Component masses in stellar and substellar binaries from Gaia astrometry and photometry

C.A.L. Bailer-Jones & L. Kreidberg

Max Planck Institute for Astronomy, Heidelberg, Germany

ABSTRACT

The masses of stars and planets can be measured dynamically in binary systems. For an unresolved binary, time series astrometry yields some orbital parameters, but it cannot provide the component masses, because we observe only the motion of the system's photocentre. However, as a star's luminosity is related to its mass, the observable photometry of both components together provides information on the system mass. Here we develop a method to determine the individual component masses of an unresolved binary using the astrometric orbit together with three-band photometry from Gaia. We use a mass–flux relation fitted from stellar isochrone models for each Gaia band to infer the unknown flux ratio. This enables our method to identify near equal-mass stellar binaries, expected to be the dominant source of false positive exoplanet candidates, without the need for additional follow-up. Using a likelihood approach, we sample the posterior probability distribution over the stellar parameters, marginalizing over system age and metallicity. We apply this to 20 000 systems with a main sequence primary within 300 pc of the Sun using data from the Gaia data release 3 non-single star catalogue. Primary masses can be determined with a precision (1σ posterior width) of 10–20% in 90% of cases. Secondary masses, which extend down to planetary-mass objects, are less precise, although half are more than 25% precise. Interestingly, adding either infrared photometry or spectroscopic orbits from Gaia does not change the mass estimates much (less than 4% and 1% respectively). Interstellar extinction likewise has little impact for this sample. We provide a catalogue of our mass estimates. This work shows that reasonably precise masses can be obtained for stars and substellar objects using just the Gaia astrometry and photometry.

1. Introduction

The mass of an astronomical object is one of its most fundamental properties. Masses can be measured in a variety of ways, but these all rely ultimately on a calibration via dynamical measurements. Observations of the mutual orbit of two gravitationally-bound bodies provides the means to determine the masses of both components.

The Gaia mission provides the data to determine orbits for millions of systems (Gaia Collaboration et al. 2016). As part of its third data release (Gaia Collaboration et al. 2023b), the Gaia Data Processing and Analysis Consortium (DPAC) has published solutions for about 840 000 spatially-unresolved non-single stars (NSS). Of these, around 443 000 have two-body Keplerian orbit solutions. This includes astrometric, spectroscopic, and eclipsing binaries as well as combinations of these.

Orbital data alone only allow us to infer the masses of both components in a limited set of cases, namely when we have radial velocities of both components (double-lined spectroscopic binaries, SB2) together with either astrometric or eclipsing orbits (Southworth 2020). This is unfortunate, because SB2 data is only available for systems that are bright enough to have good single-epoch radial velocity measurements and where the secondary is not too faint (mass ratio not too large). It would be helpful if we could obtain masses just from astrometric orbits, because these are available for many more systems, including fainter sources, sources in crowded fields, and highly inclined orbits. To achieve this we need additional information on the component masses. This we can obtain from broad band photometry.

For insight into how photometry helps infer mass, an unresolved binary comprising two identical stars is twice as bright as just one such star, and will therefore lie 0.75 mag higher in the colour–absolute–magnitude diagram (CAMD). To identify this shift we need to convert the observed apparent magnitude into the absolute magnitude. We can do this via the parallax, which Gaia determines self-consistently along with the NSS astrometric orbit. For an unequal-mass binary, the shift in the CAMD will be different. But as there is a relationship between mass and absolute magnitude for each component of the binary, the magnitude shift in the CAMD contains useable information on the system's mass ratio. The absolute magnitude does not only depend on the mass, however, so using (PARSEC) stellar models, we can fit a relationship between stellar mass, metallicity, and age on the one hand and absolute magnitude on the other.

In this paper we use the NSS astrometric orbit together with the parallax and three-band Gaia photometry (G, BP, RP) to estimate the masses of both components of an unresolved binary. For such a binary, Gaia can only observe the motion of the system's photocentre, so the measured semi-major axis (SMA) depends not only on the components' mass ratio but also on their flux ratio (see figure 1). We can predict this flux ratio for given stellar parameters from the stellar model, subject to the constraint that their summed flux agrees with the observed flux. Assuming a common metallicity and age for the two components, this gives us four stellar parameters to infer from the three-band photometry (each for the two components together), the parallax, the photometric SMA, and the orbital period. We perform this inference by defining the likelihood of the data given the stellar parameters, adopting a broad prior, and then sampling the resulting posterior probability density function (PDF). The age and metallicity are only weakly determined by the photometry – and not at all by the astrometry – but their presence is necessary to account for the broad scatter in the mass–absolute magnitude relationship.

Although NSS refers to its subject as non-single stars, there is nothing in the Gaia measurement technique that precludes the
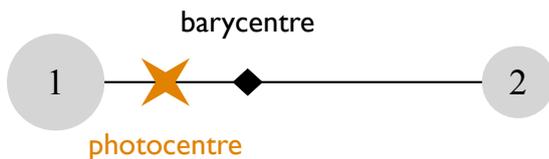
Fig. 1. Sketch of two objects, 1 and 2, of different masses in orbit about their barycentre, with a flux ratio that places the photocentre away from the barycentre. In an unresolved binary we observe the motion of the photocentre about the barycentre.

detection of substellar mass companions, namely brown dwarfs or exoplanets. These are, however, harder to detect reliably on account of their low masses, and therefore a smaller perturbation of the more massive primary star. When the companion is also very faint, inference is made easier by assuming the companion contributes no light to the observed photometry. But we cannot make this assumption a priori. This is relevant because a faint low-mass object around a bright massive object has an identical astrometric orbit as a stellar binary comprising two stars of similar mass and brightness. By neglecting the light from the companion we force the former solution. We should instead permit a variable flux ratio in the inference.

Several other studies have estimated masses, or limits thereon, of unresolved binaries using the Gaia astrometry. As part of the Gaia DR3 validation, Gaia Collaboration et al. (2023a) explored estimating masses for one or both components of the system for several different types of NSS solution. Similar in spirit to our work, Pérez-Couto et al. (2024) adopted a mass–luminosity relation to help estimate masses. This exploited the binary star results provided by the Multiple Star Classifier (MSC) in the Gaia processing (Fouesneau et al. 2023), which estimates stellar parameters of unresolved binaries using their combined BP/RP spectrum. Other studies have combined Gaia data with astrometry from Hipparcos and/or radial velocities from other surveys, often to extend the temporal baseline (e.g. An et al. 2025) The Gaia data have also been used to identify black hole (El-Badry et al. 2023; Gaia Collaboration et al. 2024) and brown dwarf (Winterhalder et al. 2024; Wallace et al. 2025) companions to stars.

The paper is laid out as follows. In section 2 we describe our inference method. We summarize the results in section 3 and compare these to other estimates in section 4. We explore the impact of adding infrared photometry and spectroscopic orbital information to the inference in sections 6 and 7 respectively. We finish up with a discussion of some of the limitations and caveats in section 8 and conclude in section 9. Appendix A presents the application of our method to simulated data to give a lower bound on the performance. Appendix B investigates the statistical contamination of stellar companions when looking for exoplanet companions.

2. Methods

2.1. What astrometry tells us

Kepler's third law relates the physical SMA a to the period T and the sum of the masses of the two components:

$$\left(\frac{a}{\mathrm{au}}\right)^3 = \left(\frac{T}{\mathrm{yr}}\right)^2 \left(\frac{\mathcal{M}_1 + \mathcal{M}_2}{\mathcal{M}_\odot}\right) \qquad (1)$$

where yr is approximately the Earth's sidereal year. From the Gaia epoch astrometry one infers not the physical SMA but

rather the SMA of the photocentre, a_p, which is related to a via (van de Kamp 1956)

$$a_p = a \left| \frac{1}{1 + \mathcal{M}_1/\mathcal{M}_2} - \frac{1}{1 + f_1/f_2} \right| \qquad (2)$$

where f_1 and f_2 are the fluxes of the two components in the band in which the astrometry is measured (in our case the Gaia G band). Note that this expression in invariant under a permutation of the subscripts.

DPAC used the Gaia epoch astrometry to compute both T and a_p. The NSS catalogue actually lists the orbital parameters as Thiele–Innes parameters (and their covariances) instead of the more familiar Campbell parameters, but Halbwachs et al. (2023) shows how to compute a_p and its uncertainty from these (and provides a python tool to do it). They provide the Thiele–Innes parameters in angular units (mas), so we divide by the parallax to put this in linear units (au). This assumes the distance equals the inverse parallax which, although generally not true (Bailer-Jones 2015), is a good approximation for the vast majority of the systems we analyse due to their high parallax signal-to-noise ratio (S/N).

Combining equations 1 and 2 gives

$$\left(\frac{a_p}{\mathrm{au}}\right)^3 \left(\frac{T}{\mathrm{yr}}\right)^{-2} = \left(\frac{\mathcal{M}_1 + \mathcal{M}_2}{\mathcal{M}_\odot}\right) \left| \frac{1}{1 + \mathcal{M}_1/\mathcal{M}_2} - \frac{1}{1 + f_1/f_2} \right|^3 . \qquad (3)$$

The Gaia astrometry gives the left side of this equation, which we call the 'astrometric measurement' and denote with symbol A. When $f_2 \ll f_1$, the right side of this equation is independent of the fluxes, so we get an estimate of a function of $\mathcal{M}_1$ and $\mathcal{M}_2$. For this reason the right side is sometimes called the 'astrometric mass function'. (If we additionally assumed $\mathcal{M}_2 \ll \mathcal{M}_1$, then the equation becomes $\mathcal{M}_2^3 \simeq A\mathcal{M}_1^2$.) To get estimates of each mass individually we need at least one more mass-sensitive measurement, such as absolute magnitude in a passband or the radial velocity variation.

### 2.2. Introducing photometry to break the mass degeneracy

The simplification of neglecting the flux of the secondary is often made when looking for faint, low mass companions to stars, as this leads to a much simpler approach. But this overlooks the degeneracy between mass and flux present in equation 3. Two stars of near-equal masses and near-equal fluxes can produce an arbitrarily small astrometric measurement. Yet two objects with the right combination of large mass and flux ratios, such as a star and a planet, can produce the same astrometric measurement. So unless we can constrain the fluxes, a given astrometric measurement can correspond to any combination of masses.

Our model accommodates the general case where f_2 is not negligible. Clearly it would be useful to know the flux from each component, but because the system is unresolved we can only measure the sum of their fluxes. If absolute magnitude depended only on stellar mass in some known way, then the sum of the fluxes together with the parallax (to convert flux to absolute magnitude), along with the astrometric measurement, would allow us to solve for the two masses (two equations, two unknowns).[1] In reality, though, the absolute magnitude on the main sequence is also a function of age, metallicity, and other parameters, so there is not a one-to-one relationship between mass and absolute magnitude. This means we need the absolute magnitude in

---

[1] This is only possible for a limited set of functional dependencies of flux on mass, and even then there can be multiple solutions.
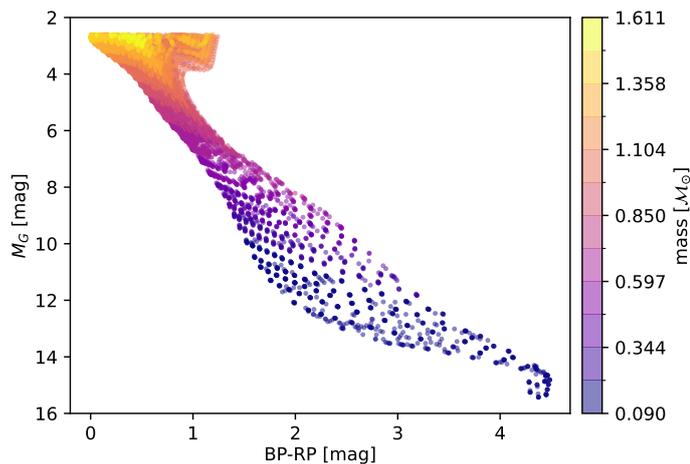
Fig. 2. Colour–absolute magnitude diagram of the PARSEC models used to fit the forward models of equation 4.
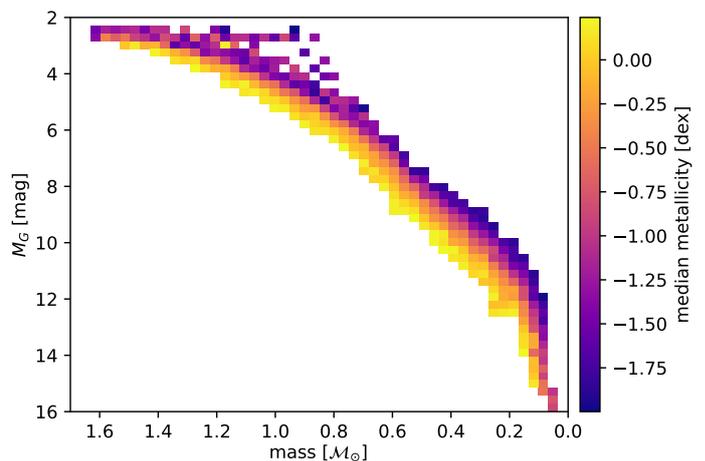
Fig. 3. Variation of M_G with mass predicted by the fitted model of equation 4, colour-coded by median metallicity.

more than one band in order to estimate individual masses (and other parameters). To do this we construct a forward model of the absolute magnitude in each band as a function of mass, age, and metallicity. These, together with equation 3, form a set of nonlinear equations which can be solved numerically.

Real stars and noisy measurements don't conform precisely to a set of deterministic equations, however, so these won't give a solution. We therefore adopt the likelihood approach in which we define the probability of the measurements given the masses (and other parameters) and then seek the most likely parameters. Rather than simply maximizing the likelihood, we define and sample the posterior probability density function (PDF), which is the product of the likelihood with a prior over the parameters. By sampling the posterior we obtain a point estimate of the parameters and a confidence interval on that estimate.

2.3. Photometric forward models from PARSEC

We use PARSEC1.2S stellar models (Bressan et al. 2012) to define the relationship between absolute magnitude (M) in each of the three Gaia bands G, B (=BP), and R (=RP) and the stellar parameters mass $\mathcal{M}$, age τ, and metallicity [M/H]. PARSEC is a stellar evolution code that solves the equations of stellar structure using information on nuclear reaction rates, material opacities, and transport processes, to predict the variation of stellar properties as a function of time. We used the public interface of PARSEC to generate a synthetic grid of stars over the mass range 0.09–5.0 $\mathcal{M}_\odot$ with a variable mass step size, a $\log_{10}(\tau/\mathrm{yr})$ range of 7.6–10.0 in steps of 0.15 dex, and a metallicity range of −2.0 to 0.25 dex in steps of 0.15 dex, retaining only evolutionary stages 1 (main sequence), 2 (sub giant branch), and 3 (red giant branch). To fit the forward models we select only those sources within the box $2.6 < M_G < 15.4$ and $0.0 < M_B − M_R < 4.5$. This reduced the mass range to 0.09–1.61 $\mathcal{M}_\odot$ (the age and metallicity ranges were unchanged) and left us with a set of 7985 models for fitting the forward models. These are shown in Figure 2. The majority (7333) are main sequence stars.

The forward model for the absolute magnitude is defined for each band independently by the three-dimensional function

$$M_{i,X} = y_X(\mathcal{M}_i, \tau, [\mathrm{M/H}]) \qquad (4)$$

where X is one of G, B and R, and $i = 1$ or 2 to indicate the primary and secondary component respectively. We experimented

with various tree and neural network architectures for fitting these models, concluding on gradient boosted trees as implemented by GradientBoostingRegressor in the scikit-learn python package Pedregosa et al. (2011). We used a squared error loss function with 200 boosting stages. 85% of the PARSEC models were used for training, the rest for testing. The median absolute deviation of the fits are 0.013 mag, 0.014 mag, and 0.010 mag in G, B, and R respectively. The bias of each fit is less than 0.001 mag in all cases.

To illustrate the models we make 4000 stars by drawing log mass, log age, and metallicity from uniform distributions over the range of the fits and simulate photometry for each. The variation of M_G with mass colour-coded by metallicity is shown in figure 3.

The forward models predict the absolute magnitudes for each component, from which we can compute the corresponding flux from

$$\hat{f}_{i,X} = 10^{-0.4*(M_{i,X}+z_X-5[1+\log_{10}\varpi])} \qquad (5)$$

where ϖ is the system parallax in arcseconds and z_X is the magnitude zeropoint of the Gaia photometric system[2]. This assumes there is no interstellar extinction, which is a reasonable approximation for many stars within a couple of hundred of parsecs, but may be a source of error for more distant stars. We assess this assumption in section 3.2. Equation 5 also assumes the parallax is the inverse of the distance which, because the parallax S/N is large for most systems in this study, is not a limiting approximation.

The lower mass limit of the grid is 0.09 $\mathcal{M}_\odot$, but we would like to be able to infer the masses of lower mass objects. As such low-mass objects contribute little flux, it might be sufficient to simply adopt zero flux contribution when the proposed mass falls below this grid limit during the inference. But this is problematic when both components in a system are below this limit (although Gaia sees such systems only when they are very nearby). In lieu of reliable grids for lower mass objects, we simply linearly extrapolate the absolute magnitude in each band as a function of log mass for masses below 0.09 $\mathcal{M}_\odot$, with no dependence on age

---

[2] Table 5.4 of the Gaia DR3 online documentation from `https://gea.esac.esa.int/archive/documentation/GDR3/Data_processing/chap_cu5pho/cu5pho_sec_photProc/cu5pho_ssec_photCal.html`

or metallicity. This is not very accurate, but is preferable to zero flux. For systems with primary masses above $0.09\,\mathcal{M}_\odot$, the exact extrapolation has little influence on the inferred masses. It turns out that none of the primaries in our Orbital300 sample (defined below) has a median mass below this limit.

Summing the fluxes for both components in a given band gives the flux we would expect to measure

$$\hat{f}_X = \hat{f}_{1,X} + \hat{f}_{2,X} .\tag{6}$$

Adopting a Gaussian noise model for the Gaia flux measurements and assuming the forward model is true, the probability of measuring a flux f_X from an unresolved binary system is described by a Gaussian with mean $\hat{f}_X$ and standard deviation Σ_{f_X}. The latter is

$$\Sigma_{f_X}^2 = \sigma_{f_X}^2 + 0.848 s_X^2 (\hat{f}_{1,X}^2 + \hat{f}_{2,X}^2)\tag{7}$$

which is the quadrature combination of the uncertainty in the measured flux σ_{f_X}, and the uncertainties in the forward model fit for each component s_X (in magnitudes). The factor $(\ln(10)/2.5)^2 = 0.848$ comes from the flux–to–magnitude conversion. We set s_X equal to the median absolute deviation of the forward model fits quoted above.

2.4. The posterior probability density function

The posterior PDF is a product of the prior over the parameters, $P(\theta)$, and the four likelihood terms:

$$P(\theta \mid A, \varpi, \{f_X\}) = P(\theta)\, P(A \mid \theta) \prod_X P(f_X \mid \theta, \varpi)\tag{8}$$

where $A = a_p^3/T^2$ is the astrometric measurement and $\theta = (\mathcal{M}_1, \mathcal{M}_2, \tau, [\mathrm{M/H}])$ is the four sampling parameters. The last term on the right side is the product of the three Gaussian likelihoods in the Gaia fluxes described in section 2.3 ($X = G, B, R$). The middle term is the probability of obtaining the astrometric measurement given the parameters. When the model and measurements are perfect these quantities are linked by equation 3. But the right side of equation 3 is model-predicted (from the parameters θ) whereas the left side is measured. Making the convenient (but not entirely accurate) assumption that deviations between these are Gaussian, the likelihood term $P(A \mid \theta)$ is then a Gaussian distribution with mean equal to the right side of equation 3 and standard deviation σ_A, where

$$\left(\frac{\sigma_A}{A}\right)^2 = \left(\frac{3\sigma_{a_p}}{a_p}\right)^2 + \left(\frac{2\sigma_T}{T}\right)^2\tag{9}$$

is obtained by combining the uncertainties in the SMA and the period in quadrature, which is reasonable for small fractional uncertainties. The posterior is therefore the product of four one-dimensional Gaussian likelihoods and the prior. As mass and age must be positive, the sampling is done over the logarithm of the masses and age (metallicity is already a logarithmic expression of abundance).

We use a separable prior in all four parameters. The mass priors are uniform in log mass between $0.001\,\mathcal{M}_\odot$ and the maximum of the forward model grid ($1.61\,\mathcal{M}_\odot$), which is minimally informative (Bailer-Jones 2017; section 5.3.1), and zero outside. We later impose absolute magnitude and colour cuts on the data to exclude stars likely to be more massive than this upper limit. The age prior corresponds to a constant star formation rate from the maximum age $\tau = \tau_{\mathrm{max}}$ in the grid to the present date $\tau = 0$,

so is $P(\log \tau) \propto \tau(\tau_{\mathrm{max}} - \tau)$, and zero outside this range. The prior in metallicity increases linearly from the lowest metallicity in the grid $[\mathrm{M/H}]_{\mathrm{min}}$, so is $P([\mathrm{M/H}]) \propto [\mathrm{M/H}] - [\mathrm{M/H}]_{\mathrm{min}}$, and is zero outside the grid. This linear increase (triangle distribution) could reflect a steady enrichment over time. Neither this metallicity prior nor the age prior is particularly realistic, but they are arguably more realistic that uniform log priors. The Gaia fluxes are only weakly sensitive to age and metallicity, so a more complex and informative prior is not justified. We later demonstrate that adopting instead the least informative priors possible (uniform log) hardly changes our results.

2.5. Sampling and summarizing the posterior

We sample the posterior using a Markov Chain Monte Carlo (MCMC) method (emcee; Foreman-Mackey et al. 2013) with 32 samplers for 1000 iterations. The initial samples are drawn uniformly (in the sampling parameters) over the full range of the grid. For each of the 32 samples the higher mass is assigned to the primary and the lower mass to the secondary. This is done just to provide a clear initial distinction between the primary and the secondary. During the sampling there is no constraint that one mass be higher than the other, and samples are never swapped. This is essential in order to retain detailed balance in the MCMC and to allow the mass posteriors to overlap.

The last 500 iterations of the sampling, thinned by a factor of five, represent the four-dimensional posterior. We summarize each marginal posterior using three quantiles: the median, the 15.87 percentile, and the 84.12 percentile. These are named – for the primary mass, for example – $\mathcal{M}_{1,\mathrm{med}}$, $\mathcal{M}_{1,\mathrm{lo}}$, and $\mathcal{M}_{1,\mathrm{up}}$ respectively. The median is taken as the point estimate. The difference between the median and the 15.87 percentile is the lower 1σ-like uncertainty ('lower error bar'), and the difference between the 84.13 percentile and the median is the upper 1σ-like uncertainty ('upper error bar'). If the median mass of the secondary is higher than the median mass of the primary, they are relabelled at this point, although this is rarely necessary.

Our Monte Carlo chains are relatively short, although the number of chains (walkers) compensates for this to some degree. As we report below, longer chains do not change the inferred masses by much.

2.6. Adding spectroscopic (radial velocity) information

In spectroscopic binaries we observe variations in the radial velocity of the brighter component due to its orbit around the barycentre (type SB1). These measurements permit the determination of two more Thiele–Innes parameters (published in Gaia DR3; Gosset et al. 2025). Combining these with the astrometric data, we can determine the spectroscopic SMA a_s (Halbwachs et al. 2023, appendix B). This quantity is related to the physical SMA via equation 3 again, but now with $f_1/f_2 = \infty$ (and so $1/(1 + f_1/f_2) = 0$), because in an SB1 we only observe the motion of the primary. We therefore define the 'spectroscopic measurement' as the left side of the following equation

$$\left(\frac{a_s}{\mathrm{au}}\right)^3 \left(\frac{T}{\mathrm{yr}}\right)^{-2} = \frac{\mathcal{M}_2^3}{(\mathcal{M}_1 + \mathcal{M}_2)^2}\tag{10}$$

with masses again in solar units.
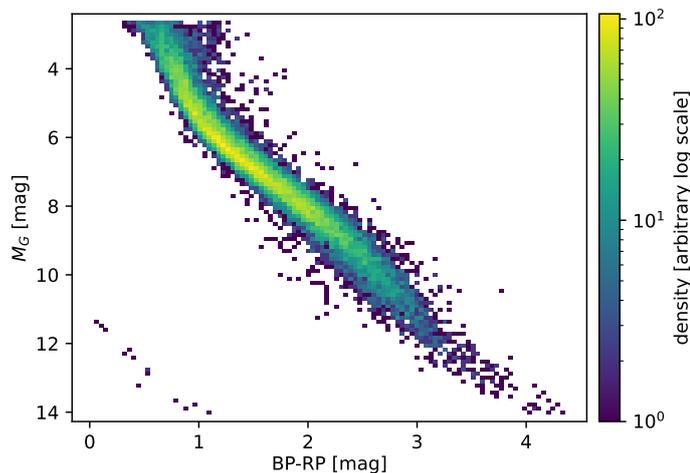
Fig. 4. Colour–absolute magnitude diagram for the Orbital300 sample assuming zero extinction.

2.7. Comparison statistics

To quantify the overall difference between two different mass estimates for a set of stars, we use two metrics. The first is the median absolute residual (MAR)

$$\text{MAR} = \text{median}(|y_i - x_i|) \tag{11}$$

where y_i and x_i are the two mass estimates for object i. This is a measure of the scatter between the estimates. The bias between the estimates is the median residual

$$\text{bias} = \text{median}(y_i - x_i) . \tag{12}$$

Note that the MAR includes the impact of a bias, in the sense that if x and y differ just by a large constant factor, both the bias and MAR will be large. When comparing our results to those by other authors, y are our results and x are the comparison results.

3. Results using astrometry and Gaia photometry

We select all systems with `nss_solution_type='Orbital'` from the `nss_two_body_orbit` table[3] in Gaia DR3 that lie within our Gaia photometry limits, have matches in the infrared 2MASS (Skrutskie et al. 2006) and AllWISE (Wright et al. 2010; Mainzer et al. 2011) point source catalogues, and that lie within 300 pc of the Sun. These Gaia photometry limits are the same as those we used to build the forward models in section 2.3, with the absolute G-band magnitude computed neglecting extinction and with distance equal to inverse parallax. We only use the infrared photometry later in section 6. This selection provides 20 334 systems, which we refer to as the 'Orbital300' sample. Its CAMD is shown in figure 4. The sample appears to consist almost entirely of main sequence stars (plus a few white dwarfs), although secondaries could be other stellar or substellar types.

The distributions of the astrometric data are shown in figure 5. The 1st and 99th percentiles of the period distribution are at 79 days and 2012 days (5.5 years) respectively. Shorter periods are harder to detect, because they have smaller astrometric signal. Periods much longer than the data span of Gaia DR3,
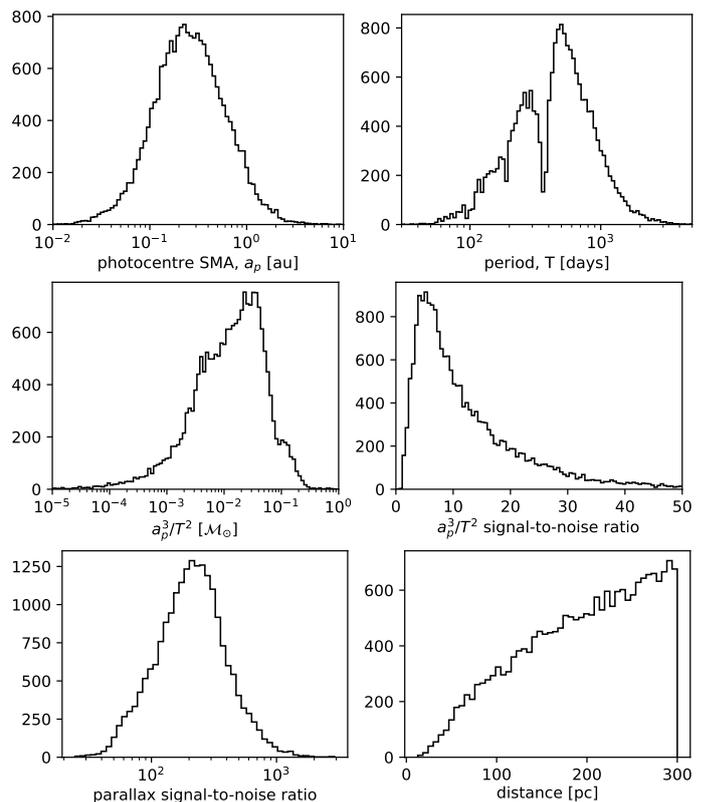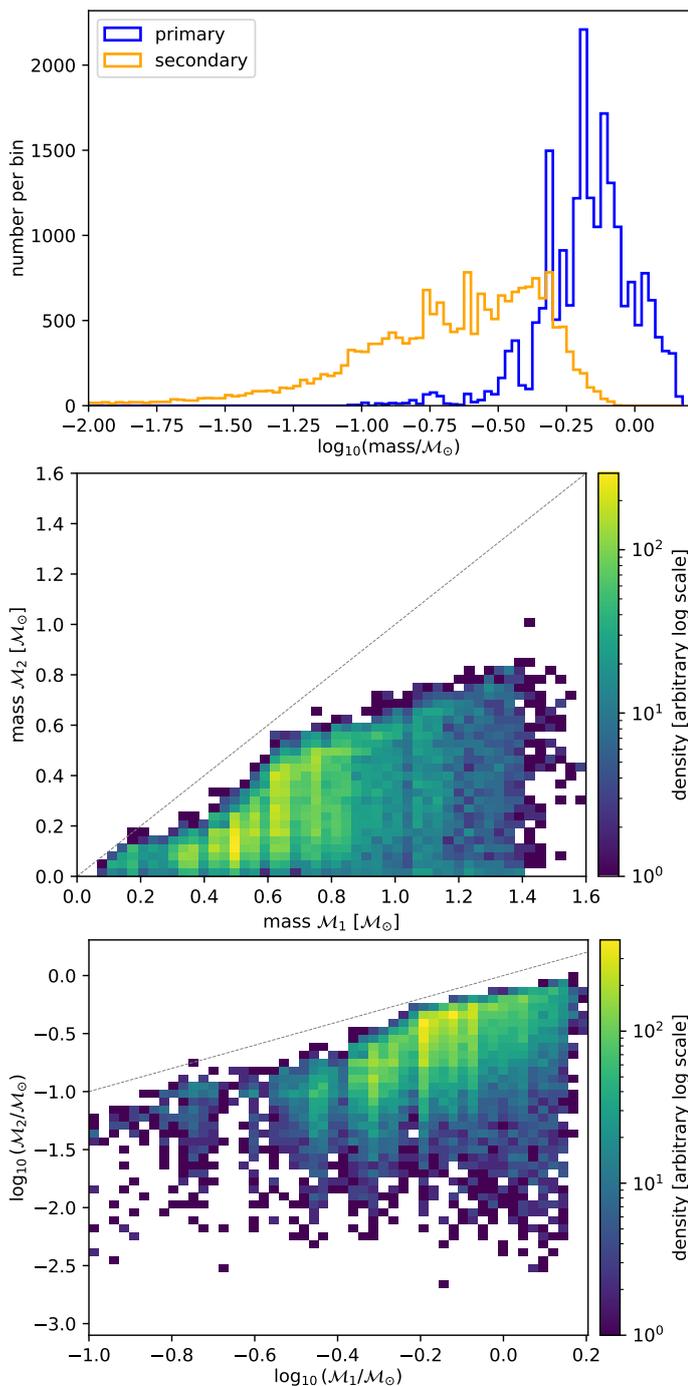
Fig. 5. Distribution of the data for the Orbital300 sample. Top row: The photometric SMA a_p and period T. Middle row: The astrometric measurement a_p^3/T^2 (with a_p in au and T in yr, from the left side of equation 3) (left) and its S/N (right). Bottom row: parallax S/N and distance (limited by construction to less than 300 pc).

which is 1038 days (2.8 years), are similarly harder to detect. The period distribution shows a sharp dip at around 365.25 days, which is the orbital period of Gaia around the Sun. The S/N of the astrometric measurement a_p^3/T^2 ranges from 1.8 and 55.8 (1st and 99th percentiles), with a median of just 9.4. The median parallax S/N is 210; 99% of sources have a parallax S/N above 49.

We run our parameter inference and summarize the posteriors as described in the previous section. The distribution of the median of the inferred masses is shown in figure 6. The median of the mass distribution is 0.44 $\mathcal{M}_\odot$ for the primaries and 0.07 $\mathcal{M}_\odot$ for the secondaries. There is a weak positive correlation between the primary and secondary masses. The lowest median primary mass is 0.093 $\mathcal{M}_\odot$ and the lowest median secondary mass is 0.0023 $\mathcal{M}_\odot$ (the limit set by the prior is 0.001 $\mathcal{M}_\odot$ in both cases). We discuss the lowest mass secondaries in section 5.
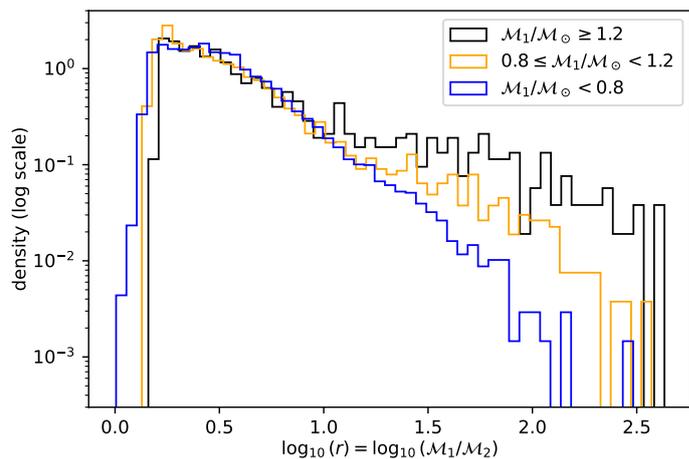
Figure 7 shows the distribution of the mass ratio for three different primary mass ranges. As we are particularly interested in the lowest mass companions, we use the parameter $r = \mathcal{M}_1/\mathcal{M}_2$ for the mass ratio, rather than its inverse. (For the log mass ratio this simply changes the sign.) The peak of $\log_{10} r$ is at 0.2–0.4, corresponding to mass ratios of 1.6–2.5. (Note that this is not the maximum of r, because a maximum is not invariant under a log transformation.) Thus for lower mass primaries we do not see as many large mass ratios. This could be in part a consequence of the lower sensitivity to lower mass secondaries.

We refrain from fitting models to the mass ratio distribution, or from inferring population statistics, because the NSS sample suffers from various selection effects, as emphasised by
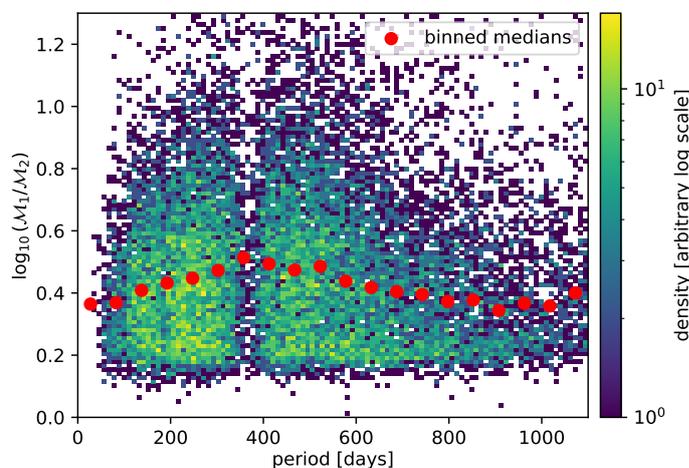
Fig. 7. Distribution of the mass ratio $r = \mathcal{M}_1/\mathcal{M}_2$ for the Orbital300 sample for three different primary mass ranges.
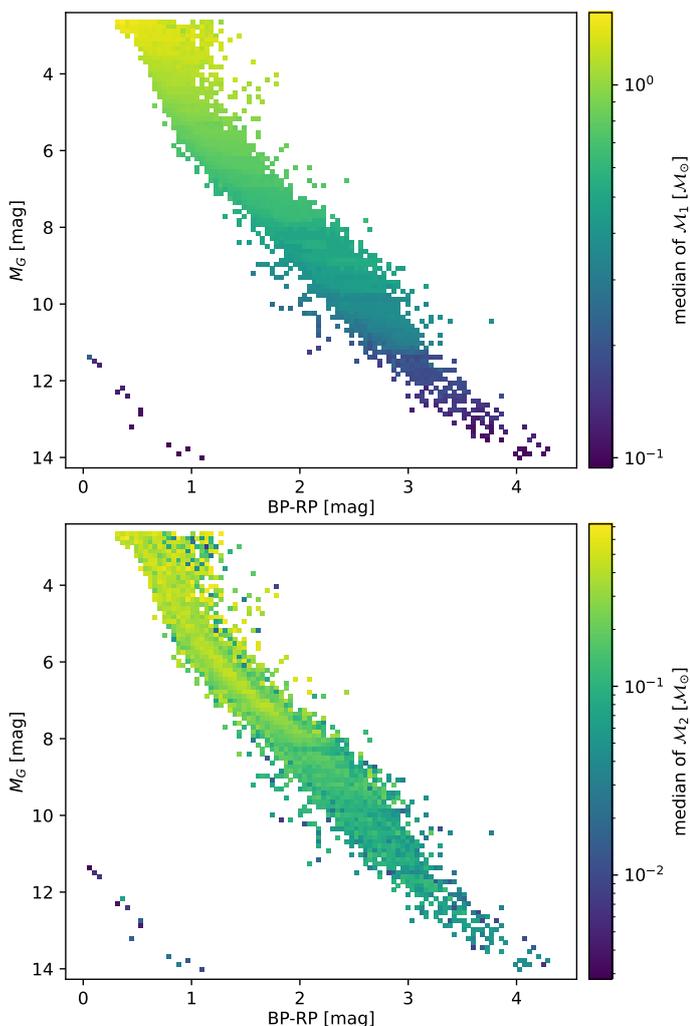

Fig. 8. Variation of the mass ratio $\mathcal{M}_1/\mathcal{M}_2$ with period for the Orbital300 sample. The red points are the median values of the mass ratio in equally-spaced period bins.

Fig. 6. Inferred masses (median of the posteriors) for the Orbital300 sample for the primary ($\mathcal{M}_1$, top) and secondary ($\mathcal{M}_2$, bottom).

The variation of the mass ratio with the orbital period is shown in figure 8. We see a slight increase in the median mass ratio at intermediate periods.

The CAMD of the sample colour-coded by the median component masses is shown in figure 9. The median over a set of stars is shown at each point. In both panels the magnitude and colour are those observed for the sytem. In the upper panel the primary mass varies fairly smoothly and continuously, with the mass decreasing for fainter and redder systems. This is what we expect when the primary dominates the system photometry. The colour-coding by the secondary mass in the lower panel shows a roughly similar trend. But for systems brighter than $M_G \simeq 8.0$ mag we see a locus of systems in the central part of the main sequence with higher average secondary masses, and thus lower mass ratios.
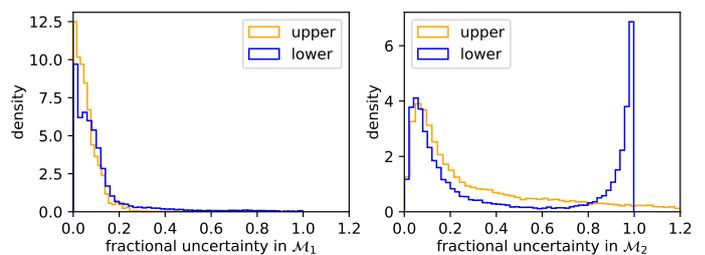
So far we have only considered the point mass estimates, which are medians of the posterior distributions. However, the posteriors can have significant widths, something we characterize with the lower and upper 1σ-like uncertainties (see section 2.5). Dividing each of these two widths by the median gives a lower and an upper fractional uncertainty respectively. These are plotted for the primary and secondary mass estimates in fig-

Halbwachs et al. (2023). It is notable, however, that for mass ratios below about 1.6, there is a steep drop in the number of binaries, meaning we find few near-equal mass binaries. Gaia's high precision astrometry makes it suitable for identifying high mass ratio systems. Nonetheless, the larger the mass ratio, the smaller the astrometric measurement (equation 3) in general, and so the more difficult it is to detect high mass ratio systems. This observational selection effect could explain some of the slope in figure 7.

Fig. 10. Distribution of the inferred fractional 1σ-like uncertainties in the masses for the Orbital300 sample for the primaries (left panel) and secondaries (right panel). The lower fractional uncertainty cannot be larger than 1.0, whereas the upper one is unbounded. There is a small number of secondaries with upper fractional uncertainties beyond the limit of the plot.

Fig. 9. System colour–absolute magnitude diagram for the Orbital300 sample colour-coded by the primary mass (top) and secondary mass (bottom).

ure 10. Note that the lower fractional uncertainty cannot be larger than 1.0, whereas the upper one is unbounded. The uncertainties for the primaries are relatively small, with median lower and upper fractional uncertainties of 0.07 and 0.05 respectively. 90% have lower and upper fractional uncertainties less than 0.23 and 0.12 respectively. The fractional uncertainties for the secondaries, in contrast, are often much larger; only 50% have lower and upper fractional uncertainties less than 0.25. A sizeable fraction have lower fractional uncertainties that are as nearly as large as 1.0. This means that the posterior permits an arbitrarily low mass (down to $0.001\,\mathcal{M}_\odot$, the lower limit of the inference). There is no obvious trend of the size of this fractional uncertainty with the size of the astrometric measurement.

Some example posteriors for the primaries and the secondaries are plotted as histograms in figure 11. The first panel shows a typical case. It is quite common for a small fraction of the MCMC samples to be 'disconnected' from the main set of samples when plotted as a histogram in this way. But the lower 'peaks' are often either near to the main one or contain few samples. Significant bimodality is very rare in the full sample. We do, however, see a few cases where there is a long low tail to low masses for the secondary, such as in panels 2 and 3 of figure 11. Panel 4 shows an example where some MCMC chains moved be-

tween primary and secondary during the sampling. However, the number of samples attributed to the other component is small, so the median is unaffected.

To test the impact of the prior on our results (section 2.4), we reran the inference using uniform log priors on age and metallicity, which are minimally informative but also less physical. Averaged over the sample, this increases the primary masses by 0.007% and the secondary masses by 0.1% (equation 12), which is negligible. There is a slightly larger scatter (equation 11) of 1% for primaries and 3% for secondaries – meaning a single source selected at random would differ by these amounts on average – but these differences are much less than the widths of the posteriors, so are accounted for by the error bars.

As MCMC is a stochastic process, the posteriors and their summaries vary if the MCMC is extended for more iterations or rerun with a different intialization. When we extend the MCMC to five times as many iterations (also for the burn-in), we find that the mass estimates change by only 1.5% for the primaries and 3.9% for the secondaries (scatter; equation 11). Rerunning with a different initialization causes less scatter. Convergence noise therefore has only a small impact on our results.

We provide our mass estimates as an electronic table with the entries show in Table 1. The log flux ratio is the ratio computed from the forward model-predicted G-band fluxes (equation 4) at the median inferred values of the masses, age, and metallicity for each system. This flux ratio is not equal to the value computed from equation 3 using the median masses and measured a_p and T, because the median stellar parameters are found by posterior sampling, and not solving that equation.

There is an obvious set of 14 white dwarfs (primaries) in the bottom left corner of the CAMD of figure 4. Our photometric forward models are not appropriate for these stars, so we do not expect to infer appropriate masses. Indeed, we get median primary and secondary masses of around $0.1\,\mathcal{M}_\odot$ and $0.008\,\mathcal{M}_\odot$ respectively for these systems. We have left them in the published catalogue because they were not excluded from our analysis (they have negligible impact on the later comparison statistics).

3.1. Relevance of flux ratio on the mass estimates

Our forward model (equation 4) predicts the G-band flux for each component of the unresolved binary from the mass, age, and metallicity. From this we compute the flux ratio at the median inferred values of the parameters. Figure 12 shows how this flux ratio varies with the inferred mass ratio for the Orbital300
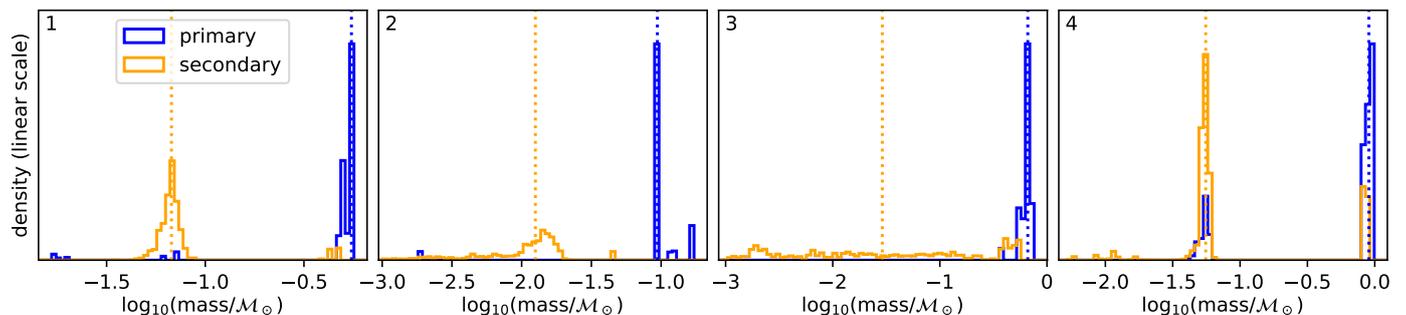
Fig. 11. Examples of the marginalized mass posteriors for the primary (blue) and secondary (orange) for four sources (one per panel). Panel 1 is typical, panels 2 and 3 show examples of long tails to low masses for the secondaries, and panel 4 shows where some of the MCMC chains moved between primary and secondary during the sampling. The vertical dotted lines show the median masses.

Table 1. Content of the catalogue of inferred masses for the Orbital300 sample.

Gaia DR3 `source_id`	$\mathcal{M}_{1,\mathrm{med}}$ [$\mathcal{M}_\odot$]	$\mathcal{M}_{1,\mathrm{lo}}$ [$\mathcal{M}_\odot$]	$\mathcal{M}_{1,\mathrm{up}}$ [$\mathcal{M}_\odot$]	$\mathcal{M}_{2,\mathrm{med}}$ [$\mathcal{M}_\odot$]	$\mathcal{M}_{2,\mathrm{lo}}$ [$\mathcal{M}_\odot$]	$\mathcal{M}_{2,\mathrm{up}}$ [$\mathcal{M}_\odot$]	a_p [au]	$\sigma(a_p)$ [au]	T [days]	$\sigma(T)$ [days]	$\log_{10}(f_1/f_2)$
1729398647131392	0.5001	0.4867	0.5493	0.1170	0.0035	0.1753	0.3118	0.0043	538.083	8.575	1.9576
2488955023504768	0.4488	0.4248	0.4525	0.0775	0.0694	0.0878	0.2136	0.0154	441.589	4.821	2.1656
3205080690546176	0.8473	0.7602	0.8642	0.0325	0.0035	0.5583	0.4405	0.0127	879.223	64.436	4.9103
3679313799492864	0.3648	0.3588	0.3882	0.1232	0.0811	0.1259	0.0897	0.0035	1074.243	56.782	1.1449
4016657710743168	1.4086	1.1721	1.5381	0.4151	0.0132	0.4474	0.4196	0.0192	822.289	26.599	2.7579
5168082607675520	0.6835	0.6479	0.7370	0.3085	0.2906	0.4449	0.1792	0.0051	711.032	4.757	1.5329
7326187710462720	1.0944	0.8926	1.1955	0.2941	0.0141	0.3830	0.1891	0.0035	300.369	1.320	2.5974
7623811764175616	0.7208	0.7176	0.7258	0.1176	0.1135	0.1241	0.1723	0.0180	241.112	1.492	2.7660
8499499760858368	0.8426	0.7643	0.8837	0.4637	0.2747	0.5320	0.2381	0.0094	448.716	0.957	1.4428
8628662312765440	0.6880	0.6622	0.7278	0.1152	0.0038	0.4162	0.1884	0.0067	1243.326	33.863	2.7783

Notes. $\mathcal{M}_{1,\mathrm{med}}$ is the median mass of the primary, $\mathcal{M}_{1,\mathrm{lo}}$ and $\mathcal{M}_{1,\mathrm{up}}$ are the 16th and 84th percentiles of the mass posterior, representing the lower and upper uncertainties respectively; and likewise for the secondary, $\mathcal{M}_2$. The table also includes the photometric SMA a_p, the period T, and their 1σ uncertainties. The final column is log base 10 of the G-band flux ratio at the median parameters. The electronic version of this table lists all 20 334 systems sorted by increasing `source_id`, with fields reported to full precision (no rounding). There are no missing data.
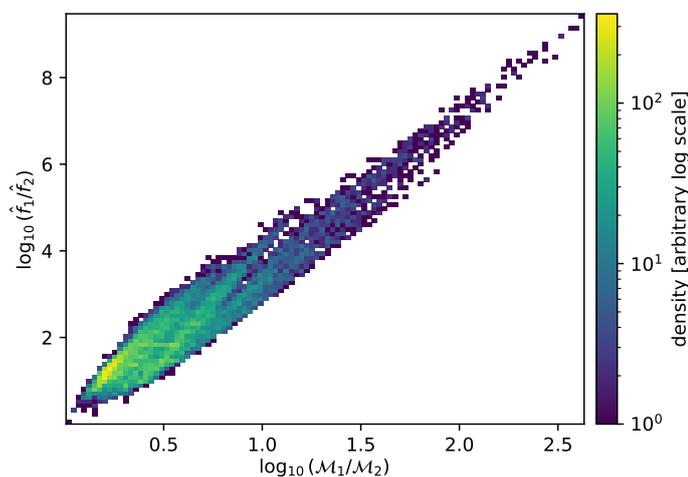
Fig. 12. Predicted G-band flux ratio vs inferred mass ratio for the Orbital300 sample.

sample. This is broadly what we expect for stellar binaries – the flux ratio increasing with mass ratio, plus additional variance due to metallicity and age variations – and is a consequence of our forward model.

The impact of the flux from the secondary on the determination of the masses via the astrometry is determined by the size of the cubed quantity on the right side of equation 3, or more conveniently by

$$R = \left| 1 - \frac{1 + \mathcal{M}_1/\mathcal{M}_2}{1 + \hat{f}_1/\hat{f}_2} \right| . \tag{13}$$

which we call the 'flux relevance ratio'. When the secondary contributes negligible flux, such that the flux ratio $\hat{f}_1/\hat{f}_2$ is very large, then R is close to one. When, instead, this flux ratio is closer to unity for a given mass ratio, then R is less than one, and the smaller R, the more relevant the flux is from the secondary in the interpretation of the astrometric measurement. The distribution of R is shown in figure 13. The majority of systems have R near to unity, but there is tail of systems with R as low as 0.4, and a handful at even lower values. The figure also shows how R varies with the primary mass. Unsurprisingly, it is larger for lower primary masses, because this is where the effect of the secondary on both the astrometry and photometry is more noticeable.

We test the importance of accounting for the flux of the secondary more directly by repeating the inference with this flux fixed to zero. In this case $f_2 = 0$ both in equation 3 and when computing the photometric likelihoods. We find that the estimated masses of the secondaries are then 9% lower on average (median), and 10% of the sources have masses that are 40% or more lower. This confirms that we need to accommodate for the unknown flux contribution of the secondary when trying to estimate its mass.
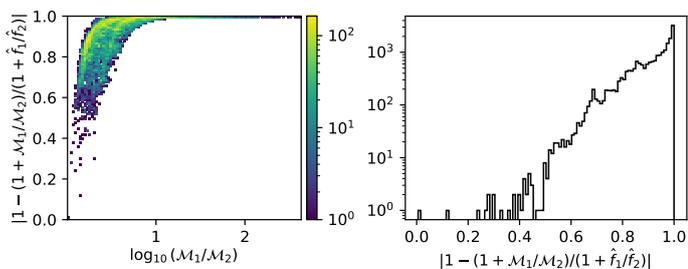
Fig. 13. Flux relevance ratio (equation 3) plotted against the inferred mass ratio (left) for the Orbital300 sample shown on a logarithmic density scale. The right panel shows the distribution of the flux relevance ratio.

## 3.2. Correction for interstellar extinction

So far we have ignored interstellar extinction. Such extinction will dim and redden the observed photometry of some stars to some degree. In reality these stars are therefore brighter and bluer, and so more massive than what we infer. Individual extinction estimates for some of our target stars are available in Gaia DR3 from the GSP-Phot algorithm Andrae et al. (2023); Fouesneau et al. (2023). This algorithm estimated the extinction parameter A_0 self-consistently together with the intrinsic stellar parameters via stellar models, and from this and the photometric passbands computed the extinction in each of the G, BP, and RP bands. Two-thirds of our Orbital300 sample – 14 173 sources – have extinction estimates, of which 13 710 have extinction-corrected magnitudes that lie within the bounds of our forward model. Their median G-band extinction A_G is 0.21 mag (10th and 90th percentiles are 0.004 and 0.56 mag). The GSP-Phot authors note that their extinction estimates may not be very accurate; in particular some may be overestimated. However, these extinction values seem plausible for sources over the whole sky within 300 pc, and so we can use them at least to assess the impact of neglected extinction.

We use the GSP-Phot extinctions to adjust the forward-model predicted fluxes in equation 5 for all three bands during the inference procedure. We then compare the inferred masses (posterior median) source-by-source with the extinction-free masses inferred earlier. When including the extinction, the primary masses are estimated to be 2.2% higher on average (median), and the secondary masses 0.4% higher. The $\mathcal{M}_1/\mathcal{M}_2$ ratio is 1.6% higher. Hence there is very little systematic increase in the mass estimates when including plausible extinctions. The scatter between the estimates is somewhat larger than these biases, indicating that some individual stars' mass estimates vary by more, although some of this is also MCMC convergence noise. Overall we can say that for stars within 300 pc, neglecting extinction has very little impact on the mass estimates with our method.

4. Comparison to other estimates

As part of the Gaia DR3 validation, Gaia Collaboration et al. (2023a) examined various combinations of Gaia NSS data to estimate the mass of the primary and/or secondary. Which data are used to estimate which masses is indicated by the flag `combination_method` in their table.[4] In several cases they had

[4] https://gea.esac.esa.int/archive/documentation/GDR3/Gaia_archive/chap_datamodel/sec_dm_performance_verification/ssec_dm_binary_masses.html
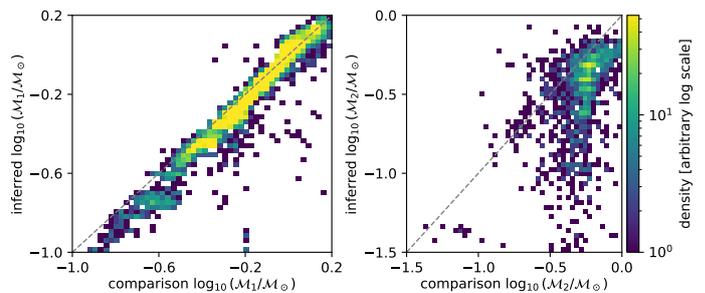
Fig. 14. Comparison of our mass estimates on the Orbital300 sample using astrometry and photometry with those from Gaia Collaboration et al. (2023a) who adopt a mass for the primary then infer the mass of the secondary from a variety of data (in many cases there is no secondary mass estimate). The two panels show the primaries (left, 19 603 stars) and secondaries (right, 1574 stars). A few sources lie outside the range of the plots.
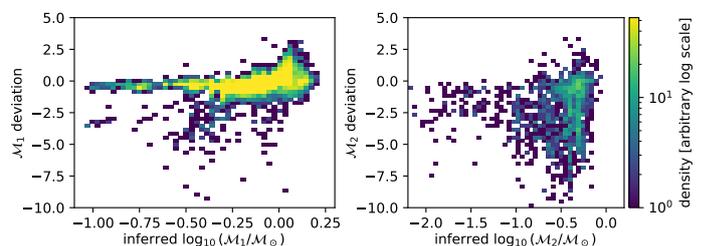
Fig. 15. Deviations, as defined by equation 14, for the results shown in Figure 14.

to adopt a value for the mass of the primary in order to estimate the mass of the secondary. Often this primary mass was estimated using Gaia photometry and isochrones. This is similar to what we are doing in the present study, except that we estimate the masses of both components simultaneously using the astrometry and photometry. Here we refer to the mass estimates from Gaia Collaboration et al. (2023a) as the 'comparison set'. Some of their $\mathcal{M}_1$ masses are lower than their $\mathcal{M}_2$ masses, which we then swap to be consistent with our results (our primary always has the higher median mass). For the Orbital300 sample analysed in the previous section, the comparison set provides estimates of the primary mass in 19 603 systems but of the secondary mass in only 1574 systems, the latter being a subset of the former. (They have more cases where the secondary has a lower and/or upper mass limit, but we don't use those here.)

Figure 14 compares the two mass estimates for the primaries and secondaries. The agreement for the primaries is reasonable. In contrast, we see some significant differences for many of the secondaries, with our estimates extending to lower masses than the comparison ones. Recall that our estimates are the medians of a posterior which for the secondaries can have a long tail to low masses, as shown earlier in figure 10.

Figure 10 compares the median mass estimates, yet the uncertainties (posterior widths) in both estimates are non-negligible. To accommodate the uncertainties in a comparison, we define the 'deviation' as the difference between the mass estimates divided by their combined error bar, using the lower error bar (subscript 'lo') of one estimate and the upper error bar (subscript 'up') of the other, according to which estimate is higher

(to get the 'neighbouring' error bar):

$$
\text{deviation} =
\begin{cases}
\dfrac{y - x}{\sqrt{\sigma_{y,lo}^2 + \sigma_{x,up}^2}} & \text{if } y - x \geq 0 \\[4ex]
\dfrac{y - x}{\sqrt{\sigma_{y,up}^2 + \sigma_{x,lo}^2}} & \text{otherwise .}
\end{cases}
\tag{14}
$$

These deviations – with y as our estimates and x as the comparisons – are plotted against our estimates in Figure 15. While the deviations for the primary masses are mostly just 'one or two sigma', in some cases the deviations for the secondary masses are considerably larger, showing that some differences are statistically significant.

We quantify the difference in the mass estimates using the statistics in section 2.7. For the primaries, the bias and scatter (MAR) in $\log_{10} \mathcal{M}_1$ are -0.024 dex and 0.028 dex respectively. Differences in log base 10 quantities correspond to ratios 2.3 times larger, so we can also interpret these values as a bias of -5.5% and a MAR of 6.4% in $\mathcal{M}_1$. In absolute terms, the bias is $-0.037\,\mathcal{M}_\odot$ and the MAR is $0.044\,\mathcal{M}_\odot$. The comparison set does include a few very high mass estimates (up to $123\,\mathcal{M}_\odot$), but as our statistics are based on the median, these have little impact.

For the secondaries, the bias and MAR in $\log_{10} \mathcal{M}_2$ are -0.16 dex (-38%) and 0.17 dex (39%) respectively. In this case the MAR is dominated by the bias: the MAR around the bias is 0.11 dex. This systematic difference is unlikely to be caused by us neglecting extinction because that should equally effect the primary mass estimates. In absolute terms, the bias and MAR are $-0.18\,\mathcal{M}_\odot$ and $0.18\,\mathcal{M}_\odot$ respectively, although these values make little sense given that $\mathcal{M}_2$ spans two orders of magnitude (see figure 6).

Among the primary masses compared, 96% have the flag `combination_method` in the comparison table equal to `Orbital+M1`. This means that $\mathcal{M}_1$ has been estimated from Gaia photometry. Among the secondary masses, 53% have `combination_method` set to `Orbital+M1`, which means that $\mathcal{M}_2$ has been estimated from the photometrically-inferred $\mathcal{M}_1$ together with the astrometric measurement (the astrometric orbit). The bias and MAR of our $\log_{10} \mathcal{M}_2$ inferences compared to this subset are higher than for the complete set, at -0.22 dex and 0.22 dex respectively. The complementary subset – the other 47% – contains mostly solutions with `combination_method` equal to `Orbital+SB1+M1`. The bias and MAR of our $\log_{10} \mathcal{M}_2$ inferences for these are lower, at -0.10 dex and 0.13 dex respectively. The comparison mass estimates in this complementary subset have used the spectroscopic orbit in addition to the astrometric orbit (and photometry), so they should be more accurate than the other comparison mass estimates. Our inferences, in contrast, have just used the photometry and astrometry in both subsets.

To investigate this further, we repeat the comparison on a different sample, namely those that have `combination_method` set to `AstroSpectroSB1+M1` and that lie within 300 pc, a set of 9674 systems we call the AstroSpectro300 sample. (This sample is defined in section 7 below.) These are all single-lined spectroscopic binaries, and so Gaia Collaboration et al. (2023a) used the combined spectroscopic and astrometric orbit along with an adopted value of $\mathcal{M}_1$ to estimate $\mathcal{M}_2$. We infer the masses with our method using just the astrometry and photometry (but not the spectroscopy) and compare results. The bias and MAR for $\log_{10} \mathcal{M}_1$ are -0.001 dex and 0.019 respectively, and for $\log_{10} \mathcal{M}_2$ are $+0.007$ dex and 0.062 dex respectively. This is
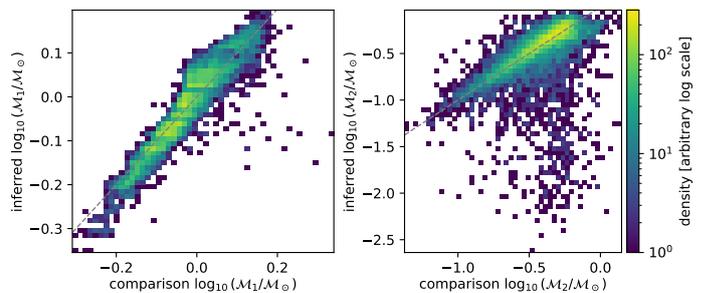
Fig. 16. Comparison of our mass estimates on the AstroSpectro300 sample when not using the spectroscopic measurement with mass estimates from Gaia Collaboration et al. (2023a) who adopt a primary mass and then used the astrometry and the spectroscopy to estimate the secondary mass. The two panels show the primaries (left) and secondaries (right), both with 9674 stars. The horizontal and vertical axis do not span a common range in each panel.
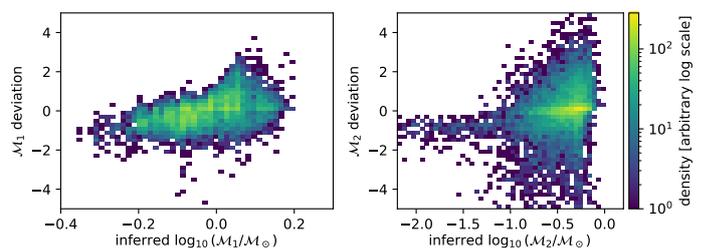
Fig. 17. Deviations, as defined in equation 14, for the results shown in Figure 16.

a much better agreement than we had above, as can be see in figure 16, which compares the masses, and figure 17, which shows the deviations. Part of the reason for the better agreement may be that the AstroSpectro300 sample is brighter – compare figures 4 and 19 – implying better quality photometry and astrometry than the Orbital300 sample. Indeed, with the Orbital300 sample, there is a trend towards increasing discrepancy in the secondary masses toward fainter magnitudes. But there is no such trend with the S/N of the astrometric measurement. In any case, the improved agreement with the AstroSpectro300 sample suggests that our method can get reasonable mass results for both components even without using the spectroscopic data.

Pérez-Couto et al. (2024) also estimated the masses of astrometric binaries from Gaia DR3 data, taking advantage of the results from the MSC algorithm in Gaia DR3. MSC attempted to estimate the stellar parameters of both components of an unresolved binary assuming the target was a binary with a flux ratio less than five, using the composite BP/RP spectrum (Fouesneau et al. 2023). The validation undertaken by the DPAC showed that these stellar parameters are not very reliable. however, and should be used with caution (Fouesneau et al. 2023; Babusiaux et al. 2023). Adopting a unique mass–luminosity relation, Pérez-Couto et al. (2024) derive the component masses from the astrometric measurement and the system's G-band photometry – corrected for extinction using MSC's estimate – by adjusting the G-band magnitude difference between the components in discrete steps and selecting the one with the best agreement with the data.

A total of 10 321 of their sources overlap with our Orbital300 sample, although 21% of their primary masses line in the range 10–$90\,\mathcal{M}_\odot$, which seems an implausibly large fraction. Exclud-

ing these, we nonetheless find a good agreement between our primary mass estimates, with a bias of +0.006 dex (1.4%) and a MAR of 0.018 dex (4.2%) in $\log_{10} \mathcal{M}_1$. The secondary masses show very poor agreement, however. Specifically, our secondary masses are invariably lower. Most of their systems have mass ratios ($\mathcal{M}_1/\mathcal{M}_2$) close to 1.0; 90% have mass ratios below 1.5. Their method is simpler than ours, but also uses less photometric data and appears not to allow for the broad intrinsic scatter in the stellar mass–luminosity relation (e.g. figure 3; see the discussion in section 2.2).

5. Planetary-mass companions

The Orbital300 sample contains 155 systems in which the secondary has a median mass below $0.012 \mathcal{M}_\odot$ (about $13 \mathcal{M}_J$). Many of these have posteriors extending to much higher masses, however, so they should not be taken as good planetary-mass candidates. We instead identify the 22 systems where $\mathcal{M}_{2,up} < 0.012 \mathcal{M}_\odot$, i.e. the 84th percentile of their mass posterior ('median plus one sigma') is below $0.012 \mathcal{M}_\odot$. Five of these systems lie on the white dwarf sequence, so their primary and secondary masses are unreliable. This leaves 17 objects that we refer to as exoplanet candidates. These are listed in Table 2.

The Gaia Consortium has published a list of 72 'astrometric objects of interest' (ASOIs) that may be binaries with an exoplanet secondary.[5] Of these, 52 are in the Orbital300 sample. Of those, 10 are in our exoplanet candidate list.

Stefánsson et al. (2025) followed up 28 of the ASOIs using radial velocities, of which five are in our exoplanet candidate list. Of these, four are designated as SB2 by Stefánsson et al. (2025). We would not expect to see the lines of an planetary-mass secondary in the spectrum, so this suggests these four objects are not planets (false positives in our list). A fifth object is designated 'Single', meaning only one set of lines (the primary) was seen. The remaining five ASOIs are designated 'Unknown', indicating there has been no known follow-up.

In total, Stefánsson et al. (2025) find 23 SBs. Of these, 22 are in our Orbital300 sample. As just noted, four are in our exoplanet candidate list. But 18 are not. Thus while our method may not be very accurate at estimating the masses of exoplanet companions – and it is not designed to do so – it shows some promise in identifying stellar or brown dwarf companions. It can therefore help to clean exoplanet target lists for follow-up.

The above caveats notwithstanding, our mass estimates for one new Gaia exoplanet and one new Gaia brown dwarf identified by Stefánsson et al. (2025) agree well with their estimates which used high-precision (10–15 m s⁻¹) multi-epoch radial velocities and multiband photometry in addition to the Gaia astrometry. For Gaia-4b (GDR3 1457486023639239296) they report a mass of $11.8 \pm 0.7 \mathcal{M}_J$, which compares to our estimate of $11.8^{+2.8}_{-8.2} \mathcal{M}_J$. For Gaia-5b (GDR3 2074815898041643520) they report a mass of $20.9 \pm 0.7 \mathcal{M}_J$, which compares to our estimate of $21.4^{+1.3}_{-1.9} \mathcal{M}_J$. Neither of these is in our exoplanet candidate list because $\mathcal{M}_{2,up} > 0.12 \mathcal{M}_\odot$ in both cases. The other three Gaia objects with exoplanet designations, Gaia-1b, -2b, and -3b (Panahi et al. 2022; Sozzetti et al. 2023) are not in our Orbital300 sample.
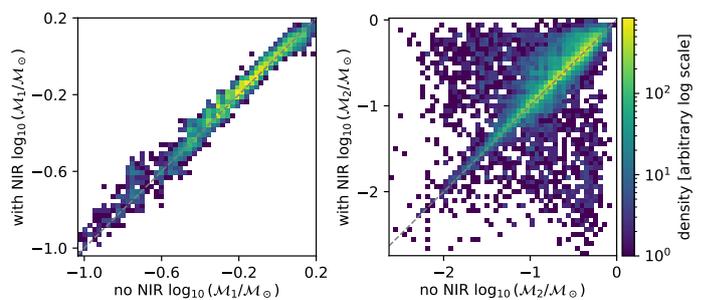
Fig. 18. Comparison of our mass estimates using infrared photometry (vertical axis) with not using infrared photometry (horizontal axis) on the Orbital300 subset for the primary (left) and secondary (right).

6. Inclusion of infrared photometry

As we are using photometry to help estimate the masses of the two components, it is reasonable to expect that more photometric bands could improve the mass estimates, either in accuracy or in precision. In this section we add the J and K_s bands from 2MASS and the W1 and W2 bands from AllWISE and include these as four extra terms in the photometric likelihood (equation 8). The astrometric likelihood is unchanged as we still use the Gaia G-band flux to compute the flux ratio in equation 3. We apply this model to the Orbital300 set used in section 3 and compare to the infrared-free results obtained there.

The overall results are very similar, as we see in figure 18. There is, however, a slight tendency for the inclusion of the infrared photometry to increase the estimated masses of both the primaries and the secondaries, by 2.2% and 3.9% respectively (median of the mass differences, i.e. the bias in equation 12). This could be a consequence of neglecting interstellar extinction. If a non-zero extinction is ignored, the system has a brighter intrinsic magnitude than we measure, which generally corresponds to a higher mass. So by ignoring extinction we expect to underestimate the mass. As infrared photometry is less affected by extinction than the visual Gaia bands, it is plausible that its inclusion delivers higher mass estimates. It is partly for this reason that we limited our analysis to systems within 300 pc. We do expect some extinction in some directions within this horizon, and this interpretation is perhaps confirmed by a slight increase in the difference in the mass estimates with distance, albeit with a weak correlation (Pearson correlation coefficients of 0.01 and 0.10 for $\mathcal{M}_1$ and $\mathcal{M}_2$ differences respectively). Yet the above mass increases are small, and the analysis in section 3.2 also indicates that neglected extinction has only a small effect on the mass estimates.

Does adding infrared photometry improve the precision of the mass estimates? To address this we compare the widths of the posteriors using the lower and upper fractional uncertainties. Recall that figure 10 shows their distributions without infrared photometry. Including infrared photometry reduces the fractional uncertainties on $\mathcal{M}_1$ by 7% and 34% for the lower and upper error bars respectively (median reduction). For $\mathcal{M}_2$ they are smaller by −1% for the lower error bars (i.e. larger, so slightly less precise) and by 14% for the upper error bar.

In summary, we find that including infrared photometry gives more precise mass estimates, but doesn't change the estimates by more than a few percent.

Table 2. All 17 exoplanet companion candidates, sorted by increasing $\mathcal{M}_{2,\text{med}}$.

Gaia DR3 `source_id`	$\mathcal{M}_{1,\text{med}}$ [$\mathcal{M}_\odot$]	$\mathcal{M}_{1,\text{lo}}$ [$\mathcal{M}_\odot$]	$\mathcal{M}_{1,\text{up}}$ [$\mathcal{M}_\odot$]	$\mathcal{M}_{2,\text{med}}$ [$\mathcal{M}_\odot$]	$\mathcal{M}_{2,\text{lo}}$ [$\mathcal{M}_\odot$]	$\mathcal{M}_{2,\text{up}}$ [$\mathcal{M}_\odot$]	a_p [au]	$\sigma(a_p)$ [au]	T [days]	$\sigma(T)$ [days]	$\log_{10}(f_1/f_2)$
1768911697089736320	1.1110	1.0148	1.1776	0.0041	0.0017	0.0086	0.1415	0.0113	656.102	15.433	8.6627
1462767459023424512	0.3362	0.2904	0.3438	0.0053	0.0021	0.0071	0.0235	0.0018	781.082	33.096	5.5126
430892357759527424	0.1279	0.1239	0.1489	0.0055	0.0049	0.0062	0.1814	0.0156	495.280	8.581	4.7508
6781298098147816192	0.3070	0.2388	0.3173	0.0058	0.0049	0.0070	0.1096	0.0031	602.376	11.731	5.2933
4702845638429469056	0.3682	0.3408	0.4063	0.0058	0.0052	0.0065	0.0782	0.0030	596.613	12.481	5.8038
4188996885011268608	0.1904	0.1850	0.1967	0.0065	0.0053	0.0072	0.0106	0.0008	406.363	3.501	4.6143
4159075462792179456	0.3336	0.2892	0.3412	0.0069	0.0042	0.0083	0.3052	0.0063	823.125	97.043	5.0892
4842246017566495232	0.3463	0.2882	0.3491	0.0072	0.0058	0.0079	0.0499	0.0012	465.430	6.461	5.2302
6694115931396057728	0.2760	0.2466	0.2898	0.0077	0.0069	0.0083	0.0418	0.0004	459.346	2.421	5.0974
557717892980808960	0.5347	0.5288	0.5384	0.0077	0.0069	0.0084	0.1982	0.0204	522.575	8.134	5.6757
5796338299045711232	0.1290	0.1264	0.1505	0.0090	0.0084	0.0098	0.0359	0.0007	488.053	3.944	3.9641
2845310284780420864	0.3509	0.3293	0.3567	0.0092	0.0085	0.0102	0.0269	0.0007	417.497	6.567	4.7233
2277249663873880576	0.4257	0.4243	0.4675	0.0094	0.0086	0.0105	0.1205	0.0023	599.359	11.777	5.2529
198464052134353536	0.1524	0.1521	0.2046	0.0094	0.0083	0.0112	0.0925	0.0007	301.518	2.699	4.0084
5486916932205092352	0.3167	0.2424	0.3181	0.0101	0.0090	0.0106	0.0181	0.0006	253.479	0.926	4.5034
2271703211828512896	0.1181	0.1172	0.1643	0.0101	0.0028	0.0119	0.0558	0.0010	622.841	10.586	3.6717
5036787935627755520	0.1871	0.1828	0.1925	0.0104	0.0094	0.0113	0.0527	0.0014	850.942	79.281	3.9374

Notes. These are the systems in the Orbital300 sample that have $\mathcal{M}_{2,\text{up}} < 0.012\,\mathcal{M}_\odot$ and are not on the white dwarf sequence of the colour–absolute magnitude diagram. Columns as in table 1.

7. Inclusion of spectroscopic orbits

Some of the stars in the Gaia DR3 NSS table have spectroscopic orbital solutions in addition to astrometric orbital solutions. To analyse the impact on the mass inference when using also the spectroscopic measurement (see section 2.6), we create a new sample single-lined spectroscopic binaries by selecting all those sources lying within 300 pc that have `nss_solution_type` set to `AstroSpectroSB1` and `combination_method` flag (section 4) set to `AstroSpectroSB1+M1`. Retaining only sources within the photometric limits of our forward models (section 2.3), this gives a set of 9674 systems which we call the AstroSpectro300 sample (used already in section 4). The CAMD of this sample is shown in figure 19. Because these sources must have (good) Gaia radial velocity spectra, this sample is brighter than the Orbital300 sample: the 95th percentile in the apparent G magnitude distribution is 12.7 mag, compared to 15.8 mag for the Orbital300 sample.

We now add an extra term for the spectroscopic measurement to the likelihood (equation 10). Everything else is unchanged. Infrared photometry is not used. The left panel of figure 20 compares the sizes of the spectroscopic and astrometric measurements. The former are larger by 11% on average (median). This is what we expect, because the non-zero flux contribution from the secondary reduces the astrometric measurement relative to the spectroscopic one (compare equations 3 and 10),

Figure 21 compares the results of the mass estimates when including spectroscopy with not including it. Overall we see very little difference. The primary masses are very similar, differing by less than 0.1% (bias; equation 12). The scatter (MAR; equation 11) is ten times larger, but is still small. The secondary masses are slightly smaller on average (bias) when using spectroscopy, but only by 0.7%, and the scatter is only 2.7%. The fractional uncertainties of these mass estimates are smaller when including spectroscopy, up 40% on average for the upper fractional uncertainty for the secondaries, but only 2–14% for the other three fractional uncertainties.

Overall, we find that including the Gaia DR3 spectroscopic information for this sample does not significantly change the
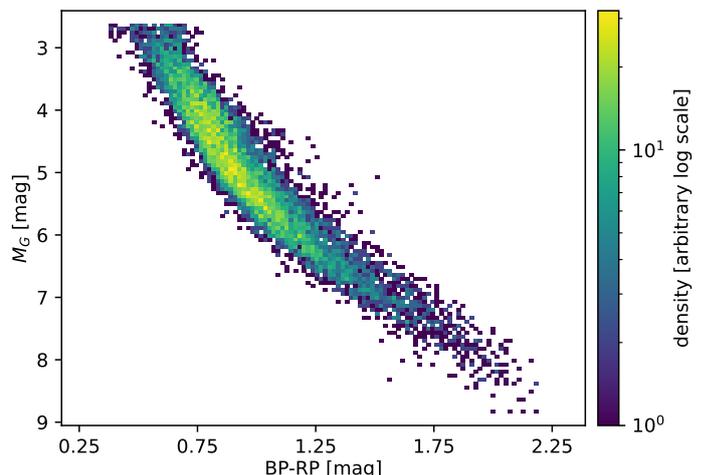

Fig. 19. Colour–absolute magnitude diagram for the AstroSpectro300 sample assuming zero extinction.
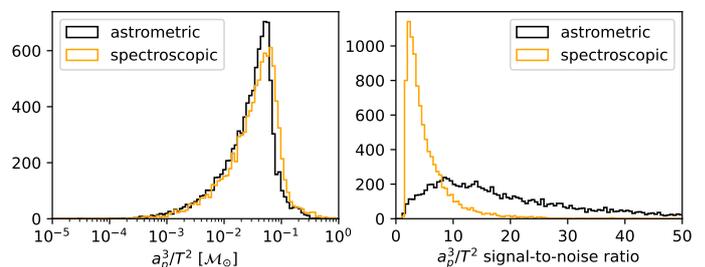
Fig. 20. Histograms of the distribution of the astrometric and spectroscopic measurements (left) and their S/Ns (right) for the AstroSpectro300 sample.

mass estimates for most systems, and only modestly improves the precision of the estimates. This is an important result, because it implies that Gaia-precision radial velocity follow-up of
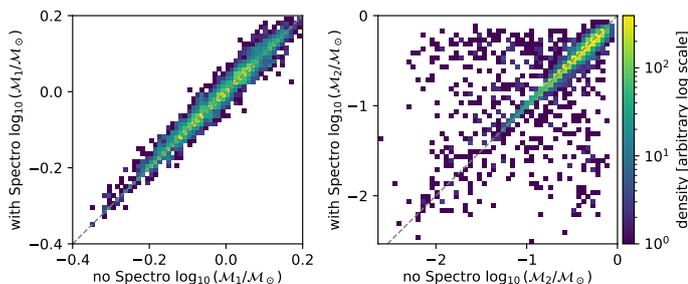
Fig. 21. Comparison of our mass estimates using spectroscopy (vertical axis) with not using spectroscopy (horizontal axis) on the AstroSpectro300 sample for the primary (left) and secondary (right) shown on a logarithmic density scale.

stars that lack spectroscopic orbits in Gaia will not help identify the lowest mass companions. However, this conclusion only applies to this sample which is quite bright and nearby. In this sample, the spectroscopic measurement S/N is on average four times lower than the astrometric measurement S/N, as we can see in figure 20, with median values of 16.4 vs 4.1. More distant stars will generally have poorer astrometric orbits (as the astrometric measurement decreases in size with increasing distance). The Gaia epoch radial velocities are not very precise – which leads to the low S/N in the spectroscopic measurement – so higher precision radial velocities would help in follow-up.

8. Limitations of the method

To determine intrinsic uncertainties in our inference method, we used it to estimate the masses of simulated binary systems. This is described in appendix A. The systems and the S/N of the data are broadly similar to the Orbital300 sample. We found that the primary masses can be estimated with an accuracy (MAR, equation 11) of 3% and a bias less than -0.2% (that is, a near negligible underestimation). For the secondary mass the MAR is 5% and the bias is -3%. This is quite a lot better than what we achieved on the Orbital300 sample.

There are a number of reasons why the performance on real data is probably worse that what these simulations yield. These could include:

1. mismatch between the PARSEC models and real stars;
2. assumption of zero interstellar extinction;
3. non-Gaussian likelihoods and/or incorrectly-estimated uncertainties.

The first of these issues is a known problem and could well be the dominant issue. In particular, our extrapolation of the fit below the lower mass limit of the PARSEC models should ideally be replaced by a brown dwarf and exoplanet model.

The second issue is a problem in principle, but our analyses using the GSP-Phot extinction values (masses systematically overestimated by less than 2%; see section 3.2), as well as the results from including infrared photometry (see section 6), suggest neglected extinction has little impact for our nearby sample (within 300 pc). We chose not to adjust the photometry using one of the various published 3D dust maps, because these tend not to fit very well in the solar neighbourhood, plus they are not designed to give accurate results for individual stars.

Concerning the third issue, we found that inflating the Gaia photometric uncertainties by a factor of five changes the inferred masses by much less than 1% on average, and changes the S/N in

these masses by less than 1% on average. On the other hand, non-Gaussianity in the sometimes noisy astrometric measurements may be an issue (figure 5), not least because it is computed from the ratio of just two quantities (equation 9), so central limit theorem convergence to a Gaussian may not apply. This uncertainty estimate has also neglected the contribution from the scatter in the forward models, which could be a limiting factor when the astrometric measurements are more precise..

Some studies, such as Winterhalder et al. (2024) and Wallace et al. (2025), adopt a simplified version of the method in this paper. By assuming $f_2 \ll f_1$, the observed photometry allows the mass of the primary to be determined independently of the secondary, and the simplified version of equation 3 then yields the mass of the secondary. However, this assumes we already know that the secondary is faint. If the single-star main sequence were an arbitrarily thin locus, the absolute photometry would betray the presence of a bright secondary. But due to age and metallicity variations the main sequence is not arbitrarily thin, and if the photometry is not dominated by the primary then the single star inference from the photometry will give an incorrect primary mass. Our method avoids this assumption to produce a self-consistent solution for the masses and fluxes of both components.

9. Summary and conclusions

We have developed a method to infer the masses of both components of an unresolved binary using the photocentre semi-major axis, orbital period, parallax, and photometric flux measurements. This involves using a forward model to map the relationship between the fluxes and the two masses and two latent variables (system age and metallicity). We fit this model using PARSEC stellar models. We sample the resulting four-dimensional posterior using a Monte Carlo method.

We tested our method on Gaia astrometry and photometry (three bands) for binaries listed in the NSS table in Gaia DR3. Our main sample (Orbital300) is a set of 20 334 systems within 300 pc that have system photometry consistent with the stellar main sequence. We can estimate primary masses with median precisions (1-σ-like uncertainties) of 6% and secondary masses with median precisions of 25%. For a significant fraction of the systems, however, the posterior for the secondaries extends to zero with non-negligible probability density, indicating a large lower uncertainty on the secondary mass.

The numbers above do not refer to deviations from the true masses. Dynamical-only 'true' mass estimates are rarely available, because they require additional information available only for a limited class of object (in particular, double-lined spectroscopic binaries). We could, however, compare our results to mass estimates obtained by Gaia Collaboration et al. (2023a). For the Orbital300 sample, the primary masses are quite similar: the scatter (median absolute residual in the log masses) is 6.4%, which includes a bias of -5.5% (our estimates are lower). The secondary masses show larger deviations: a scatter of 39%, which includes a bias of -38%. However, on a sample of brighter stars (AstroSpectro300), the agreement is much better. For the secondaries the scatter is 14%, which includes a bias of just $+1.6\%$. For this, Gaia Collaboration et al. (2023a) also used the spectroscopic orbits, whereas we did not.

Our method accommodates the initially unknown flux contribution of the secondary that affects the size of the observed semi-major axis. We found that neglecting this underestimates the mass of the secondary by 9% on average (median), but by

40% or more for 10% of the sources. Neglecting this could therefore be a source of false positive low mass companions.

We did not account for interstellar extinction, and this could in principle lead to mass underestimates. However, we found that correcting for the extinction using values from Gaia DR3 had very little impact on our mass estimates. This is supported by the fact that including infrared photometry (J, K_s, W1, W2) in the inference does not change the mass estimates by more than a few percent. This additional photometry does, however, produce more precise mass estimates (narrower posteriors).

Including Gaia's spectroscopic orbit in the inference hardly changes the mass estimates. This suggests that Gaia-like spectroscopic follow-up of these systems will not improve the mass estimates. More precise radial velocities at the appropriate epochs will help, however.

Using our method on the larger and more complete set of astrometry that will be published in Gaia DR4, we plan to estimate the masses of both components of many more binaries using only the astrometric orbit and photometric data. With this we hope in particular to identify substellar companions without being significantly affected by contamination of stellar companions having the same astrometric signal.

References

An, Q., Brandt, T. D., Brandt, G. M., & Venner, A. 2025, ApJS, 280, 61

Andrae, R., Fouesneau, M., Sordo, R., et al. 2023, A&A, 674, A27

Babusiaux, C., Fabricius, C., Khanna, S., et al. 2023, A&A, 674, A32

Bailer-Jones, C. A. L. 2015, PASP, 127, 994

Bailer-Jones, C. A. L. 2017, Practical Bayesian Inference (Cambridge University Press)

Bressan, A., Marigo, P., Girardi, L., et al. 2012, MNRAS, 427, 127

El-Badry, K., Rix, H.-W., Quataert, E., et al. 2023, MNRAS, 518, 1057

Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306

Fouesneau, M., Frémat, Y., Andrae, R., et al. 2023, A&A, 674, A28

Gaia Collaboration, Arenou, F., Babusiaux, C., et al. 2023a, A&A, 674, A34

Gaia Collaboration, Panuzzo, P., Mazeh, T., et al. 2024, A&A, 686, L2

Gaia Collaboration, Prusti, T., de Bruijne, J. H. J., et al. 2016, A&A, 595, A1

Gaia Collaboration, Vallenari, A., Brown, A. G. A., et al. 2023b, A&A, 674, A1

Gosset, E., Damerdji, Y., Morel, T., et al. 2025, A&A, 693, A124

Halbwachs, J.-L., Pourbaix, D., Arenou, F., et al. 2023, A&A, 674, A9

Hunter, J. D. 2007, Computing in Science and Engineering, 9, 90

Lammers, C. & Winn, J. N. 2026, AJ, 171, 18

Mainzer, A., Bauer, J., Grav, T., et al. 2011, ApJ, 731, 53

Marcussen, M. L. & Albrecht, S. H. 2023, AJ, 165, 266

Panahi, A., Zucker, S., Clementini, G., et al. 2022, A&A, 663, A101

Pass, E. K., Winters, J. G., Charbonneau, D., et al. 2023, AJ, 166, 11

Pedregosa, F., Varoquaux, G., Gramfort, A., et al. 2011, Journal of Machine Learning Research, 12, 2825

Pérez-Couto, X., Docobo, J. Á., & Campo, P. P. 2024, A&A, 682, A12

Skrutskie, M. F., Cutri, R. M., Stiening, R., et al. 2006, AJ, 131, 1163

Southworth, J. 2020, in Stars and their Variability Observed from Space, ed. C. Neiner, W. W. Weiss, D. Baade, R. E. Griffin, C. C. Lovekin, & A. F. J. Moffat, 329–334

Sozzetti, A., Pinamonti, M., Damasso, M., et al. 2023, A&A, 677, L15

Stefánsson, G., Mahadevan, S., Winn, J. N., et al. 2025, AJ, 169, 107

van de Kamp, P. 1956, Vistas in Astronomy, 2, 1040

Wallace, A. L., Casey, A. R., Brown, A. G. A., & Castro-Ginard, A. 2025, MNRAS, 536, 2485

Wenger, M., Ochsenbein, F., Egret, D., et al. 2000, A&AS, 143, 9

Winterhalder, T. O., Lacour, S., Mérand, A., et al. 2024, A&A, 688, A44

Wright, E. L., Eisenhardt, P. R. M., Mainzer, A. K., et al. 2010, AJ, 140, 1868
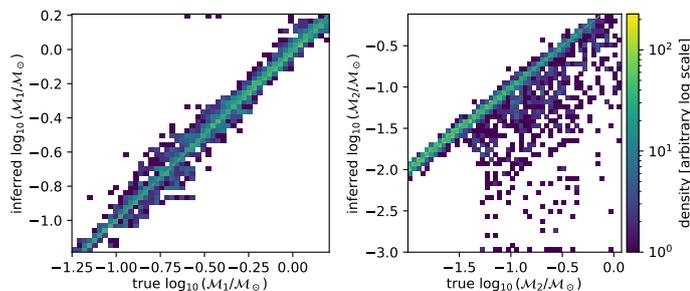
Fig. A.1. Performance of mass estimation on the simulated data set using astrometry and Gaia photometry. Note that in each panel, the horizontal and vertical axis do not span a common range.

Appendix A: Performance on simulated astrometric binaries

To determine the limiting performance of our inference method, we use it here to infer the masses of simulated astrometric binaries. To generate our sample, we first draw parameters of individual stars at random from uniform distributions in $\log_{10} \mathcal{M}$, $\log_{10}(\tau/\mathrm{yr})$, and [M/H]. The limits of these distributions are the range of the forward model grids (defined at the beginning of section 2.3), except for the lower mass limit, which is set to a lower mass of $0.01\,\mathcal{M}_\odot$. We then use the fitted forward models described in section 2.3 to predict the photometry for each star, including the extrapolation for the lower mass stars.

We generate binaries by randomly pairing stars. This is done by drawing a period at random for each system from a distribution uniform in log period between 100 and 1500 days, then computing the SMA from equations 1 and 2 and the true (noise-free) astrometric measurement from equation 3. We assign a parallax to each system at random by drawing a distance from a uniform space density distribution ($P(r) \propto r^2$) up to 1 kpc with a lower limit set for each system to ensure that the photocentre SMA is below 100 mas and so would be unresolved by Gaia.

To generate noisy data for the inference we add Gaussian noise to all inputs. For simplicity we assume fixed S/Ns for each input with the following values: 9.9 for the astrometric measurement (this is dominated by the uncertainty in the SMA); 200 for the G band flux; 100 for each of the BP and RP band fluxes; 150 for the parallax. This S/N in the astrometric measurement is close to the median of the Orbital300 sample of 9.4. The high, constant S/Ns for the Gaia fluxes essentially assume that the uncertainties are dominated by systematic errors rather than photon noise, which is consistent with the NSS sources being bright. The uncertainties corresponding to these S/Ns are also used as the standard deviations σ_A (equation 9) and σ_{f_X} (equation 7) in the likelihoods.

We compute and summarize the posteriors as done for the real data (section 2.5). The median of the posteriors for the primary and secondary masses are compared to the true masses in figure A.1. For the primaries, the bias and MAR $\log_{10} \mathcal{M}_1$ are -0.0008 dex and 0.013 dex respectively, that is, negligible bias and a median scatter of 3%. For the secondaries the bias and MAR are -0.0136 dex and 0.0236 dex respectively. The former means the inference underestimates the masses by 2% on average, it being generally lower for low mass systems and higher for higher mass systems.

If we include the four infrared photometric bands as additional measurements, the results are very similar. The biases for the primaries are secondaries are quite similar, as is the MAR

on the secondaries. But the MAR on the primaries is three times smaller when using infrared photometry. The S/Ns on the mass estimates of the primaries are also about twice as large, for both the lower and upper uncertainties, when using the infrared photometry.

This performance on simulated data is considerably better than the consistency we see in section 4 between our real results and other literature estimates, in particular for the secondaries. Possible reasons for this are discused in section 8. The simulations obviously present an ideal situation, in which the measurements – one astrometric measurement and the three fluxes – differ from the true values by only random Gaussian noise. This nonetheless indicates that in principle our method of determining the four system parameters – two masses, age, and metallicity – from the four inputs can infer accurate masses, even for the secondary. These results on simulated data represent an approximate limit on the mass estimation accuracy that we can achieve on real data.

Appendix B: Contamination of stellar companions in detecting exoplanet companions

A follow-up goal of this work when Gaia DR4 data become available is to estimate the masses of exoplanet companions in unresolved astrometric star–planet binaries. As the planet contributes negligible light to the integrated photometry, the estimation of the mass of both the stellar and planet mass from the astrometric measurement and the flux is easier than considered in this paper, because we can then assume $f_1/f_2 = \infty$ in equation 3 and need the forward model to map mass to luminosity for the stellar primary only.

However, we do not know a priori that the secondary is a planet. As discussed in section 2.2, a near equal-mass and equal-luminosity star–star binary can have an arbitrarily small photometric SMA and thus be confused as a star–planet binary. Although the method outlined in this paper is designed to overcome this, the mass uncertainties on the secondary can nonetheless be large (figure 10). But how often do 'contaminating' star–star binaries occur in practice?

Here we make some simple assumptions about the expected populations of star–star and star–planet binaries to investigate statistically how likely it is that the former have photometric SMAs similar to the latter. We look at stars with masses between 0.8 and $1.5\,\mathcal{M}_\odot$, which corresponds roughly to F and G dwarfs.

To model both types of binaries, we draw the mass of the primary from a truncated power law $P(\mathcal{M}_1) \propto \mathcal{M}_1^{-1}$. For the star–star binaries we then draw the mass ratio from a flat distribution and use this to assign the mass of the secondary. The resulting distributions are shown in the upper panel of figure B.1. For the star–planet binaries we draw the planet mass from a gamma distribution truncated between 0.5 and $5.0\,\mathcal{M}_J$. The resulting distributions are shown in the lower panel of figure B.1 We simulate ten times as many star–star as star–planet binaries to reflect a higher occurence of stellar companions over massive planet companions around F and G stars (Pass et al. 2023).

The orbital period for each system is drawn from a uniform distribution between 30 days and 5 years. From the period and masses we compute the physical SMA from Kepler's third law (equation 1). To compute the SMA of the photocentre (a_p) for the star–star binaries, we need to take into account the flux contributions from both components. To do this we first assign a log age and a metallicity for the system by drawing these quantities from uniform distributions with limits equal to those of the forward model grid (section 2.3). We then use the forward models
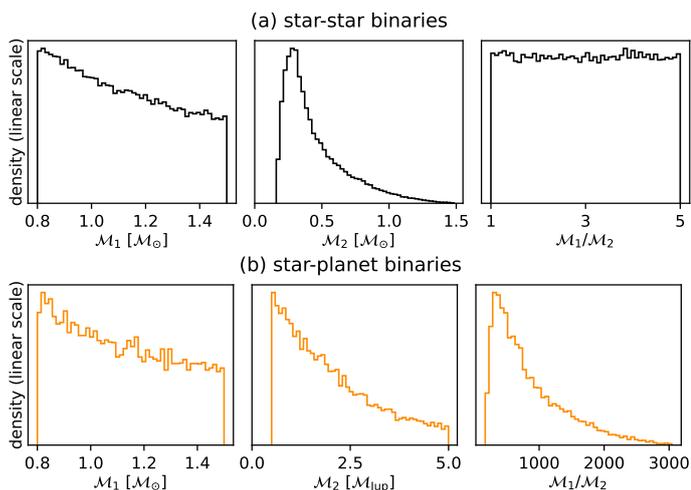
Fig. B.1. Distribution of the masses and mass ratio in the simulated star–star systems (top), and the simulated star–planet systems (bottom).
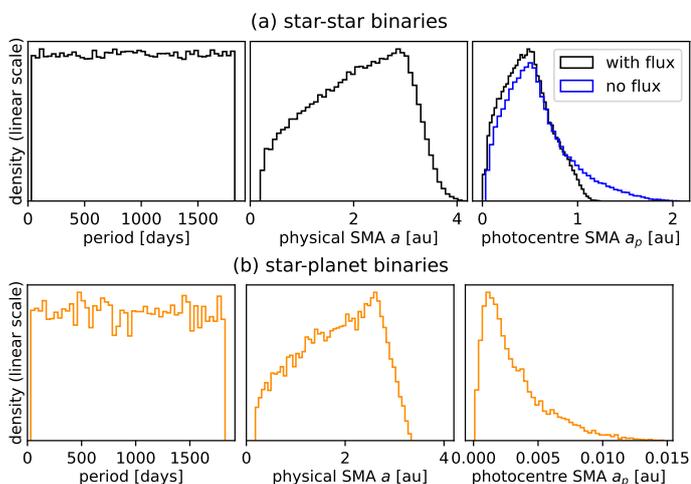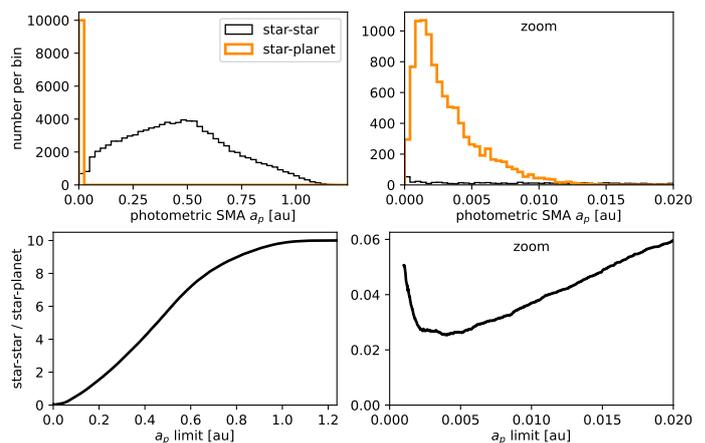

Fig. B.2. Distribution of the period and SMAs in the simulated star–star systems (top), and the simulated star–planet systems (bottom).The two left panels span a common range, as do the two centre ones, but the right ones do not. The blue curve in the upper right panel shows the distribution if we erroneously neglect the flux from the secondary in computing the photocentre of the SMA.

Fig. B.3. Comparison of the photometric SMA distributions for the simulated star–star and star–planet systems. The right panel is a zoom of the left panel in both rows. The top row shows the actual number of systems per bin (taking into account that star–planet binaries are ten times rarer). The bottom row shows the ratio of the number of star–star to star–planet systems that have a photometric SMA less than the value shown on the horizontal axis.

mimic a star–planet binary. This result is quite robust to changes both in the shapes of the mass distributions adopted and to the lower period limit. Even if we now assume stars–star binaries outnumber star–planet ones by a factor of 50, and so scale up the star–star curve by a factor of five, the above fraction becomse 0.18, a 'contamination' rate of 1 in 6.5.

A more detailed analysis of the expected stellar binary contamination in the Gaia DR4 exoplanet yield is given by Lammers & Winn (2026), and is also studied in Marcussen & Albrecht (2023).

to predict the flux for both components and finally equation 2 to compute a_p. The resulting distributions are shown in the upper panel of figure B.2.

For the star–planet binaries we assume the planet contributes no flux, so we don't need to simulate its photometry. The SMA of the photocentre is given by equation 3 with $f_1/f_2 = \infty$. The period and SMA distributions for these binaries are shown in the lower panel of figure B.2.

It is already apparent from the above figures that the photometric SMAs of the star–planet binaries are usually much smaller than those of the star–star binaries. This is more obvious in figure B.3, which plots them together. The bottom row shows the ratio of the number of systems of each type with photometric SMA below some limit. We see that below 0.01 au, there are only 0.037 star–star binaries for every star–planet binary. So although is possible for a near-equal mass star–star binary to have a very small photometric SMA, it is rare that it is small enough to