Mass estimates for visual binaries with incomplete orbits

L. B. Lucy

Astrophysics Group, Blackett Laboratory, Imperial College London, Prince Consort Road, London SW7 2AZ, UK
 e-mail: l.lucy@imperial.ac.uk

Received 11 September 2013 / Accepted 12 December 2013

ABSTRACT

The problem of estimating the total mass of a visual binary when its orbit is incomplete is treated with Bayesian methods. The posterior mean of a mass estimator is approximated by a triple integral over orbital period, orbital eccentricity and time of periastron. This reduction to 3D from the 7D space defined by the conventional Campbell parameters is achieved by adopting the Thiele-Innes elements and exploiting the linearity with respect to the four Thiele-Innes constants. The formalism is tested on synthetic observational data covering a variable fraction of a model binary's orbit. The posterior mean of the mass estimator is numerically found to be unbiased when the data cover $\geq 40\%$ of the orbit.

Key words. binaries: visual – stars: fundamental parameters – methods: statistical

1. Introduction

Visual binaries are a fundamental source of data on stellar masses. An individual system provides definitive data if it has a precise parallax ϖ and a high quality orbit yielding accurate values for the orbital period P and the semi-major axis a . The total mass of the binary in solar units is then given by Kepler's third law

$$\mathcal{M}_1 + \mathcal{M}_2 = \frac{1}{\varpi^3} \frac{a^3}{P^2} \quad (1)$$

where P is in years and a and ϖ are in arcseconds. However, for many long-period binaries, the accumulated measurements do not cover even one full orbit. The question therefore arises: can a high quality orbit be derived from such incomplete data?

In his classic monograph, Aitken (1918) gave the following answer: “in general, it is not worthwhile to compute the orbit of a double star until the observed arc not only exceeds 180° , but also defines both ends of the apparent ellipse”. Since the first observation will seldom coincide with either end point, Aitken's criterion typically requires $f_{\text{orb}} \gtrsim 0.75$, where f_{orb} is the fraction of the orbit covered. Aitken also warns that when $f_{\text{orb}} < 0.5$ it will generally be possible to draw several very different ellipses each of which will satisfy the data about equally well.

Although definitive orbits are surely a prerequisite for definitive masses, Eggen (1967) noted that a^3/P^2 is often reasonably well determined even when a and P are not, a conclusion he reached by compiling data on systems with multiple computed orbits. He reports (Eggen 1962) that “In some cases there have been drastic revaluations of the period of a given pair but there is then a compensating change in a and the resulting values of a^3/P^2 suffer little alteration”. These same compensating changes in a and P also arise in the Monte Carlo search technique of Schaefer et al. (2006) – see their Fig. 13 for DF Tau.

Eggen's examination of the historical record provides strong empirical evidence that useful masses can be obtained from incomplete orbits. But restricting the discussion to *observed* systems does not allow an assessment of the statistical properties

of such mass estimates. Accordingly, in this paper, synthetic observations with $f_{\text{orb}} < 1$ are created for a binary with specified elements in order to test the accuracy of the derived total mass.

The problem of inferring masses from incomplete orbits will be particularly acute for data acquired by the *Gaia* satellite. Nurmi (2005) and Pourbaix (2011) estimate that $\gtrsim 10^7$ visual binaries will be discovered. Many of these will have periods longer than the five years planned for the mission, and their orbits could perhaps be analysed by the technique developed in this paper.

2. Bayesian estimation

As noted already by Aitken (1918), when an orbit is incomplete, several theoretical orbits will provide satisfactory fits. Two possible responses to this circumstance are: 1) scan parameter space to find the global minimum (e.g., Hartkopf et al. 1989) and assume that this orbit is the closest achievable approximation to the truth; or 2) scan parameter space to make a complete census of acceptable orbits and then compute an appropriate average of the individual mass determinations (e.g., Schaefer et al. 2006). However, a Bayesian approach is preferred here. Specifically, by scanning parameter space, the posterior means of the orbital elements or any function thereof can be computed without locating minima.

2.1. Orbital elements

The standard orbital elements describing a secondary's motion relative to its primary are the Campbell elements $\theta = (P, T, e, a, i, \omega, \Omega)$. In addition to P and a previously defined, T is a time of periastron passage, e is the eccentricity, i is the inclination, ω is the longitude of periastron, and Ω is the position angle of the ascending node. However, for scanning parameter space, it is convenient to define an alternative to T . Noting that T can be replaced by $T \pm nP$, where n is any integer, we choose $T \in (0, P)$ and then define the alternative element $\tau = T/P \in (0, 1)$.

The above Campbell elements are standard. But an alternative set, the Thiele-Innes elements (Appendix A), simplify

Bayesian integrations. In these elements, the vector (a, i, ω, Ω) is replaced by $\psi = (A, B, F, G)$, so that the 7-element vector of elements is now (ϕ, ψ) , where $\phi = (P, e, \tau)$.

Numerous authors have preferred the Thiele-Innes elements for conventional least-squares determinations of orbital parameters for both visual binaries and exoplanets (e.g., Hartkopf et al. 1989; Pourbaix et al. 2002; Schaefer et al. 2006; Casertano et al. 2008; Wright & Howard 2009). This paper argues that these elements are also advantageous for Bayesian estimation of orbital parameters.

2.2. Posterior mean

Let $Q(\theta)$ be a quantity whose value would normally be computed after deriving a definitive orbit. But now in anticipation of there being no single definitive orbit, we instead compute Q 's posterior mean, given by

$$\langle Q \rangle = \int Q \mathcal{L} p \, d\phi d\psi / \int \mathcal{L} p \, d\phi d\psi. \quad (2)$$

Here $\mathcal{L}(\phi, \psi|D)$ is the likelihood of the elements (ϕ, ψ) given the data D , and $p(\phi, \psi)$ is a probability density function (pdf) quantifying our prior beliefs or knowledge about the orbital elements. In this problem, D comprises the secondary's measured relative Cartesian sky coordinates $(\tilde{x}_n, \tilde{y}_n)$ at times t_n .

As remarked in Sect. 2.1, the Thiele-Innes elements simplify the integrations in Eq. (2). For fixed $\phi = (P, e, \tau)$, the family of Keplerian orbits is *linear* in A, B, F, G . Accordingly, their least-squares values $\hat{A}, \hat{B}, \hat{F}, \hat{G}$ can be computed without iteration – see Sect. A.2. (Note that with normally-distributed errors the least-squares solution is also the point of maximum likelihood (ML).)

The second simplification concerns topology. When $f_{\text{orb}} \lesssim 0.5$, several orbits fit the data (Sect. 1). This implies severe topological complexity for $\mathcal{L}$ in the 7D (ϕ, ψ) -space. However, at each point in ϕ -space, the 4D function $\mathcal{L}(\psi)/\mathcal{L}(\hat{\psi})$ is a quadrivariate normal distribution (Sect. A.3), which therefore has a *single* peak at $\hat{\psi} = (\hat{A}, \hat{B}, \hat{F}, \hat{G})$, the point of ML.

This topological simplification can be used in estimating the 7D integrals in Eq. (2). To achieve this, we approximate the quadrivariate normal distribution by the delta function $\delta(\psi - \hat{\psi})$, so that

$$\mathcal{L}(\phi, \psi|D) = \mathcal{L}^\dagger \delta(\psi - \hat{\psi}) \quad (3)$$

where

$$\mathcal{L}^\dagger = \mathcal{L}(\phi, \hat{\psi}|D). \quad (4)$$

Thus, for each point in ϕ -space, $\mathcal{L}^\dagger$ is the value of $\mathcal{L}$ at $\hat{\psi} = \hat{\psi}(\phi)$, the ML point in ψ -space.

If we now replace $\mathcal{L}$ in Eq. (2) by the approximation given in Eq. (3) and integrate over ψ -space, we obtain

$$\langle Q \rangle = \int Q^\dagger \mathcal{L}^\dagger p^\dagger \, d\phi / \int \mathcal{L}^\dagger p^\dagger \, d\phi \quad (5)$$

where the superscript $\dagger$ indicates evaluation at $(\phi, \hat{\psi})$ as in Eq. (4).

On the assumption of normally-distributed measurement errors and with constants of proportionality omitted,

$$\mathcal{L}^\dagger = \exp \left(-\frac{1}{2} \chi^2 \right) \quad (6)$$

where $\chi^2(\phi, \hat{\psi})$ is the conventional goodness-fit criterion – see Eq. (A.5) – at $\phi = (P, e, \tau)$.

In the statistics literature $\mathcal{L}^\dagger$ is the *profile likelihood* (e.g., Severini 2000). This function is used in high energy physics for evaluating the significance of particle detections (e.g., Ranucci 2012).

2.3. Priors

In order to calculate the approximate posterior mean $\langle Q \rangle$ from Eq. (5), $p^\dagger = p(\phi, \hat{\psi})$ must be specified. Guided by common practice in Bayesian exoplanet detection (e.g., Ford & Gregory 2007), we assume the elements to have *independent* priors, so that $p^\dagger$ is the product of seven independent functions. The priors on the ψ elements are taken to be uniform and so cancel between numerator and denominator in Eq. (5).

For the ϕ elements e and τ , the chosen priors are uniform in $(0, 1)$. For P , we adopt a Jeffreys prior – i.e., $\log P$ uniform in $(\log P_L, \log P_U)$.

2.4. Numerical integration

The integrals in Eq. (5) are 3D as against 7D in Eq. (2). This reduction allows $\langle Q \rangle$ to be evaluated simply by computing the integrands at every point in a 3D grid and then summing. Without this reduction, a Markov chain Monte Carlo (MCMC) calculation would perhaps be required. Direct summation would, however, still be feasible if regions of negligible likelihood were efficiently excluded (Mikkelsen et al. 2012).

The integration domain in $(\log P, e, \tau)$ -space is taken to be $(\log P_L, \log P_U)$ for $\log P$, and $(0, 1)$ for both e and τ . This domain is partitioned into a 3D grid with constant steps for each variable. The quantity Q is then evaluated at the mid-points of the cells labelled (i, j, k) with $i \in (1, I)$, $j \in (1, J)$, $k \in (1, K)$. The resulting formula for the posterior mean is

$$\langle Q \rangle = \Sigma_{ijk} Q_{ijk}^\dagger \mathcal{L}_{ijk}^\dagger / \Sigma_{ijk} \mathcal{L}_{ijk}^\dagger. \quad (7)$$

The priors adopted in Sect. 2.3 are implicitly incorporated since the cells are weighted equally.

3. Calculation of $\langle Q \rangle$

In this section, the calculation of $\langle Q \rangle$ is described step-by-step.

3.1. Model binary

The model binary has the following Campbell elements:

$$\begin{aligned} P_* &= 100y & \tau_* &= 0.4 & e_* &= 0.5 & a_* &= 1'' \\ i_* &= 60^\circ & \omega_* &= 250^\circ & \Omega_* &= 120^\circ. \end{aligned} \quad (8)$$

From Eqs. (A.1), the Thiele-Innes elements corresponding to the above values of $a_*, i_*, \omega_*, \Omega_*$ are

$$\begin{aligned} A_* &= +0''.578 & B_* &= -0''.061 \\ F_* &= -0''.322 & G_* &= +0''.899. \end{aligned} \quad (9)$$

3.2. Synthetic data

Given these elements, the binary is “observed” at times

$$t_n = f_{\text{orb}} P_* \times (n - 1) / (N - 1) \quad (10)$$

for $n = 1, 2, \dots, N$, so that the fraction f_{orb} of the orbit is uniformly sampled. At each t_n , the secondary's exact coordinates (x_n, y_n) are computed from Eqs. (A.2)–(A.4). The measured positions are then

$$\tilde{x}_n = x_n + \sigma z_G \quad \tilde{y}_n = y_n + \sigma z_G \quad (11)$$

where each z_G is a random Gaussian variate drawn from $\mathcal{N}(0, 1)$ and σ is the standard error of unit weight. The vector $(f_{\text{orb}}, N, \sigma)$ defines an observing campaign.

Uniform sampling is here assumed for simplicity: it is not required by the Bayesian technique. Random sampling, the obvious alternative, is an inferior model for extant ground-based data on long-period binaries where typically an observation was made every observing season.

3.3. Calculation procedure

Given positions $(\tilde{x}_n, \tilde{y}_n)$, a posterior mean $\langle Q \rangle$ is computed as follows:

- 1) The grid parameters I, J, K and P_L, P_U are specified.
- 2) At each grid point, the vector $\phi_{ijk} = (P_i, e_j, \tau_k)$ follows from the stepping procedure described in Sect. 2.4.
- 3) Given ϕ_{ijk} , the least-squares (ML) values of the Thiele-Innes elements $\hat{\psi} = (\hat{A}, \hat{B}, \hat{F}, \hat{G})$ are computed from Eq. (A.7).
- 4) Given $\hat{\psi}$, the predicted positions (x_n, y_n) are computed from Eqs. (A.2)–(A.4) and the resulting χ^2_{ijk} derived from Eq. (A.5).
- 5) Given χ^2_{ijk} , the profile likelihood $\mathcal{L}^\dagger_{ijk}$ is given by Eq. (6).
- 6) Given $\hat{\psi}$, the corresponding Campbell elements are computed from Eqs. (A.11)–(A.15) and used to evaluate $Q_{ijk} = Q(\theta_{ijk})$.
- 7) Given Q_{ijk} and $\mathcal{L}^\dagger_{ijk}$ throughout the grid, $\langle Q \rangle$ is obtained from Eq. (7).

Note that steps 1)–4) also arise in the parameter search techniques of Hartkopf et al. (1989) and Schaefer et al. (2006).

4. Feasible orbits

In this section, we exploit the scanning of parameter space required by the calculation of $\langle Q \rangle$ to relate the findings of Aitken and Eggen (Sect. 1) to the behaviour of $\mathcal{L}^\dagger(\phi|D)$.

4.1. A particular case

With orbital parameters from Eq. (8) and campaign parameters $f_{\text{orb}} = 0.4, N = 15, \sigma = 0''.05$, synthetic data are calculated from Eq. (11). Then, with grid parameters $P_L = 0.7f_{\text{orb}}P_*$, $P_U = 300P_*$, $I = 800, J = K = 200$, the values χ^2_{ijk} at all grid points are determined as described in Sect. 3.3.

If at grid point (i, j, k) ,

$$P(\chi^2 > \chi^2_{ijk}) > 0.05 \quad (12)$$

the elements θ_{ijk} are deemed to represent a *feasible* orbit, and the ensemble of such orbits define the feasible domain(s) $\mathcal{D}$ in $(\log P, e, \tau)$ -space.

In Fig. 1, $\mathcal{D}$ is projected onto the $(\log P, e)$ plane. Thus a filled circle appears at the point $(\log P_i, e_j)$ if Eq. (12) is fulfilled

Fig. 1. Feasible domain $\mathcal{D}$ in $(\log P, e)$ -space for the test binary defined by Eq. (8) when observed in a campaign with parameters $f_{\text{orb}} = 0.4, N = 15, \sigma = 0''.05$. The true values $\log P_*, e_*$ are indicated as are the coordinates of the orbits plotted in Fig. 2.

for at least one τ_k . From this plot, we see that for this poorly-observed incomplete orbit, there is an extended domain of feasible orbits with P ranging from 40 to 5000y and e from 0.19 to >0.99 .

Because $\mathcal{D}$ extends to the grid's upper boundary at $e = 1$, a more general treatment would, in this particular case, show that some parabolic and even hyperbolic orbits are feasible – i.e., fit the data. By restricting the analysis to elliptical orbits, unbound orbits are excluded. But this is appropriate for the practical problem of concern: an incomplete observed orbit is highly unlikely to be a close encounter of an unbound pair.

In Fig. 2, the true orbit for this particular case is plotted together with the measurements $(\tilde{x}_n, \tilde{y}_n)$. Also shown are three feasible orbits indicative of the range seen in Fig. 1. The Campbell elements of these are:

$$\begin{aligned} P = 47.0y \quad T = 41.0y \quad e = 0.79 \quad a = 0''.87 \\ i = 62^\circ 3' \quad \omega = 273^\circ \quad \Omega = 157^\circ \end{aligned} \quad (13)$$

for orbit I;

$$\begin{aligned} P = 98.7y \quad T = 38.3y \quad e = 0.19 \quad a = 0''.85 \\ i = 51^\circ 1' \quad \omega = 243^\circ \quad \Omega = 98^\circ \end{aligned} \quad (14)$$

for orbit II; and

$$\begin{aligned} P = 4423.8y \quad T = 33.2y \quad e = 0.95 \quad a = 15''.2 \\ i = 73^\circ 9' \quad \omega = 207^\circ \quad \Omega = 110^\circ \end{aligned} \quad (15)$$

for orbit III.

Figure 2 illustrates and confirms Aitken's warning (Sect. 1) that when $f_{\text{orb}} < 0.5$ several orbits will fit the data. Similar plots for real binaries are given by Schaefer et al. (2006).

Orbit III merits further comment. This orbit is nearly parabolic and has its periastron during the observing campaign. Thus, the observations were taken in time interval $(0, 40y)$ and periastron occurs at $T = 33.2y$. If this were the true orbit with $P = 4423.8y$, we would count ourselves fortunate to catch a periastron passage in such a short campaign. This is a generic feature of fitting orbits to short observed arcs: the orbit is not constrained

Fig. 2. Feasible orbits. Orbits I, II, III fit the measurements (filled circles) with an acceptable χ^2 . The bold curve is the true orbit with elements given in Eq. (8).

when not observed and so can balloon out to large angular separations. This family of feasible orbits violate the Copernican Principle since the observer is inferred to be at a special epoch. Although this possibly justifies their exclusion, these orbits are retained in the calculation of posterior means.

4.2. Eggen's effect

To investigate Eggen's empirical discovery that even for poor orbits the ratio a^3/P^2 is relatively reliable, the ensemble of feasible orbits plotted in Fig. 1 is now projected onto the $(\log P, \log a)$ plane: Fig. 3. From this plot, we see a high density of points along the line

$$\log \frac{a}{a_*} = \frac{2}{3} \log \frac{P}{P_*}. \quad (16)$$

Thus, despite ranging in period from 40 to 5000y, the values of a^3/P^2 are relatively concentrated. This effect becomes even more striking with improved campaigns having $f_{\text{orb}} = 0.5$ and 0.6 (see Fig. 3).

These experiments indicate that with the chosen (N, σ) , the feasible orbits are:

- 1) closely confined to (a_*, P_*) when $f_{\text{orb}} \gtrsim 0.6$;
- 2) dispersed widely in P but narrowly in a^3/P^2 when $0.5 \gtrsim f_{\text{orb}} \gtrsim 0.4$;
- 3) dispersed widely in P and increasingly so in a^3/P^2 when $f_{\text{orb}} \lesssim 0.4$.

These results provide strong theoretical support for Eggen's discovery and for his use of the derived masses in studies of the mass-luminosity relation.

This confirmation derives from investigating the domain of high likelihood in $(\log P, e, \tau)$ -space. Since this is the domain in which high weight is assigned to Q when its posterior mean is calculated, Eggen's effect is evidently incorporated when this formalism is used to estimate the binary's mass. Accordingly, Fig. 3 implies that we can anticipate useful results in this standard case provided that $f_{\text{orb}} \gtrsim 0.4$.

Fig. 3. Eggen's effect. Feasible domains $\mathcal{D}$ in the $(\log P, \log a)$ plane. Results for three campaigns with $f_{\text{orb}} = 0.4, 0.5, 0.6$ are shown, each with $N = 15, \sigma = 0'.05$. The $f_{\text{orb}} = 0.5$ and 0.6 domains are displaced downwards by 0.5 and 1.0 dex, respectively. The plotted lines have slope 2/3 and pass through the points $(2.0, 0.0), (2.0, -0.5), (2.0, -1.0)$, each corresponding to $(\log P_*, \log a_*)$.

The explanation of Eggen's effect is that as soon as secondary's relative motion reveals a significant departure from rectilinear motion ($f_{\text{orb}} \ll 1$) then acceleration has been detected and this is $\propto (M_1 + M_2) \propto a^3/P^2$. As f_{orb} increases, the uncertainty due to orientation factors decreases and so a^3/P^2 is well-determined before the orbit is complete – see also Heintz (1978) and Schaefer et al. (2006).

In addition to the linear trends, Fig. 3 shows a dramatic increase in the volume of $\mathcal{D}$ as f_{orb} decreases from 0.6 to 0.4, supporting Aitken's warning about the onset of multiple acceptable orbits. Note that this increase in volume is a generic effect that is not sensitive to the choice of 0.05 in Eq. (12).

This section has shown that the insights of Aitken and Eggen from 50–100 years ago find their modern explanation in how the topology of the likelihood function $\mathcal{L}^\dagger(\phi|D)$ responds to changes in the observational data D .

5. Mass estimates

The analysis of Sects. 2 and 3 is now applied to simulated partial orbits for the standard visual binary. Results with variations of e and i are reported in Appendix C.

5.1. Mass estimator

At the grid point (i, j, k) , the ϕ elements are (P_i, e_j, τ_k) , and the $\hat{\psi}$ elements are the least-squares values $(\hat{A}, \hat{B}, \hat{F}, \hat{G})$. From the latter, the remaining Campbell elements $(\hat{a}, \hat{i}, \hat{\omega}, \hat{\Omega})$ are computed as described in Sect. A.4, and we denote this semi-major axis as $\hat{a}_{ijk}$.

Now if $\mathcal{M}_*$ is the true total mass of the model binary with orbit defined by Eq. (8), the inferred mass with the orbit corresponding to the point (i, j, k) is

$$\mathcal{M}_{ijk} = \mathcal{M}_* \left(\frac{\hat{a}_{ijk}}{a_*} \right)^3 \left(\frac{P_i}{P_*} \right)^{-2}. \quad (17)$$

The natural choice for the mass variable to be substituted in Eq. (7) is therefore $\mathcal{M}_{ijk}$. However, numerical experiments demonstrate that $\langle \log \mathcal{M} \rangle$ is more accurate than $\log(\mathcal{M})$. Accordingly, we take

$$Q_{ijk}^\dagger = \log \frac{\mathcal{M}_{ijk}}{\mathcal{M}_*} = 3 \log \frac{\hat{a}_{ijk}}{a_*} - 2 \log \frac{P_i}{P_*}. \quad (18)$$

Exploiting the freedom to set $\mathcal{M}_* = 1$, we now substitute Eq. (18) into Eq. (7) to obtain the posterior mean $\langle \log \mathcal{M} \rangle$, where $\mathcal{M}$ is the ratio of the inferred to the exact mass. Thus, if the exact mass is recovered, then $\langle \log \mathcal{M} \rangle = 0$.

5.2. Numerical experiments

In these experiments, the true orbit is again given by Eq. (8). In the first sequence of campaigns, $N = 15$ and $\sigma = 0.05$ as in Sect. 4, but now f_{orb} is treated as a continuous variable, thus investigating the deterioration of mass estimates as $f_{\text{orb}} \rightarrow 0$. Note that when f_{orb} changes so does the seed for the random number generator. The pattern of Gaussian variates in Eq. (11) is therefore never repeated.

With steps of 0.01 in f_{orb} , the values of $\langle \log \mathcal{M} \rangle$ are plotted in Fig. 4. We see that for $f_{\text{orb}} \gtrsim 0.40$, the mass estimates scatter about the exact value $\mathcal{M} = 1$ with errors < 0.1 dex. However, for $f_{\text{orb}} \lesssim 0.40$, the scatter increases sharply, with the error first exceeding 0.1 dex at $f_{\text{orb}} = 0.38$. These results are consistent with the implications of Fig. 3.

Accuracy should improve with more data or data of higher precision – i.e., with decreasing precision parameter $\eta = \sigma / \sqrt{N}$. This is investigated by reducing σ from 0.05 to 0.005 while keeping $N = 15$. This second sequence is plotted in Fig. 4 as filled circles. As expected, the quality of the estimates dramatically improves: the errors are < 0.012 dex for $f_{\text{orb}} \geq 0.36$, but increase sharply thereafter. Note that decreasing η by a factor of 10 has only slightly extended the domain of reliable solutions – the error first exceeds 0.1 dex at $f_{\text{orb}} = 0.34$.

These experiments demonstrate that $\langle \log \mathcal{M} \rangle$ provides a seemingly unbiased estimate of a visual binary's total mass even when standard orbit fitting generates multiple acceptable solutions with a large range of orbital periods.

5.3. Posterior pdf of $\log \mathcal{M}$

For a real binary with an incomplete orbit, f_{orb} is unknown. Moreover, perhaps counterintuitively, it cannot even be reliably estimated. Figure 2 and Eqs. (13)–(15) illustrate this: for orbit I, $f_{\text{orb}} = 0.85$; for orbit II, $f_{\text{orb}} = 0.41$; and for orbit III, $f_{\text{orb}} = 0.009$, as against the true value 0.4. It follows that an investigator cannot judge the accuracy of $\langle \log \mathcal{M} \rangle$ from an estimate of f_{orb} . Instead, the posterior pdf of $\log \mathcal{M}$ should be derived and credible intervals computed.

Inspection of Eq. (7) shows that the implied approximation to the posterior pdf is

$$\Theta(Q|D) = \sum_{ijk} W_{ijk} \delta(Q - Q_{ijk}^\dagger) \quad (19)$$

where δ is the delta function, and

$$W_{ijk} = \mathcal{L}_{ijk}^\dagger / \sum_{ijk} \mathcal{L}_{ijk}^\dagger. \quad (20)$$

From these formulae, we immediately find

$$\int \Theta dQ = 1 \quad \text{and} \quad \int Q \Theta dQ = \langle Q \rangle \quad (21)$$

where $\langle Q \rangle$ is the approximation given by Eq. (7).

Fig. 4. Mass estimator. The posterior mean of $\log \mathcal{M}$ is plotted against f_{orb} . The campaigns with $N = 15$, $\sigma = 0.05$ are plotted as open circles; those with $N = 15$, $\sigma = 0.005$ as filled circles. The dashed lines indicate errors of ± 0.1 dex.

A histogram representation of the pdf $\Theta(\log \mathcal{M}|D)$ is obtained by convolving Eq. (19) with a top-hat function. Examples are plotted in Fig. 5 for $f_{\text{orb}} = 0.4$ and 0.5. These show the loss of precision with decreasing f_{orb} that was expected from Fig. 3 and seen in Fig. 4.

The equivalent of 1σ Gaussian error bars can be computed from these pdf's by finding the equal area tails that together comprise 31.7% of the probability. Thus, we find $\langle \log \mathcal{M} \rangle = 0.012^{+0.019}_{-0.019}$ for $f_{\text{orb}} = 0.5$ and $= 0.068^{+0.077}_{-0.080}$ for $f_{\text{orb}} = 0.4$.

Because the pdf $\Theta(\log \mathcal{M}|D)$ is approximately Gaussian (see Fig. 5) the above error bars can with sufficient accuracy be replaced by the standard deviation of the pdf, given by

$$s_Q^2 = \sum_{ijk} W_{ijk} (Q_{ijk}^\dagger - \langle Q \rangle)^2. \quad (22)$$

This formula gives $s_{\log \mathcal{M}} = 0.019$ and 0.085 for $f_{\text{orb}} = 0.5$ and 0.4, respectively. Evidently, the dramatic loss in precision when the orbital coverage is insufficient is apparent from $s_{\log \mathcal{M}}$ even without knowing f_{orb} .

5.4. Bias

For $f_{\text{orb}} \gtrsim 0.4$, Fig. 4 shows no evidence that the estimator $\langle \log \mathcal{M} \rangle$ is biased. But for $f_{\text{orb}} \lesssim 0.4$, any bias is obscured by the dramatic increase in scatter. Accordingly, since quantifying the onset of bias is of interest in understanding the limits of this Bayesian approach, further calculations are now reported.

In order to beat down the noise, the campaigns with $N = 15$, $\sigma = 0.05$ are repeated $n = 100$ times for selected values of f_{orb} . The average and variance

$$\bar{Q} = \frac{1}{n} \sum_i \langle Q \rangle_i \quad s^2 = \frac{1}{n-1} \sum_i (\langle Q \rangle_i - \bar{Q})^2 \quad (23)$$

Fig. 5. Posterior pdf $\Theta(\log M|D)$ for $f_{\text{orb}} = 0.4$ and 0.5 both with $N = 15$, $\sigma = 0.05$. The vertical dashed lines are the posterior means $\langle \log M \rangle$.

of the $\langle Q \rangle_i$ are computed, so that the resulting estimate is $\bar{Q} \pm s/\sqrt{n}$, and this indicates bias if it differs significantly from the exact value of Q .

Figure 6 presents evidence of bias for the estimator $\langle \log(M) \rangle$. The average values of $\langle \log(M) \rangle$ and their standard deviations from 100 independent realisations of the orbits are plotted for $f_{\text{orb}} = 0.30(0.05)0.95$. There is no evidence of bias for $f_{\text{orb}} \geq 0.6$, but strong evidence for $f_{\text{orb}} \leq 0.4$ where the displacements are $>3.5s$. Similar plots are given in Appendix B for the posterior means of the Campbell elements.

The onset of bias coincides with the onset of scatter (see Fig. 4) at which point accurate mass determinations are no longer possible. Nevertheless, from a technical standpoint, it is of interest to understand the origin of bias. Comparison of Figs. 2, 3, and 6 shows that the onset of bias corresponds to the dramatic increase in the volume of $\mathcal{D}$, and this implies that the choice of priors becomes increasingly important. Accordingly, since the expansion of $\mathcal{D}$ brings in orbits of high eccentricity (Fig. 2), we might seek to penalize such orbits via the prior on e . But this would be an *ad hoc* fix for this *particular* model binary. A fundamental solution probably invokes the Copernican Principle (Sect. 4.1) to reduce the weight assigned to highly eccentric orbits.

5.5. Coverage fractions

When the estimate of a parameter derived by least-squares analysis is reported as $\bar{Q} \pm \sigma_Q$, we automatically interpret this as an assertion that there is a 68.3% probability that the true answer lies in the interval $(\bar{Q} - \sigma_Q, \bar{Q} + \sigma_Q)$. Implicit in this interpretation, however, are the following assumptions: 1) normally-distributed measurement errors; 2) the validity of the linearization used in deriving the normal equations; and 3) that the number of data points N greatly exceeds M , the number of parameters estimated. If one or more these assumptions does not hold then the above interval's *coverage fraction* will differ from 0.683.

In Bayesian statistics, credibility intervals derived from the posterior pdf replace the confidence intervals of frequentist

Fig. 6. Bias of the posterior mean $\langle \log(M) \rangle$. The averages and standard errors from Eq. (23) plotted against f_{orb} with $N = 15$, $\sigma = 0.05$. The dashed line is the exact value.

statistics, and the different meanings of these intervals has been topic of much discussion in the statistical literature over many years. One approach to reconciling this aspect of the Bayesian-frequentist debate is that of calibrated Bayes (e.g., Dawid 1982; Little 2006): a credible interval is *well-calibrated* if the enclosed probability matches the fraction of times that it contains the true value.

In order to test whether credible intervals calculated with this formalism for visual binaries are well-calibrated, 1000 independent realizations of two campaigns with $N = 15$, $\sigma = 0.05$ have been analysed, one with $f_{\text{orb}} = 0.60$, the other with $f_{\text{orb}} = 0.95$. For each realization, we check if Q_{exact} is in the intervals $(\langle Q \rangle - ms_Q, \langle Q \rangle + ms_Q)$ for $m = 1, 2$, where $\langle Q \rangle$ is given by Eq. (7) and s_Q by Eq. (22).

The results of this exercise are in Table 1. The first row gives the desired fraction $\mathcal{E}(f)$ – i.e., the fraction expected for an *ideal* frequentist analysis. The remaining rows give the results for the posterior means of the Campbell elements and $\log M$.

Inspection of these results shows that the coverage is close to ideal for the ϕ elements P, e, τ . Accordingly, their credible intervals are well-calibrated. However, for the remaining elements and for $\log M$, the error bars are too small by factors of 1.1–2.1 at 1σ and 1.1–1.6 at 2σ . These error bars could of course be calibrated from such simulations. However, these shortfalls occur for the elements that were replaced by the Thiele-Innes constants and thus had the widths of their posterior pdfs reduced by the delta function approximation in Eq. (3). Further research based on the analytical formulae of Sect. A.3 may alleviate or eliminate this problem.

6. Conclusion

The aim of this paper has been to apply Bayesian methods to a problem that dates back a century or more, namely how to estimate the total mass of a visual binary that has a measured parallax but an incomplete orbit. This problem has previously

Table 1. Coverage fractions f from 1000 trials.

Q	f_{orb}	1σ	2σ
$\mathcal{E}(f)$		0.683 ± 0.015	0.954 ± 0.007
$\langle \log P \rangle$	0.60	0.680 ± 0.015	0.945 ± 0.007
	0.95	0.697 ± 0.015	0.962 ± 0.006
$\langle e \rangle$	0.60	0.656 ± 0.015	0.954 ± 0.007
	0.95	0.670 ± 0.015	0.948 ± 0.007
$\langle \tau \rangle$	0.60	0.677 ± 0.015	0.945 ± 0.007
	0.95	0.656 ± 0.015	0.957 ± 0.006
$\langle \log a \rangle$	0.60	0.536 ± 0.016	0.858 ± 0.011
	0.95	0.388 ± 0.015	0.709 ± 0.014
$\langle i \rangle$	0.60	0.386 ± 0.015	0.680 ± 0.015
	0.95	0.491 ± 0.016	0.826 ± 0.012
$\langle \omega \rangle$	0.60	0.607 ± 0.015	0.892 ± 0.010
	0.95	0.516 ± 0.016	0.831 ± 0.012
$\langle \Omega \rangle$	0.60	0.398 ± 0.015	0.702 ± 0.014
	0.95	0.331 ± 0.015	0.611 ± 0.015
$\langle \log \mathcal{M} \rangle$	0.60	0.371 ± 0.015	0.691 ± 0.015
	0.95	0.480 ± 0.016	0.786 ± 0.013

been effectively treated by Schaefer et al. (2006) with a heuristic Monte Carlo search technique. In contrast, the Bayesian approach has a firm theoretical foundation and is, moreover, now the preferred methodology in many areas of astronomy, notably in cosmology (e.g., Liddle 2009) and in the discovery and characterization of the orbits of exoplanets (e.g., Ford & Gregory 2007).

A further merit of Bayesian estimation for visual binaries is that only triple integrals need to be evaluated when the quadrivariate normal distribution of the Thiele-Innes constants (Sect. A.3) is approximated by a delta function (Sect. 2.2). The use of this approximation might be expected to introduce systematic errors in the derived elements. However, the failure to detect bias when $f_{\text{orb}} \gtrsim 0.5$ (Sect. 5.4, Appendix B) demonstrates that such errors are not of practical concern.

In Sect. 1, the possible application to binaries discovered by *Gaia* was mentioned. However, the simulations reported in this paper relate to ground-based data where *two* sky coordinates are measured. In contrast, the scanning mode of *Gaia* will provide ~ 70 high precision *one* dimensional measurements of a secondary's displacement (e.g., Pourbaix 2002; Lattanzi et al. 2000). Simulations of a model binary observed in this manner are required to demonstrate the merits of this Bayesian approach for *Gaia*.

Acknowledgements. I thank D.J. Mortlock for a helpful discussion on nuisance parameters and profile likelihoods. I am obliged to the referee for suggestions that improved the paper.

Appendix A: Thiele-Innes elements

In this Appendix, a concise self-contained account of the Thiele-Innes elements is presented – see also Wright & Howard (2009) and Hartkopf et al. (1989). The explicit least-squares formulae for A, B, F, G and for their variances and covariances are new – Eqs. (A.7), (A.9), (A.10).

A.1. Definitions

The Thiele-Innes constants A, B, F, G are defined in terms of four of the Campbell elements. Specifically,

$$\begin{aligned} A &= a(+\cos \Omega \cos \omega - \sin \Omega \sin \omega \cos i) \\ B &= a(+\sin \Omega \cos \omega + \cos \Omega \sin \omega \cos i) \\ F &= a(-\cos \Omega \sin \omega - \sin \Omega \cos \omega \cos i) \\ G &= a(-\sin \Omega \sin \omega + \cos \Omega \cos \omega \cos i). \end{aligned} \quad (\text{A.1})$$

Given these constants, the coordinates (x, y) of the secondary relative to the primary at time t are

$$\begin{aligned} x(t) &= A X(E) + F Y(E) \\ y(t) &= B X(E) + G Y(E) \end{aligned} \quad (\text{A.2})$$

where

$$X = \cos E - e, \quad Y = \sqrt{1 - e^2} \sin E \quad (\text{A.3})$$

and $E(t)$, the eccentric anomaly, is given by Kepler's equation

$$\mu(t - T) = E - e \sin E \quad \text{with} \quad \mu = 2\pi/P. \quad (\text{A.4})$$

A.2. Least-squares estimates

On the assumption that P, e, T are known, the determination of the four Thiele-Innes elements is a straightforward problem of fitting a *linear* model to the data. If the observed position at t_n is $(\tilde{x}_n, \tilde{y}_n)$, then estimates for A, B, F, G may be obtained by minimizing

$$\chi^2 = \frac{1}{\sigma^2} \sum_n w_n (x_n - \tilde{x}_n)^2 + \frac{1}{\sigma^2} \sum_n w_n (y_n - \tilde{y}_n)^2 \quad (\text{A.5})$$

where (x_n, y_n) is the theoretical position given by Eq. (A.2). Here w_n is the weight of the n th observation and σ is the standard error of a measurement of unit weight.

This least-squares problem can be solved without iteration since x and y are linear in A, B, F, G . Moreover, the first term in Eq. (A.5) depends only on the pair (A, F) and the second term only on (B, G) . Accordingly, the estimates for A, F are obtained by minimizing the first term in Eq. (A.5) and those for B, G by minimizing the second term. Thus, the problem reduces to solving two independent pairs of linear equations, each with two unknowns.

To simplify the resulting formulae, we define the following quantities

$$\begin{aligned} a &= \sum_n w_n X_n^2 & b &= \sum_n w_n Y_n^2 & c &= \sum_n w_n X_n Y_n \\ r_{11} &= \sum_n w_n \tilde{x}_n X_n & r_{12} &= \sum_n w_n \tilde{x}_n Y_n \\ r_{21} &= \sum_n w_n \tilde{y}_n X_n & r_{22} &= \sum_n w_n \tilde{y}_n Y_n. \end{aligned} \quad (\text{A.6})$$

The least-squares estimates for the Thiele-Innes constants are then

$$\begin{aligned} \hat{A} &= (br_{11} - cr_{12})/\Delta & \hat{F} &= (-cr_{11} + ar_{12})/\Delta \\ \hat{B} &= (br_{21} - cr_{22})/\Delta & \hat{G} &= (-cr_{21} + ar_{22})/\Delta \end{aligned} \quad (\text{A.7})$$

where $\Delta = ab - c^2$.

The above analysis is readily generalized to allow $\tilde{x}_n, \tilde{y}_n$ to have different weights w_n^x, w_n^y . In the formulae for $(\hat{A}, \hat{F})$, the quantities a, b, c, r_{11}, r_{12} are evaluated with $w_n = w_n^x$. Correspondingly, in the formulae for $(\hat{B}, \hat{G})$, the quantities a, b, c, r_{21}, r_{22} are evaluated with $w_n = w_n^y$.

A.3. Covariance matrix

The two sets of normal equations solved above for $(\hat{A}, \hat{F})$ and $(\hat{B}, \hat{G})$ have the same matrix of coefficients and therefore the same inverse. These are

$$\mathbf{M} = \begin{pmatrix} a & c \\ c & b \end{pmatrix} \quad \text{and} \quad \mathbf{M}^{-1} = \begin{pmatrix} +b/\Delta & -c/\Delta \\ -c/\Delta & +a/\Delta \end{pmatrix}. \quad (\text{A.8})$$

From the coefficients of the covariance matrix $\mathbf{M}^{-1}$, we derive

$$\sigma_{A,B}^2 = \frac{b}{\Delta} \sigma^2 \quad \sigma_{F,G}^2 = \frac{a}{\Delta} \sigma^2 \quad (\text{A.9})$$

and

$$\text{cov}(A, F) = \text{cov}(B, G) = -\frac{c}{\Delta} \sigma^2. \quad (\text{A.10})$$

The covariances of other pairings of A, B, F, G are zero. Note that these results apply for fixed P, e, T .

The above analysis shows that (A, F) have a bivariate normal distribution centred on $(\hat{A}, \hat{F})$ with variances and covariance from Eqs. (A.9) and (A.10) – and correspondingly for the pair (B, G) . Accordingly, the pdf $p(A, B, F, G|\phi)$ is a quadrivariate normal distribution centred on $(\hat{A}, \hat{B}, \hat{F}, \hat{G}) = \hat{\psi}(\phi)$. This explicit derivation of $p(\psi|\phi)$ could be used to replace the delta function in Eq. (3).

A.4. Campbell elements

Having derived the least-squares values of A, B, F, G from Eqs. (A.7), we must invert the transformation given by Eq. (A.1). The angle $\omega + \Omega$ is the solution of the equation

$$\omega + \Omega = \arctan\left(\frac{B - F}{A + G}\right) \quad (\text{A.11})$$

such that $\sin(\omega + \Omega)$ has the same sign as $B - F$. Similarly, $\omega - \Omega$ is the solution of the equation

$$\omega - \Omega = \arctan\left(\frac{-B - F}{A - G}\right) \quad (\text{A.12})$$

such that $\sin(\omega - \Omega)$ has the same sign as $-B - F$. This pair of equations can be solved for ω and Ω . But Ω may violate the convention that $\Omega \in (0, \pi)$. Accordingly, if $\Omega < 0$, we set $\Omega = \Omega + \pi$ and $\omega = \omega + \pi$. But if $\Omega > \pi$, we set $\Omega = \Omega - \pi$ and $\omega = \omega - \pi$.

If we now define

$$q_1 = \frac{A + G}{\cos(\omega + \Omega)} \quad \text{and} \quad q_2 = \frac{A - G}{\cos(\omega - \Omega)} \quad (\text{A.13})$$

then the inclination is

$$i = 2 \arctan\left(\sqrt{q_2/q_1}\right) \quad (\text{A.14})$$

and the semi-major axis is

$$a = (q_1 + q_2)/2 \quad (\text{A.15})$$

Appendix B: Bias of posterior means

The numerical search for bias reported in Sect. 5.4 also included the Campbell elements. The results are plotted in Figs. B.1–B.7. The behaviour revealed by these plots is broadly similar to that seen in Fig. 6 except that for some elements bias is evident at somewhat larger f_{orb} . Thus biases $>3s$ occur at $f_{\text{orb}} = 0.5$ for $\langle \log P \rangle$, $\langle \log a \rangle$, $\langle \tau \rangle$ and $\langle e \rangle$.

Fig. B.1. Bias of the posterior mean $\langle \log(P/P_*) \rangle$. The averages and standard errors from Eq. (23) are plotted against f_{orb} with $N = 15$, $\sigma = 0''.05$. The dashed line is the exact value.

Fig. B.2. Bias of the posterior mean $\langle \log(a/a_*) \rangle$. The averages and standard errors from Eq. (23) are plotted against f_{orb} with $N = 15$, $\sigma = 0''.05$. The dashed line is the exact value.

Fig. B.3. Bias of the posterior mean $\langle \tau \rangle$. The averages and standard errors from Eq. (23) are plotted against f_{orb} with $N = 15$, $\sigma = 0''.05$. The dashed line is the exact value.

Fig. B.4. Bias of the posterior mean $\langle e \rangle$. The averages and standard errors from Eq. (23) are plotted against f_{orb} with $N = 15, \sigma = 0'.05$. The dashed line is the exact value.

Fig. B.5. Bias of the posterior mean $\langle i \rangle$. The averages and standard errors from Eq. (23) are plotted against f_{orb} with $N = 15, \sigma = 0'.05$. The dashed line is the exact value.

Fig. B.6. Bias of the posterior mean $\langle \omega \rangle$. The averages and standard errors from Eq. (23) are plotted against f_{orb} with $N = 15, \sigma = 0'.05$. The dashed line is the exact value.

Fig. B.7. Bias of the posterior mean $\langle \Omega \rangle$. The averages and standard errors from Eq. (23) are plotted against f_{orb} with $N = 15, \sigma = 0'.05$. The dashed line is the exact value.

Appendix C: Variation of parameters

In this Appendix, posterior means are computed when selected orbital parameters are varied from those of the standard model; see Eq. (8). Specifically, in one sequence $e = 0.00(0.19)0.95$, and in a second sequence $i = 0^\circ(18^\circ)90^\circ$, with, in each case, the other elements as in Eq. (8). For both sequences, the campaign parameters are $f_{\text{orb}} = 0.7, N = 15, \sigma = 0'.05$, and each orbit is computed with a different random number seed. The orbits are poor and incomplete but f_{orb} is in the domain ($f_{\text{orb}} \gtrsim 0.6$) where bias due to priors is negligible; see Sect. 5.4 and Appendix B.

Results for the e -sequence are in Table C.1. Aspects to note are as follows: The solution for $e = 0$ has large error bars for $\langle \omega \rangle$ and $\langle \tau \rangle$. These arise because ω and $\tau = T/P$ become meaningless and indeterminate as $e \rightarrow 0$ - i.e., for a circular orbit. Also since e is non-negative, $\langle e \rangle$ is positively-biased when the true value is $e = 0$. Despite these issues, an accurate mass is obtained.

At the other extreme of highly elliptical orbits, Table C.1 reveals large error bars for all elements except $\langle e \rangle$.

Results for the i -sequence are in Table C.2. Aspects to note are as follows: The solution $i = 0^\circ$ - i.e., a face-on orbit - has large error bars for ω and Ω . From Eqs. (A.1), we see that when $i = 0^\circ$, $A = G = a \cos(\omega + \Omega)$ and $B = -F = a \sin(\omega + \Omega)$. Accordingly, for an exactly face-on orbit, the difference $\omega - \Omega$ is indeterminate, and the error bars on ω and Ω reflect this. Nevertheless, the mass determination is accurate.

As was the case for $\langle e \rangle$, the quantity $\langle i \rangle$ is positively-biased since i is non-negative. The expected magnitude of this bias can be computed independently of the Bayesian code. If we assume that P, e, τ have the values given in Eq. (8), then the least-squares fit to a synthetic orbit with $i = 0$ is derived with Eqs. (A.7) and then the corresponding Campbell elements a, i, ω, Ω obtained with Eqs. (A.11)–(A.15). With the chosen campaign parameters, $n = 1000$ repetitions of this calculation gives the average value $\bar{i} = 13'.9$, consistent with $14'.4 \pm 4'.0$ in Table C.2. When σ is reduced to $0'.005$ as in Sect. 5.2, $\bar{i} = 4'.4$. Thus, as expected, the bias of i decreases with improved data.

A further point of note in the i -sequence is that the standard error of $\langle \Omega \rangle$ is much smaller than that of $\langle \omega \rangle$ when $i = 90^\circ$ - i.e., an edge-on orbit. Although the error bars of these elements

Table C.1. Sequence with varying eccentricity e .

Q	$e = 0$	0.19	0.38	0.57	0.76	0.95
$\langle \log P \rangle$	1.994 ± 0.013	2.024 ± 0.019	2.014 ± 0.017	2.011 ± 0.018	1.997 ± 0.019	1.976 ± 0.038
$\langle e \rangle$	0.022 ± 0.018	0.237 ± 0.029	0.402 ± 0.025	0.570 ± 0.024	0.783 ± 0.025	0.941 ± 0.026
$\langle \tau \rangle$	0.524 ± 0.249	0.376 ± 0.020	0.382 ± 0.016	0.392 ± 0.016	0.405 ± 0.018	0.424 ± 0.036
$\langle \log a \rangle$	-0.002 ± 0.004	0.003 ± 0.005	-0.005 ± 0.004	-0.014 ± 0.007	0.036 ± 0.018	0.011 ± 0.109
$\langle i \rangle$	$59^\circ 7 \pm 0^\circ 6$	$59^\circ 1 \pm 0^\circ 3$	$60^\circ 0 \pm 0^\circ 6$	$57^\circ 7 \pm 1^\circ 1$	$63^\circ 8 \pm 1^\circ 6$	$62^\circ 6 \pm 5^\circ 5$
$\langle \omega \rangle$	$225^\circ 6 \pm 100^\circ 7$	$246^\circ 9 \pm 4^\circ 0$	$245^\circ 3 \pm 2^\circ 2$	$247^\circ 5 \pm 1^\circ 8$	$253^\circ 4 \pm 1^\circ 6$	$247^\circ 0 \pm 7^\circ 3$
$\langle \Omega \rangle$	$122^\circ 1 \pm 1^\circ 1$	$120^\circ 9 \pm 1^\circ 3$	$121^\circ 8 \pm 1^\circ 1$	$121^\circ 2 \pm 1^\circ 3$	$117^\circ 8 \pm 1^\circ 3$	$125^\circ 4 \pm 5^\circ 5$
$\langle \log \mathcal{M} \rangle$	0.011 ± 0.026	-0.048 ± 0.037	-0.028 ± 0.034	-0.022 ± 0.036	0.007 ± 0.039	0.059 ± 0.124

Table C.2. Sequence with varying inclination i .

Q	$i = 0^\circ$	18°	36°	54°	72°	90°
$\langle \log P \rangle$	1.998 ± 0.013	1.999 ± 0.013	2.043 ± 0.018	1.994 ± 0.014	2.011 ± 0.017	1.965 ± 0.015
$\langle e \rangle$	0.504 ± 0.016	0.532 ± 0.016	0.546 ± 0.019	0.482 ± 0.020	0.541 ± 0.032	0.441 ± 0.045
$\langle \tau \rangle$	0.405 ± 0.013	0.396 ± 0.012	0.361 ± 0.016	0.406 ± 0.014	0.395 ± 0.016	0.425 ± 0.015
$\langle \log a \rangle$	0.006 ± 0.003	0.023 ± 0.003	0.001 ± 0.005	0.000 ± 0.004	0.007 ± 0.008	-0.030 ± 0.008
$\langle i \rangle$	$14^\circ 6 \pm 4^\circ 0$	$25^\circ 9 \pm 2^\circ 4$	$32^\circ 6 \pm 1^\circ 6$	$51^\circ 0 \pm 1^\circ 0$	$73^\circ 2 \pm 0^\circ 7$	$91^\circ 2 \pm 0^\circ 2$
$\langle \omega \rangle$	$263^\circ 2 \pm 56^\circ 4$	$260^\circ 4 \pm 4^\circ 9$	$245^\circ 6 \pm 4^\circ 5$	$251^\circ 2 \pm 1^\circ 8$	$252^\circ 0 \pm 1^\circ 6$	$249^\circ 5 \pm 2^\circ 6$
$\langle \Omega \rangle$	$94^\circ 6 \pm 25^\circ 1$	$108^\circ 8 \pm 5^\circ 4$	$123^\circ 9 \pm 4^\circ 4$	$117^\circ 7 \pm 1^\circ 4$	$119^\circ 5 \pm 0^\circ 6$	$118^\circ 1 \pm 0^\circ 1$
$\langle \log \mathcal{M} \rangle$	0.004 ± 0.026	0.003 ± 0.026	-0.087 ± 0.037	0.012 ± 0.028	-0.021 ± 0.035	0.069 ± 0.029

are somewhat underestimated (Sect. 5.5), this is a real effect. From Eqs. (A.1) and (A.2), it follows that the exact edge-on orbit obeys the equation $y/x = \tan \Omega$, so that Ω is determined by the slope of the straight line fitted to $\tilde{x}_n, \tilde{y}_n$. The Bayesian code does not formally make such a fit, but this mathematical aspect is implicit.

References

- Aitken, R. G. 1918, *The Binary Stars* (New York: D. C. McMurtrie), 92
Casertano, S., Lattanzi, M. G., Sozzetti, A., et al. 2008, *A&A*, 48, 699
Dawid, A. P. 1982, *Journal of the American Statistical Association*, 77, 605
Eggen, O. J. 1962, *QJRAS*, 3, 259
Eggen, O. J. 1967, *ARA&A*, 5, 105
Ford, E. B., & Gregory, P. C. 2007, *ASPC*, 371, 189
Hartkopf, W. I., McAlister, H. A., & Franz, O. G. 1989, *AJ*, 98, 1014
Heintz, W. D. 1978, *Double Stars* (Reidel, Dordrecht), 63
Lattanzi, M. G., Spagna, A., Sozzetti, A., & Casertano, S. 2000, *MNRAS*, 317, 211
Liddle, A. R. 2009, *Ann. Rev. Nucl. Particle Sci.*, 59, 95
Little, R. 2006, *The American Statistician*, 60, 1
Mikkelsen, K., Nss, S. K., & Eriksen, H. K. 2012, *ApJ*, submitted
Nurmi, P. 2005, *ESA SP*, 576, 463
Pourbaix, D. 2002, *A&A*, 385, 686
Pourbaix, D. 2011, *AIP Conf. Proc.*, 1346, 122
Ranucci, G. 2012, *Nucl. Instr. Meth. Phys. Res. A*, 661, 77
Schaefer, G. H., Simon, M., Beck, T. L., Nelan, E., & Prato, L. 2006, *AJ*, 132, 2618
Severini, T. A. 2000, *Likelihood Methods in Statistics* (Oxford: OUP)
Wright, J. T., & Howard, A. W. 2009, *ApJS*, 182, 205