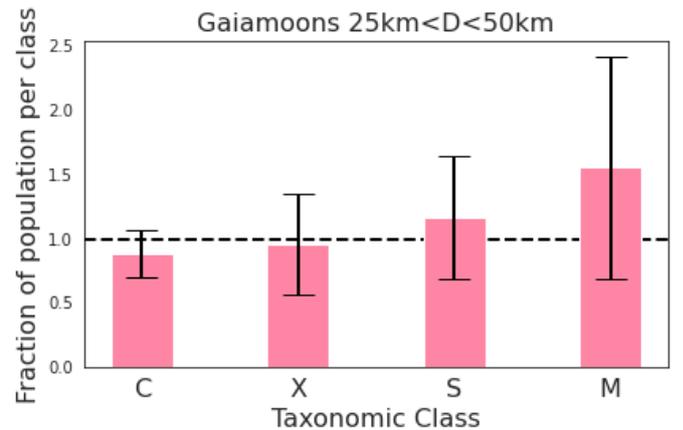

Fig. 22. Fraction of binary candidates with $30 < D < 50$ km in each taxonomic complex, with respect to the reference sample of background asteroids.

6. Discussion

As explained above, this first attempt of binary search in *Gaia* DR3 astrometry is performed by adopting some conservative criteria and an approach that does not take into account all the observations available for each asteroid. Other choices are possible, such as those described in Segev et al. (2023), where a solution of the complete inverse problem leading to satellite orbits, and involving all the AL residuals of each object, is proposed. Their search is not successful and the main explanation

given is the rejection mechanism of astrometric outliers implemented by the validation pipeline in DR2 (Gaia Collaboration 2018).
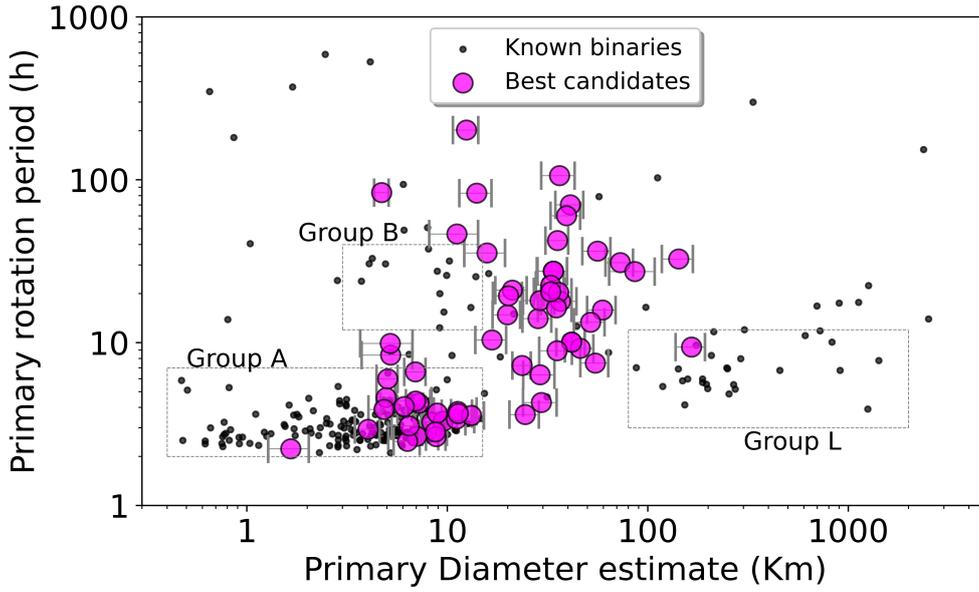
Fig. 23. Plot of the primary rotation period vs. the estimated primary diameter for the best candidates. The black dots are known binaries in the shown intervals and the magenta circles are the best binary candidates obtained.
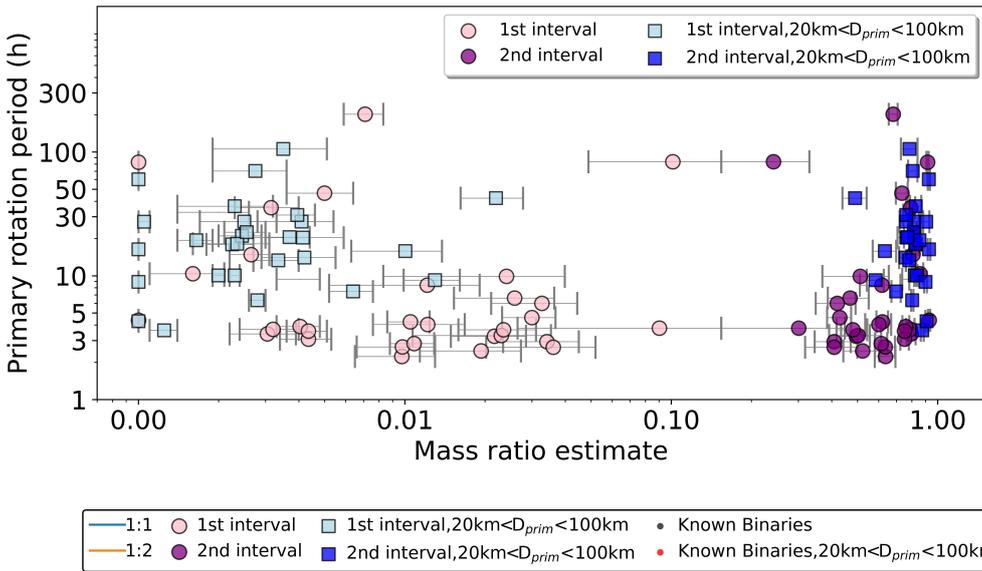
Fig. 24. Plot of the primary rotation period vs. the estimated mass ratio intervals for the 67 selected candidates. The lighter and darker coulored circles represent, respectively, the first and second intervals of possible mass ratio values for the best binary candidates.
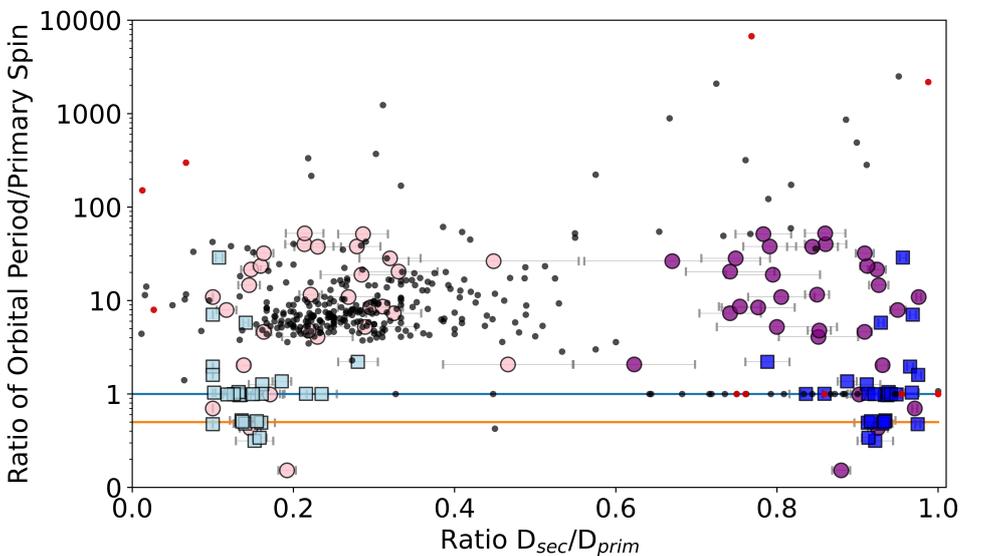
Fig. 25. Plot of the ratio of the estimated orbital period (or wobble period) of the secondary over the photometric rotation period of the primary vs. the estimated size ratio of the binary system. The black dots represent all of the known binaries and the red dots highlight those with intermediate sizes (20 km < D_{primary} < 100 km). The light and dark purple circles represent the average value of the best-estimated size ratio intervals for objects smaller than 20 km and the light and dark blue circles represent the larger objects. The blue line shows the condition where the estimated orbital period of the secondary is almost equal to the rotation period of the primary object, representing a 1:1 period ratio, while the orange line represents the 1:2 period ratio.

However, our results show that periodic fluctuations in the residuals are compatible with the expected amplitude and orbital periods of satellites. We thus think that the missed detection in the cited article is due to three main reasons: the strong increase in the time range and number of asteroids of *Gaia* DR3 concerning DR2 (by factors 1.5 and 11, respectively); the improved astrometric quality of DR3; and, of lesser importance, the possible presence of systematic errors in isolated astrometric measurements that are included in the analysis.

Concerning the first argument, by simple statistical scaling we can estimate that, if DR2 astrometry were good enough to reveal our best candidates, only a few of them (<10) would have been found. However, the probability of detection is not a linear function of the time range, as a reduction in the observed arc can have a strong adverse impact on the quality of the orbital fit and, in turn, on its residuals.

The fact that DR3 improves astrometric quality in a significant way over DR2, is well visible not only in the overall statistics but also in specific cases such as (4337) Arecibo, whose residuals show a clear trend in DR3 (Tanga et al. 2023) but less clearly in DR2 (Fig. 19 in Segev et al. 2023), over the same time span. Hence, this is the first work that provides a large list of binary candidates obtained from astrometry and crosses dedicated statistical and physical selection criteria. At this stage, without independent confirmation of the binary status and a clear discrimination between the two solutions that we obtain for each candidate, the impact of our possible discoveries on formation scenarios cannot be very specific. However, some general considerations can be attempted.

First, the intermediate size range (20–100 km) where very few binaries are currently known (only ~5%n of the total population) is probably the most interesting, and the one having a stronger potential impact. If they are synchronous and have large mass ratios, a well-known prototype at large sizes can be considered the asteroid (90) Antiope (size ratio ~0.95, diameter ~88 km, Merline et al. 2000). Even though there are several theories for the formation of the Antiope system (Pravec & Harris 2007; Weidenschilling et al. 2001), observations support the hypothesis that the Antiope system objects were formed simultaneously from the disruption of a parent body(Descamps et al. 2009; Marchis et al. 2011).

With a similar mechanism, as shown in numerical simulations (Doressoundiram et al. 1997; Durda et al. 2004), Escaping Ejecta Binaries (EEBs) result from the creation of collisional fragments at low relative velocity, that remain gravitationally bound. They preferentially have highly eccentric orbits, that could be circularised by tidal forces in their secular evolution. They should not have particularly fast primary rotation as they are not created by fission or fragment ejection. However, the current sample of known binaries, if simulations are correct, has a clear under-representation of this category. The only good candidate is (317) Roxane with its satellite Olympias (Drummond et al. 2021). The population of EEBs could also be compatible with binary Mars impactors revealed by Vavilov et al. (2022). Once validated, our sample could clearly contribute to revealing many details about the presence, origin and evolution of EEBs.

We also notice the compatibility of our highest mass ratio solutions with the findings by Scheeres (2007) and Pravec et al. (2010), that determine (theoretically and by observations) the conditions for splitting of pairs created by fission. Their main finding is a general tendency to create separate pairs when the mass ratio is <0.2, which is about the inferior threshold for our distribution of solutions for high values of mass ratio when small

asteroids <20 km are considered (magenta circles in Fig. 24). Some of the systems close to (and above) that critical splitting threshold, could have been produced by fission. Considering that mass shedding and re-accumulation in orbit is the most likely formation process for small asteroids ($D \lesssim 20$ km) (Pravec & Harris 2007; Walsh et al. 2008), we expect that future observations characterising our sample should find results compatible with the known population in this size range.

7. Conclusions

Our work presented a simple but robust method to detect periodic signals in the residuals of the orbital fit in *Gaia* astrometry, which is compatible with the presence of binary asteroids. We explored all astrometric data of asteroids available in the DR3 catalogue and, after a careful selection process combining statistical and physical analyses, we obtained a consolidated list of candidates that are likely to be binary asteroids.

Among the strengths of our results, we can highlight that the presence fraction of our candidates has physical properties similar to the known binary population (small asteroids with companions discovered by photometry), but also spans a parameter range that tends to fill the gap where other techniques have been failing to find binaries, as expected from preliminary studies (Mignard et al. 2007; Pravec & Scheirich 2012). The distribution of the taxonomy of our candidate binaries also provides interesting matches to the known population and some potentially new evidence. We also find a large fraction of possible synchronous binaries, which are notoriously difficult to discover by photometry, especially if the secondary rotation is also synchronised (doubly synchronous binaries).

Some limitations of our approach are also clear. As we deal with an astrometric accuracy of the order of 1 mas, we cannot exclude that some signals are not due to companions, but to the photocentre displacement occurring during the rotation of irregular, single asteroids. This could be the case, especially for the largest objects approaching the 100 km diameter, although the relatively small phase angles should mitigate the effect. However, our preliminary analysis of the detection of photocentre variation shows that even though for some cases it could be the source of the signal detected, the results are not significant enough to discard the possibility of the astrometric detection of a satellite.

We also note that our comparison with known satellites is limited to a very small sample, but seems to point out that in some cases we detect companions, although not with the parameters derived by previous surveys. The parameters estimated by making use of taxonomy must be considered with caution since they could be misclassified, which would significantly change the separations. Additionally, we face problems due to the limitations of the data. For instance, the length of the observation window can be shorter than the wobble period, the irregular sampling that causes the rise in aliases in the period determination, and the usually small number of data points along with low S/N that complicates the detections.

Future improvements may involve many aspects: the exploitation of the astrometry could adopt a more refined error model at the transit level; the inclusion of observations more distant in time, with an increase in the accessible period range; a more complex model for the astrometric signal, involving flattening of the components and non-uniform illumination; a more detailed analysis of the spectral properties and the choices of sizes and plausible densities. In addition, a full simulation of the whole process starting from a synthetic binary population

and simulated *Gaia* astrometry is a complex task that should provide a better view of the biases present in our method, and its limitations. However, such extensive investigation is beyond the scope of this work.

Other *Gaia* data releases, such as the Focused Product Release (which appeared in October 2023) and, in particular, the DR4 (including all the asteroid astrometry from the nominal mission) can certainly offer opportunities to consolidate our approach and refine the results. We expect the astrometric quality to increase further, and the number of consecutive data sequences (also for the same asteroid) to be more frequent.

Despite these limitations, we think that we have revealed the capabilities of *Gaia* in discovering asteroid satellites by the astrometric method, and we have set the foundations for future, more refined searches. Stellar occultation campaigns focused on our candidates may pin down the orbits, sizes, and shapes of the companions. In the absence of our *Gaia* binary candidates, a blind search of asteroid companions by stellar occultation observations would be unlikely to succeed. In addition, without the stellar occultation technique, it is hard to better assess some of the candidate system properties.

For some of the candidates with similar characteristics to the known binary population discovered by photometry, light curves could be sufficient to determine the presence of a companion. Therefore, we warmly invite the community to collaborate on the campaigns that we are organising, by photometry and stellar occultations, to validate our findings and physically characterise the *Gaia* astrometric binaries.

References

Astropy Collaboration (Robitaille, T. P., et al.) 2013, A&A, 558, A33
Astropy Collaboration (Price-Whelan, A. M., et al.) 2018, AJ, 156, 123
Astropy Collaboration (Price-Whelan, A. M., et al.) 2022, ApJ, 935, 167
Benishek, V., Pravec, P., & Pilcher, F. 2023a, (2871) Schober, in IAU Central Bureau Electronic Telegrams, 5215, 1
Benishek, V., Pravec, P., Vachier, F., et al. 2023b, (18301) Konyukhov is a binary asteroid, in IAU Central Bureau for Astronomical Telegrams, 5320, 1
Berthier, J., Carry, B., Mahlke, M., & Normand, J. 2023, A&A, 671, A151
Carry, B. 2012, Planet. Space Sci., 73, 98
Chapman, C., Veverka, J., Thomas, P., et al. 1995, Nature, 374, 783
Ćuk, M. 2007, ApJ, 659, L57
Descamps, P., Marchis, F., Michalowski, T., et al. 2009, Icarus, 203, 102
Doressoundiram, A., Paolicchi, P., Verlicchi, A., & Cellino, A. 1997, Planet. Space Sci., 45, 757
Drummond, J. D., Merline, W., Carry, B., et al. 2021, Icarus, 358, 114275
Durda, D. D., Bottke Jr, W. F., Enke, B. L., et al. 2004, Icarus, 167, 382
Durda, D. D., Bottke Jr, W. F., Nesvornỳ, D., et al. 2007, Icarus, 186, 498
Ďurech, J., & Hanuš, J. 2023, A&A, 675, A24
Durech, J., Kaasalainen, M., Warner, B. D., et al. 2009, A&A, 493, 291
Durech, J., Sidorin, V., & Kaasalainen, M. 2010, A&A, 513, A46
Ďurech, J., Hanuš, J., & Vančo, R. 2019, A&A, 631, A2
Ďurech, J., Tonry, J., Erasmus, N., et al. 2020, A&A, 643, A59
Efron, B. 2010, Significance Testing Algorithms, Institute of Mathematical Statistics Monographs (Cambridge University Press), 30
Eyer, L., & Bartholdi, P. 1999, A&ASS, 135, 1
Fairbairn, M. B. 2005, JRASC, 99, 92
Gaia Collaboration (Spoto, F., et al.) 2018, A&A, 616, A13
Gault, D., Nosworthy, P., Nolthenius, R., Bender, K., & Herald, D. 2022, Minor Planet Bull., 49, 3
Hanuš, J., Ďurech, J., Brož, M., et al. 2011, A&A, 530, A134
Hanuš, J., Viikinkoski, M., Marchis, F., et al. 2017, A&A, 601, A114
Hanuš, J., Pejcha, O., Shappee, B. J., et al. 2021, A&A, 654, A48
Hestroffer, D., Cellino, A., Tanga, P., et al. 2010, in Dynamics of Small Solar System Bodies and Exoplanets (Springer), 251
Hunter, J. D. 2007, Comput. Sci. Eng., 9, 90
Jacobson, S. A., Scheeres, D. J., & McMahon, J. 2013, ApJ, 780, 60
Jutzi, M. 2019, Planet. Space Sci., 177, 104695
Kaasalainen, M., Torppa, J., & Piironen, J. 2002, Icarus, 159, 369
Lambert, J.-H. 1760, Photometria sive de mensura et gradibus luminis, colorum et umbrae (Sumptibus viduae Eberhardi Klett, typis Christophori Petri Detleffsen)
Lomb, N. R. 1976, Astrophys. Space Sci., 39, 447
Madeira, G., Charnoz, S., & Hyodo, R. 2023, Icarus, 115428
Marchis, F., Descamps, P., Berthier, J., et al. 2008, Icarus, 195, 295
Marchis, F., Enriquez, J. E., Emery, J. P., et al. 2011, Icarus, 213, 252
Marchis, F., Enriquez, J. E., Emery, J. P., et al. 2012, Icarus, 221, 1130
Marciniak, A., Michałowski, T., Kaasalainen, M., et al. 2007, A&A, 473, 633
Marciniak, A., Bartczak, P., Santana-Ros, T., et al. 2012, A&A, 545, A131
Marciniak, A., Bartczak, P., Müller, T., et al. 2018, A&A, 610, A7
Marciniak, A., Ďurech, J., Choukroun, A., et al. 2023, A&A, 679, A60
Margot, J.-L., Pravec, P., Taylor, P., Carry, B., & Jacobson, S. 2015, Asteroids IV, 355, 373
McKerns, M., Strand, L., Sullivan, T., Fang, A., & Aivazis, M. G. 2012, Building a Framework for Predictive Science, arXiv e-prints [arXiv:1202.1056]
Merline, W., Close, L., Dumas, C., et al. 2000, in AAS/Division for Planetary Sciences Meeting Abstracts 32, 13-06
Merline, W. J., Weidenschilling, S. J., Durda, D. D., et al. 2002, Asteroids III, 1, 289
Merline, W. J., Close, L. M., Tamblyn, P. M., et al. 2003, IAU Circ., 8075, 2
Merline, W. J., Tamblyn, P. M., Drummond, J. D., et al. 2009, Central Bureau Electronic Telegrams, 2054, 1
Michel, P., Benz, W., Tanga, P., & Richardson, D. C. 2001, Science, 294, 1696
Mignard, F., Cellino, A., Muinonen, K., et al. 2007, Earth Moon Planets, 101, 97
Minker, K., & Carry, B. 2023, A&A, 672, A48
Ortiz, J. L., Cikota, A., Cikota, S., et al. 2011, A&A, 525, A31
Ostro, S., Hudson, R., Benner, L., et al. 2002, Asteroids III, 151
Pedrotti, F. L., Pedrotti, L. M., & Pedrotti, L. S. 2017, Introduction to OPTICS (Cambridge University Press)
Pravec, P., & Harris, A. W. 2007, Icarus, 190, 250
Pravec, P., & Scheirich, P. 2012, Planet. Space Sci., 73, 56
Pravec, P., Scheirich, P., Kušnirák, P., et al. 2006, Icarus, 181, 63
Pravec, P., Vokrouhlickỳ, D., Polishook, D., et al. 2010, Nature, 466, 1085
Richardson, D. C., & Walsh, K. J. 2006, Annu. Rev. Earth Planet. Sci., 34, 47
Scargle, J. D. 1982, ApJ, 263, 835
Scheeres, D. J. 2007, Icarus, 189, 370
Segev, N., Ofek, E. O., & Polishook, D. 2023, MNRAS, 518, 3784
Stephens, R. D., & Warner, B. D. 2020, Minor Planet Bull., 47, 224
Tanga, P., Pauwels, T., Mignard, F., et al. 2023, A&A, 674, A12
VanderPlas, J. T. 2018, ApJS Series, 236, 16
VanderPlas, J. T., & Ivezic, Ž. 2015, ApJ, 812, 18
Vavilov, D. E., Carry, B., Lagain, A., et al. 2022, Icarus, 383, 115045
Walsh, K. J., & Jacobson, S. A. 2015, Asteroids IV (University of Arizona Press), 375
Walsh, K. J., & Richardson, D. C. 2006, Icarus, 180, 201
Walsh, K. J., & Richardson, D. C. 2008, Icarus, 193, 553
Walsh, K. J., Richardson, D. C., & Michel, P. 2008, Nature, 454, 188
Weidenschilling, S. J., Marzari, F., Davis, D. R., & Neese, C. 2001, in Lunar and Planetary Science Conference, 1890
Zechmeister, M., & Kürster, M. 2009, A&A, 496, 577

Appendix A: Amplitude offset

To evaluate the statistical significance of the highest GLSP peak, we explained in Sect. 4.1 that we use Monte Carlo simulations. Their purpose is to generate a large number of synthetic times series that are consistent with our model when there is no sinusoid, that is, an unknown constant sampled at the same time sampling grid as the data set under investigation plus noise consistent with the data error bars.

Therefore, to generate unknown constants (offsets) that are consistent with our *Gaia* residual time series, we estimated this offset from a large number of time series and plotted its distribution. The result is in Fig. A.1, limited between -2 and 2 mas for visualisation purposes.

The figure compares the empirical distribution with a Laplacian (magenta, parameters μ and b) and Gaussian (red, parameters μ and σ) best fit. Clearly, the Laplacian fit is better and we use this distribution to generate offsets in the Monte Carlo simulations.
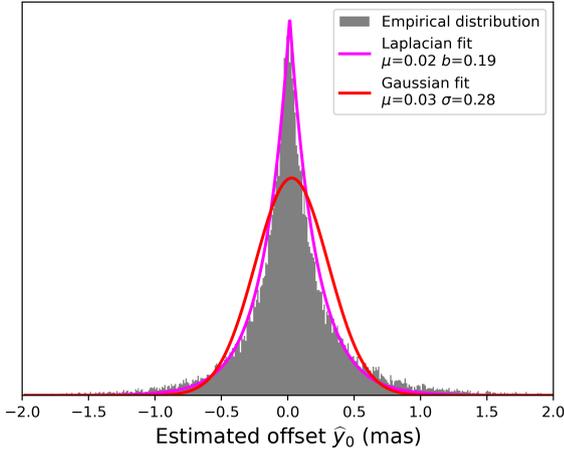
Fig. A.1: Histogram of distribution of the estimated offset $\widehat{y_0}$ for each window of search explored. The grey bars represent the empirical distribution, while the coloured lines represent different distributions fitted to the data and their respective best-fit parameters.

Appendix B: Detection of eccentric system

Highly eccentric binary asteroid systems are rare and not expected to be frequently found in our sample. However, it is interesting to evaluate how efficiently we detect eccentric binaries' astrometric signal. For such analysis, we start by creating samples of data from synthetic eccentric binary asteroid systems in one dimension, which mimics the projection of the residuals from the orbital fit of the *Gaia* DR3 astrometric data in the AL direction.

The generic binary system's parameters were chosen arbitrarily but consistent with what we expect to detect. It consists of a primary with a diameter of $D_1 = 40$ km and a secondary with $D_2 = 8$ km resulting in a mass ratio of $q = 0.008$. The bulk density of both objects was chosen to be $\rho = 2.5$ g cm^{-3}, a separation of $a = 120$ km equivalent to 3 radii of the primary and eccentricities of $e = 0, 0.2, 0.5,$ and 0.8. Assuming that it was observed at a distance of 2.8 au, the resulting signal has a nominal wobble

amplitude of $\alpha = 1.8$ *mas* and a period of $T = 30.5$h, as shown in Fig. B.1.
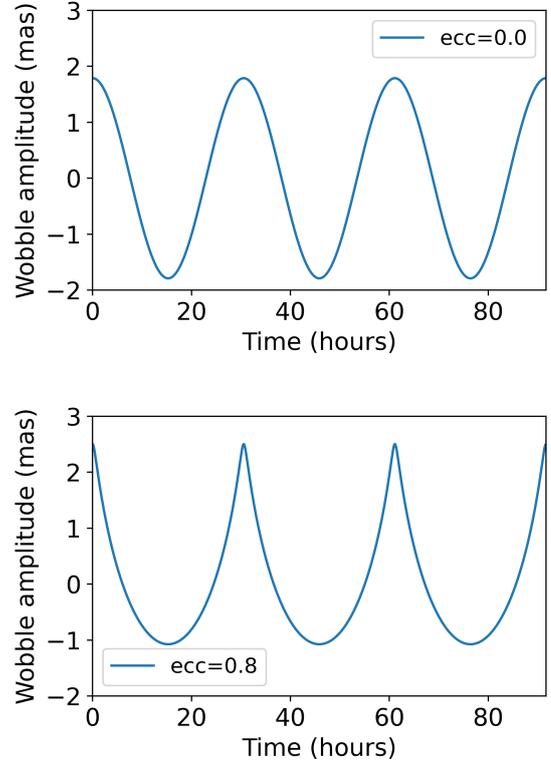
Fig. B.1: Nominal signal from the generic binary system with extreme eccentricity values.

We randomly select 20 data points from the nominal signal over 3.75 days and create 1000 different samples of data with a process similar to the one described in Sect. 4.2, as it is also done for the white noise samples.

For each one of the 5 situations analysed (4 different values of eccentricity + noise only), we then calculate the empirical probability of detection ($\hat{P}_{det}$) and of false alarm ($\hat{P}_{FA}$). When plotting $\hat{P}_{det}$ as a function of $\hat{P}_{FA}$ we obtain Fig. B.2, which provides a visual representation of how well our test distinguishes between the two classes (noise only, or noise plus wobble). The plot shows that the detection method is efficient since all curves are concentrated on the top-left corner.

For instance, if we choose to accept a false alarm probability of 5% we find that about 98% of the signals from the circular case are detected against approximately 85% from the highly eccentric systems. Turning to the estimated $\widehat{S/N}$ and estimated wobble amplitude for the case where $\hat{P}_{FA} = 5\%$, we obtained the results of Fig. B.3. Here, the considered wobble amplitudes correspond to a nominal S/N of approximately 1.7. We see that the estimated S/N are consistent with this value. We notice that as we increase the eccentricity, the estimated amplitude decreases and the uncertainty grows. Since the periodogram uses a sinusoidal model to fit the signal from the eccentric systems, we lose some precision on the amplitude estimate when the signal comes from an eccentric system, which also impacts the $\widehat{S/N}$ estimate as shown. Still, the statistical detection process succeeds at selecting at least 90% of the samples of data from the most extreme eccentricity case. Note finally that when there is no wobble in the
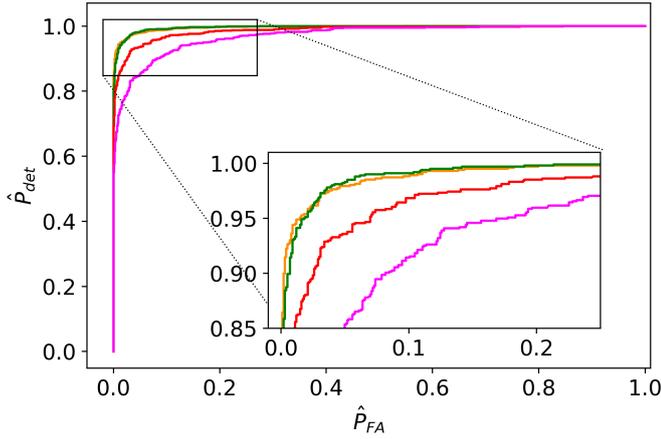
Fig. B.2: ROC curve for the different eccentricity values considered.

data, the estimated S/N tends to be much smaller (below) 1 than in the wobble-present case.

Furthermore, when analysing the frequencies at which the maximum of the periodogram is found, we find that for the circular case about 97% of the detections are within the correct signal frequency±10% against and 79% for the case with e=0.8. Even in extreme eccentricity cases, when a signal is correctly classified as a detection, the period is correctly estimated 63% of the time. We can conclude that the detection of eccentric binaries is indeed more difficult, but this study shows that the method is still efficient under these conditions. We finally note that highly eccentric binaries are not commonly found and they are also not expected to survive for long periods of time under these extreme configurations (Marchis et al. 2008; Walsh & Richardson 2008; Walsh & Jacobson 2015), so they must be exceptions and thus rarely found in our data.
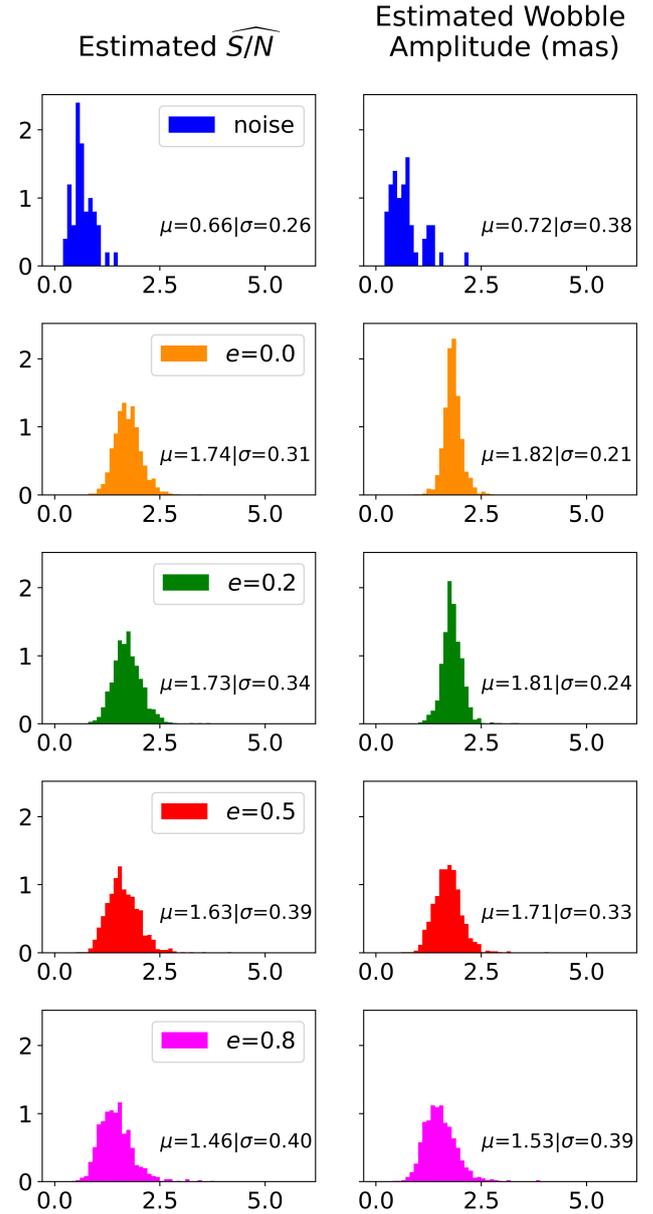
Fig. B.3: Density histograms for estimated $\widehat{S/N}$ (top) and estimated wobble amplitude (bottom) for the cases with $\hat{P}_{FA} = 5\%$.