UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA ELÉCTRICA

BAYESIAN INFERENCE IN HIERARCHICAL STELLAR SYSTEMS

TESIS PARA OPTAR AL GRADO DE
MAGÍSTER EN CIENCIAS DE LA INGENIERÍA, MENCIÓN ELÉCTRICA

MEMORIA PARA OPTAR AL TÍTULO DE
INGENIERO CIVIL ELÉCTRICO

MIGUEL IGNACIO VIDELA ARAYA

PROFESOR GUÍA:
JORGE SILVA SÁNCHEZ

PROFESOR CO-GUÍA:
RENÉ MÉNDEZ BUSSARD

MIEMBROS DE LA COMISIÓN:
MARCOS ORCHARD CONCHA
GAËL CHAUVIN

Este trabajo ha sido parcialmente financiado por FONDECYT 1170854, FONDECYT 1210315 y Proyecto Basal AC3E FB0008

SANTIAGO DE CHILE
2021

RESUMEN DE LA TESIS PARA OPTAR AL GRADO DE
MAGÍSTER EN CIENCIAS DE LA INGENIERÍA ELÉCTRICA
Y AL TÍTULO DE INGENIERO CIVIL ELÉCTRICO
POR: MIGUEL IGNACIO VIDELA ARAYA
FECHA: 2021
PROF. GUÍA: JORGE SILVA SÁNCHEZ PROF. CO-GUÍA: RENÉ MÉNDEZ BUSSARD

INFERENCIA BAYESIANA EN SISTEMAS ESTELARES JERÁRQUICOS

Se presenta una metodología para la inferencia bayesiana en sistemas estelares binarios basada en el algoritmo Markov chain Monte Carlo No-U-Turn sampler, que proporciona una estimación precisa y eficiente de la distribución posterior conjunta de los parámetros orbitales. La metodología bayesiana permite incorporar directamente información previa sobre el sistema para restringir la solución y estimar parámetros orbitales que no se pueden determinar por falta de observaciones o mediciones imprecisas. La incorporación de información previa del paralaje y la masa del objeto primario se estudia extensamente para determinar las masas individuales de los componentes de sistemas binarios espectroscópicos visuales de una sola línea. Este estudio se realiza analizando las distribuciones posteriores y su respectiva proyección en los espacios de observación. La metodología se extiende para la inferencia bayesiana en sistemas estelares jerárquicos de cualquier multiplicidad, arquitectura y falta de fuentes de observación. Finalmente, se propone una metodología para determinar el tiempo óptimo de medición en sistemas binarios y jerárquicos basado en el criterio de muestreo de máxima entropía. Esta metodología hace uso directo de la distribución posterior estimada para proporcionar una caracterización temporal de la información obtenida de nuevas observaciones del sistema y estima una distribución de probabilidad del tiempo de medición óptimo.

RESUMEN DE LA TESIS PARA OPTAR AL GRADO DE
MAGÍSTER EN CIENCIAS DE LA INGENIERÍA ELÉCTRICA
Y AL TÍTULO DE INGENIERO CIVIL ELÉCTRICO
POR: MIGUEL IGNACIO VIDELA ARAYA
FECHA: 2021
PROF. GUÍA: JORGE SILVA SÁNCHEZ PROF. CO-GUÍA: RENÉ MÉNDEZ BUSSARD

BAYESIAN INFERENCE IN HIERARCHICAL STELLAR SYSTEMS

A methodology for Bayesian inference in binary stellar systems based on the No-U-Turn sampler Markov chain Monte Carlo algorithm is presented, providing a precise and efficient estimation of the joint posterior distribution of the orbital parameters. The Bayesian methodology allows to directly incorporate prior information about the system to constrain the solution and estimate orbital parameters that cannot be determined due to lack of observations or imprecise measurements. The incorporation of prior information of the parallax and the primary object mass is extensively studied to determine the individual masses of the components of single-lined visual-spectroscopic binary systems. This study is made by analyzing the posterior distributions and their respective projection in the observation spaces. The methodology is extended for the Bayesian inference in hierarchical stellar systems of any multiplicity, architecture, and lack of observation sources. Finally, a methodology to determine the optimal measurements time in binary and hierarchical systems is proposed based on the maximum entropy sampling criterion. This methodology makes direct use of the estimated posterior distribution to provide a temporal characterization of the information gain of new observations of the system and estimates a probability distribution of the optimal measurement time.

Agradecimientos

Inicio este breve extracto con lo típico: un reconocimiento a mi familia, ya que sea para bien o para mal, su influencia configura en gran medida mi persona y mi obra (y por tanto esta tesis). Agradezco especialmente a mis abuelos Haydée y Miguel (mi mami y mi papi). Su apoyo incondicional y ejemplo de virtud constituyeron una invaluable y muy necesaria guía de cómo debe ser una buena y correcta persona. A dicho ejemplo le debo algunos de mis principales valores: la responsabilidad y la disciplina, sin los cuales, el presente trabajo conclusivo de una extenuante carrera universitaria jamás hubiese existido.

Imposible no reconocer también a los contribuidores en la concepción de este modesto documento: a mis compañeros/as del departamento de ingeniería eléctrica, al grupete de "inteligentes" de especialidad, a mis nobles *borraxitos*, a los distinguidos miembros del grupo de investigación IDS y a mis excelsos profesores de tesis. Sin su permanente compañía, constante apoyo, extraordinaria disposición y a las innumerables, variadas y enriquecedoras discusiones entabladas, este trabajo *casi seguramente* no hubiese conocido buen término.

Para preservar el balance de este emotivo relato, procedo a exponer a la luz pública a todos los individuos que intentaron sabotear este trabajo: a mis amistades del colegio, de la universidad y de la vida, a la prestigiosa *liga* (de no se qué), al nefasto *team montixello 5am*, a la insensata *guarida del gato* y a mis despreciables *borraxitos*. Sus permanentes propuestas reñidas con la ética, la moral, las buenas constumbres, la legislación vigente y el sentido común, pusieron en permanente riesgo la conclusión de este trabajo de tesis, mi integridad y mi futuro. Gracias por ser la vía de escape de esta ajetreada, agobiante e ingrata rutina universitaria.

A todas las personas que siempre están, a las que están a veces y a las que alguna vez estuvieron, sepan que su huella en este humilde y exasperante sujeto es y será siempre valorada enormemente.

Muchas gracias.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Motivation

Mass is the most critical parameter in determining the structure and evolution of stars. In binary stellar systems, the masses of their individual components (stars) can be directly calculated through the orbital parameters. These can be determined through Kepler's laws using astrometric and spectroscopic observations. In some spectroscopic binary systems, the spectral lines of both components are visible (double-line spectroscopic binaries). However, in most cases, only one component can be seen in the spectra (single-lined spectroscopic binaries). In the absence of companion star spectra, the mass ratio of the stellar binary system can not be determined. This lack of information limits the astrophysical study of these abundant stellar systems.

A similar problem occurs in the case of hierarchical stellar systems, where the system is decoupled into multiple binary subsystems that interact with each other gravitationally, forming different stellar architectures. The binary subsystems can be very different in size, period, and brightness of their components depending on the hierarchical system. In addition, only some of the subsystems can be observed through observational tools with acceptable precision. In practice, the observations are available only for some binary subsystems that conforms the hierarchical system, making the joint estimation of the orbital parameters of the whole system intractable. Fortunately, the joint estimation can be performed, in both the single-lined spectroscopic binary and the hierarchical systems, if some suitable external information about the system is provided.

The present work introduces a Bayesian methodology based on the No-U-Turn sampler inference tool to address the orbital parameters estimation problem in two scenarios: the individual masses determination on single-lined spectroscopic binaries and the joint orbital parameters estimation on hierarchical systems of any multiplicity, architecture, and lack of observation sources. This methodology provides a precise characterization of the uncertainty of the estimates in the form of the joint posterior distribution of the orbital parameters addressing the lack of observation sources by incorporating suitable prior distributions on some critical parameters of the system. The methodology is evaluated on several systems, providing an exhaustive analysis of the obtained results. For this, a comparison of the posterior distribution is performed by applying different priors and the corresponding projected posterior distributions on the observational space.

In order to highlight the relevance of the posterior distribution characterization provided by the proposed methodology, the current works additionally introduce a Bayesian-based framework for the determination of the optimal time of measurement in binary systems. The optimal time of measurement is implemented using the maximum entropy sampling principle. This information-driven principle allows characterizing the measurement uncertainty in a dense range of time with a low computationally cost, providing an estimation of the probability distribution of time of optimal measurement. This work also shows that the proposed sampling methodology is indeed optimal in the Bayesian sense.

1.2 Hypotheses

All the developed work is based on the following hypotheses:

- No other force besides the two bodies' mutual gravitational interaction perturbs the system in the binary stellar systems. Both bodies present symmetric spherical geometry that allows to treat them as point masses.
- The dynamics of the multiple stellar systems studied in this work (hierarchical stellar systems) can be approximated by decomposing the system into subgroups of stars that form multiple binary systems.
- The whole set of possible new measurements presents the same uncertainty in the Bayesian-based methodology for estimating the optimal measurement time.
- All the algorithms used in this work (for optimization, sampling, and estimation) are efficient, in the sense that few iterations are required to reach accurate solutions.

1.3 Objectives

1.3.1 Main Objectives

The main objectives of this work are the following:

- The implementation of an efficient and precise Bayesian inference framework to estimate the orbital parameters of binary and hierarchical stellar systems.
- The mathematical formalization, development and implementation of a Bayesian-based methodology to estimate the optimal measurement time in binary and hierarchical stellar systems.

1.3.2 Specific Objectives

The specific objectives of this work are the following:

- Analyze the estimated posterior distributions provided by the proposed Bayesian inference framework (in the parameters and observational spaces) using astrometric and radial velocities observations of several binary and hierarchical stellar systems.
- Analyze and evaluate the performance of the proposed Bayesian inference framework for the estimation of the individual masses of single-lined visual-spectroscopic binary systems, providing an extensive study of the effect of the incorporation of prior information on the parallax and the mass of the primary component of the system.

- Formulate succinct analytic expressions of the kinematic equations of hierarchical stellar systems that generalize to any architecture and multiplicity of these systems.
- Analyze and evaluate the performance of the proposed Bayesian inference framework for the estimation of the orbital parameters in hierarchical stellar systems with different architectures, multiplicities, and sources of available observations, studying the effect of the incorporation of suitable prior information on the system.
- Analyze and validate the proposed methodology for estimating the optimal measurement time in binary and hierarchical stellar systems, contrasting the results obtained on some binary and hierarchical stellar systems with the mathematical theory that supports it.

1.4 Structure

The structure of the document is as follows:

- Section 2 describes the Keplerian orbital models used to characterize the binary and hierarchical stellar systems.
- Section 3 presents the fundamental concepts related to the Bayesian inference, as well as a theoretical review on the Markov chain Monte Carlo sampling method for Bayesian inference. This section also presents the most relevant Markov chain Monte Carlo algorithms, including the No-U-Turn sampler used in this work.
- Section 4 presents a brief chronological review of some of the most relevant works that addressed the problem of orbital parameters estimation on binary stellar systems.
- Section 5 addresses the problem of the individual masses estimation on single-lined spectroscopic binaries, describing the Bayesian model considered in this work and exploring two sources of prior information: the parallax and the mass of the primary object. This section also provides an extensive experimental evaluation by comparing the posterior distributions obtained and their projection on the observational space, ending with applying the methodology on previously unresolved single-lined spectroscopic binaries.
- Section 6 addresses the problem of the estimation of the orbital parameters on hierarchical stellar systems, describing the Bayesian model considered and evaluating the methodology on several hierarchical systems of different multiplicities, architectures, and lack of observational sources.
- Section 7 presents a direct application of the Bayesian inference tool developed in this work by introducing a Bayesian-based optimal measurement time strategy for binary systems. This section covers the theoretical guarantees and capacity of the proposed method. It also introduces an experimental framework to evaluate the method on some binary and hierarchical systems within an exhaustive analysis of the obtained results.
- Section 8 presents the main conclusions of the developed work and provides some guidelines on future extensions and new applications of the proposed methodologies.

Chapter 2

Stellar System Model

2.1 Binary Star System

A binary star system is defined as a stellar system composed of two stars that are bounded by their mutual gravitational attraction [Heintz, 1975].

Assuming there is no other force besides the mutual gravitational interaction of the two bodies perturbing the system[1] and both bodies present symmetric spherical geometry that allows to treat them as point masses, the mechanics of the system can be simplified to a *two-body problem*. Let m_1, r_1 and m_2, r_2 be the mass and position vector of each body. Newton's third law implies that the force experimented by each body is determined by:

$$\begin{aligned} m_1\ddot{\vec{r}}_1 &= \frac{Gm_1m_2}{r^2}\hat{r}, \\ m_2\ddot{\vec{r}}_2 &= -\frac{Gm_1m_2}{r^2}\hat{r}. \end{aligned} \tag{2.1}$$

Considering the position of the primary object 1 relative to its companion 2, i.e., $\vec{r} = \vec{r}_2 - \vec{r}_1$, the force experimented by the relative position is:

$$\ddot{\vec{r}} = \ddot{\vec{r}}_2 - \ddot{\vec{r}}_1 = -\frac{G(m_1 + m_2)}{r^2}\hat{r}, \tag{2.2}$$

which solution is a Keplerian elliptical orbit determined by:

$$r(\nu) = \frac{a(1 - e^2)}{1 + e\cos\nu}. \tag{2.3}$$

In (2.3), a determines the semi-major axis of the orbit, e the eccentricity of the orbit and ν the true anomaly defined as the angle between the current position of the companion object in the orbit and the periapsis (point of closest distance between both bodies in the orbit).

Depending on the source of observational data, the binary system is classified as:

[1] Neglecting the effects of gravitational interaction with external objects, mass transference, and complex relativistic phenomena, among others.

- Visual binary system, if positional observations are available.
- Spectroscopic binary system, if radial velocities observations are available.
- Spectroscopic-binary or combined system, if both sources of observations are available.

Different formulations are developed to describe the observational data depending on the case.

2.1.1 Visual Binary System

A visual binary (or visual double system) corresponds to a binary system in which the relative position of both components is observable. The positional or astrometric observations in binary systems measure the position of the fainter companion object relative to the brighter primary object. Depending on the technique used to obtain the astrometric observations, these can be classified into micrometric, interferometric or photographic position observations.

The solution of the differential equation of the motion (Eq. (2.2)) is described by the Kepler laws, wherein the case of binary stars, it corresponds to an elliptical orbit where the primary star is in the focus and the area swept by the radius vector is constant per unit of time. This elliptical orbit, denoted as real orbit, is characterized by four orbital parameters:

- **Period (P):** The revolution period in years.
- **Time of periastron passage (T):** One epoch of passage through periastron (minimum true distance between the components) in years and fractions.
- **Semimajor axis (a):** The major semiaxis of the elliptical true orbit in seconds of arc.
- **Eccentricity (e):** The numerical eccentricity.

The astrometric observations are position measurements of the projection of the real orbit in the plane of the sky relative to the observer (plane of reference), denoted as apparent orbit. Three additional parameters are necessary to project the real orbit into the apparent orbit:

- **Longitude of the ascending node (Ω):** The position angle from a reference direction to the ascending node[2] in the plane of reference (ranging from 0° to 360°).
- **Argument of periapsis (ω):** The angle from the node to the periastron in the real orbit, following the direction of motion (ranging from 0° to 360°).
- **Inclination (i):** The angle between the plane of projection and that of the true orbit (ranging from 0° to 180°).

An illustration of the angular parameters of the apparent orbit is presented in Figure 2.1.

It is worth pointing out that two values for the longitude of the ascending node (Ω and $\Omega + \pi$) result in identical apparent orbits. Therefore, the ascending node cannot be identified by positional observations. By convention in astronomy, if the ascending node is undetermined, the value of Ω is ranged from 0° to 180°.

On the specifics, the position on the apparent orbit (ρ, θ) at a certain time t (the ephemerides formulae) involves the determination of the position in the real orbit and its projection to the apparent orbit. The position on the real orbit involves the determination of the three orbital

[2]Point where the real orbit of the object passes through the plane of reference.

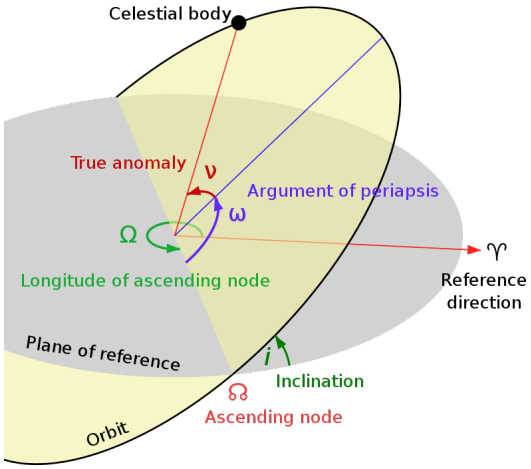

Figure 2.1: Angular parameters of the apparent orbit (`https://commons.wikimedia.org/wiki/File:Orbit1.svg`).

anomalies: the true anomaly $\nu(t)$, the eccentric anomaly $E(t)$ and the mean anomaly $M(t)$, in terms of the orbital parameters P, T, a and e.

The true anomaly $\nu(t)$ is defined as the angle between the periapsis and the current position of the companion object in the orbit, seen from its main focus (the position of the primary object). The following geometrical identity can determine this angular parameter:

$$\tan\frac{\nu(t)}{2} = \sqrt{\frac{1+e}{1-e}}\tan\frac{E(t)}{2}. \tag{2.4}$$

where $E(t)$ is the eccentric anomaly. The eccentric anomaly is defined as the angle between the periapsis and the intersection of a perpendicular line to the semi-major axis of the orbit and the position of the companion object in the orbit, seen from it central point. This angular parameter can be determined by the numerical resolution of the Kepler equation[3]:

$$M(t) = \frac{2\pi(t-T)}{P} = E(t) - e\sin E(t), \tag{2.5}$$

where $M(t)$ is the mean anomaly of the orbit. The mean anomaly represents the angular movement of the companion object in the orbit (similar to the true anomaly) but at a uniform rate. The uniform rate of movement is represent by a circle circumscribed to the orbit. Therefore, the mean anomaly is defined as the angle between the periapsis and the point in the circle circumscribed to the orbit. It seems from its central point, which covers the same area as the true anomaly. This angular parameter corresponds to the revolution period of the companion object in the orbit, i.e., $M(t) = 2\pi(t-T)/P$. An illustration of the orbital anomalies is presented in Figure 2.2.

Finally, the position on the real orbit is projected to the apparent orbit through the angular parameters ω, Ω, and i:

$$\begin{aligned}\tan(\theta(t)-\Omega) &= \tan(\nu(t)+\omega)\cos(i),\\ \rho(t) &= r(t)\cos(\nu(t)+\omega)\sec(\theta(t)+\Omega),\end{aligned} \tag{2.6}$$

[3]The Kepler equation is commonly resolved using the Newton-Raphson method [Ypma, 1995].

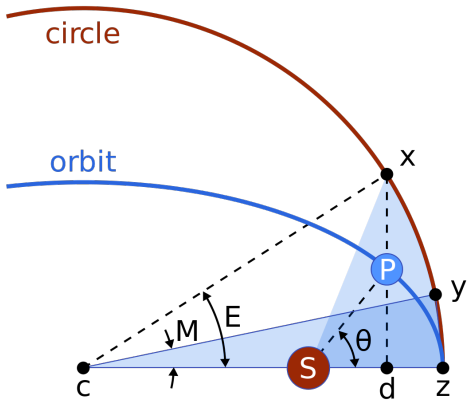

Figure 2.2: Orbital anomalies (`https://commons.wikimedia.org/wiki/File:Mean_Anomaly.svg`). In blue the elliptical orbit and in red the circumscribed circle. In the figure: θ is the true anomaly (ν in this work); E is the eccentric anomaly; M is the mean anomaly, C is the center of the orbit; S is the main focus of the orbit (the position of the primary object) and P is the position of the companion object on the orbit.

with $r(t) = a(1-e^2)/(1+e\cos(\nu(t)))$ the radius vector.

The procedure to compute the position on the apparent orbit in rectangular coordinates (x, y) at a certain time t involves determining the normalized rectangular coordinates in the true orbit X, Y:

$$\begin{aligned} X(t) &= \cos E(t) - e, \\ Y(t) &= \sqrt{1-e^2}\sin E(t), \end{aligned} \tag{2.7}$$

with $E(t)$ the eccentric anomaly determined in (2.5). Therefore, the position on the true orbit is computed by ponderation of the normalized coordinates:

$$\begin{aligned} x(t) &= AX(t) + FY(t), \\ y(t) &= BX(t) + GY(t), \end{aligned} \tag{2.8}$$

with A, B, F and G the Thiele-Innes elements defined as:

$$\begin{aligned} A &= a(\cos\omega\cos\Omega - \sin\omega\sin\Omega\cos i), \\ B &= a(\cos\omega\sin\Omega + \sin\omega\cos\Omega\cos i), \\ F &= a(-\sin\omega\cos\Omega - \cos\omega\sin\Omega\cos i), \\ G &= a(-\sin\omega\sin\Omega + \cos\omega\cos\Omega\cos i). \end{aligned} \tag{2.9}$$

The terms $(A/a, B/a)$ and $(F/a, G/a)$ are interpreted as the direction cosines of the major and minor axis, respectively, of the orbit in the rectangular coordinate system formed by the tangential plane and the North direction. The Thiele-Innes elements form a one-to-one correspondence with the elements a, Ω, ω, i.

2.1.2 Spectroscopic Binary System

A spectroscopic binary system corresponds to a binary system in which the spectral lines of the light emitted by its components are observable. The movement of the stars in the orbit produces

variations on the spectral lines observed due to the Doppler's effect: Blue or red spectra are perceived, respectively, when the stars move towards or away from the observer. The Doppler shift of the components' spectral lines measured through a spectrometer result in radial velocities observation of the objects.

The radial velocity V at a certain time t can be calculated as the sum of the radial velocity of the system center of mass V_{CoM} (a constant) and the radial part of the orbital velocity of the observed component relative to the center of mass $\dot{z} = dz/dt$:

$$V(t) = V_0 + \dot{z}(t). \tag{2.10}$$

Following the resolution of the *two-body problem*, the radial component of the system can be expressed as:

$$z(t) = r\sin(\nu(t) + \omega)\sin(i), \tag{2.11}$$

and therefore, by taking the first temporal derivative, we have that:

$$\dot{z}(t) = \frac{2\pi a \sin(i)}{P\sqrt{1-e^2}}[e\cos\omega + \cos(\nu(t) + \omega)]. \tag{2.12}$$

Denoting $K = 2\pi a \sin i/(P\sqrt{1-e^2})$ as the semi-amplitude of the radial velocity curve, the above expression becomes:

$$V(t) = V_0 + K[e\cos\omega + \cos(\nu(t) + \omega)]. \tag{2.13}$$

Therefore, the radial velocity $V(t)$ is characterized by six orbital parameters:

- **Period (P):** The revolution period in days.
- **Time of periastron passage (T):** One epoch of passage through periastron in Julian Date (J.D.).
- **Eccentricity (e):** The numerical eccentricity.
- **Argument of periapsis (ω):** The periastron longitude, counted from the maximum of the radial velocity curve.
- **Semi-amplitude (K):** The semi-amplitude of the radial velocity curve in km/s.
- **Velocity of center of mass (V_0):** The radial velocity of the center of mas of the system in km/s.

Although the parameters P and T do not appear directly in the expression of the radial velocity (2.13), they are implicit in the determination of the true anomaly ν (2.4). According to the convention for the units of the orbital parameters involved, the semi-amplitude K is measured in km/s, whereby $a\sin(i)$ must be measured in km and P must be converted to seconds through $P[s] = 86400 \cdot P[days]$.

The expression (2.13) is derived with respect to the relative orbit $\vec{r} = \vec{r}_2 - \vec{r}_1$, however, the radial velocity observations are relative to the center of mass of the system. To correct this discrepancy, the semi-major axis a of the relative orbit must be replaced by their counterparts a_1 and a_2 of the components relative to the center of mass of the system.

By the definition of the center of mass of a system composed of two particles of mass m_1 and m_2, the following relation is directly obtained:

$$m_1 a_1 = m_2 a_2, \tag{2.14}$$

and noticing that $a = a_1 + a_2$, the following expressions for a_1 and a_2 are obtained:

$$\begin{aligned} a_1 &= a \cdot \frac{m_2}{m_1 + m_2} = a \cdot \frac{q}{1+q}, \\ a_2 &= a \cdot \frac{m_1}{m_1 + m_2} = a \cdot \frac{1}{1+q}, \end{aligned} \tag{2.15}$$

where $q = m_2/m_1$ is defined as the mass ratio. Consequently, the semi-amplitude of each component becomes:

$$\begin{aligned} K_1 &= \frac{1}{86400} \frac{2\pi a_1 \sin(i)}{P\sqrt{(1-e^2)}}, \\ K_2 &= \frac{1}{86400} \frac{2\pi a_2 \sin(i)}{P\sqrt{(1-e^2)}}. \end{aligned} \tag{2.16}$$

Finally, noting that the argument of periapsis of the primary ω_1 and companion ω_2 components of the system differs by 180°, we have that $\omega = \omega_1 = \omega_2 + \pi$, the expression for the radial velocities of each component of the binary system becomes:

$$\begin{aligned} V_1(t) &= V_0 + K_1[e\cos(\omega) + \cos(\nu(t) + \omega)], \\ V_2(t) &= V_0 - K_2[e\cos(\omega) + \cos(\nu(t) + \omega)]. \end{aligned} \tag{2.17}$$

It is important to note that the terms $a_1 \sin(i)$ and $a_2 \sin(i)$ in the definition of the semi-amplitudes (2.16) can not be separated through radial velocities observations.

When the spectra of both components are distinguishable, i.e., the radial velocities of the primary and the companion object are observable, the system is denoted as **double-lined spectroscopic binary** ($SB2$). It is characterized by the set of orbital parameters $\vartheta_{SB2} = \{P, T, e, K_1, K_2, V_0\}$. However, this is an infrequent case since most of the spectroscopic binary systems observations are only from the primary (brighter) object ($\sim 80\%$). When the primary object spectra is the only observable, the system is denoted as **single-lined spectroscopic binary** ($SB1$) and is characterized by the set of orbital parameters $\vartheta_{SB2} = \{P, T, e, K_1, V_0\}$, since the parameter K_2 is undetermined.

2.1.3 Visual-Spectroscopic Binary System

The visual-spectroscopic binary system corresponds to a binary system in which the relative position and the radial velocities of its components are observable. Since four orbital parameters P, T, e, ω are common in the visual and spectroscopic binary systems, the schemes are dependent, and joint modeling that describes both sources of information (positional and radial velocity observations) allows to determine the ambiguities and indeterminacy of each scheme. Furthermore, the joint modeling allows to determine the individual masses of the system.

The radial velocities observations allow solving the indeterminacy of the longitude of the ascending node Ω in the visual binary case since the maximum/minimum of the radial velocity curve

of each component is reached in the ascending/descending node. Conversely, the positional observations allow decoupling the term $a\sin(i)$ in the spectroscopic binary case since it allows to determine the inclination i of the orbit.

The individual masses of a binary system can be computed by determining the total mass of the system $m_1 + m_2$ and the mass ratio $q = m_2/m_1$. According to the Third Law of Kepler, the total mass of the system can be obtained through the following expression:

$$m_1 + m_2 = \frac{a[AU]^3}{P^2}, \tag{2.18}$$

where a[AU] is the relative semi-major axis of the system (in astronomical units), P the period of the system (in seconds) and m_1, m_2 the mass of the primary and the companion object (in solar masses), respectively. Since the positional observations allow to determine the semi-major axis a in angular distance units (seconds of arc), the conversion to linear distance units (AU) is determined by the following expression:

$$a[AU] = \frac{a[''] }{\pi}, \tag{2.19}$$

where π is the system parallax (in seconds of arc), an additional orbital parameter required to determine the individual masses.

For the computation of the mass ratio q, the combination of positional and radial velocities observations are required. Considering that the radial velocity observations can determine the terms $a_1\sin(i)$ and $a_2\sin(i)$ and the positional observations can determine the inclination i, both sources of observations allow to determine the individual semi-major axis of each component (a_1 and a_2) and the mass ratio $q = a_1/a_2$ (2.14).

Considering the equation (2.19), the radial velocity expression in (2.17) becomes[4]:

$$\begin{aligned} V_1(t) &= V_0 + \frac{2\pi a_1 \sin i}{P\sqrt{1-e^2}}[\cos(\omega + \nu(t)) + e\cos\omega], \\ V_2(t) &= V_0 - \frac{2\pi a_2 \sin i}{P\sqrt{1-e^2}}[\cos(\omega + \nu(t)) + e\cos\omega], \end{aligned} \tag{2.20}$$

with $a_1 = a''/\pi \cdot q/(1+q)$, $a_2 = a''/\pi \cdot 1/(1+q)$ and a'' the semi-major axis in arcseconds.

If the radial velocity observations of each component ($V_1(t)$ and $V_2(t)$) are available ($SB2$ case), the combined model that describes the positional and radial velocities observations is characterized by the set of orbital parameters $\vartheta_{VB+SB2} = \{P, T, e, a, \omega, \Omega, i, V_0, \pi, q\}$. However, if the radial velocity observations of only one component are available ($SB1$ case), the parameters q and π cannot be simultaneously determined.

2.2 Hierarchical System

A stellar system composed of more than two stars is called a multiple stellar system. The dynamics of a multiple stellar system composed of $n > 2$ stars can be modeled as a n-body problem. However,

[4]The units conversion between the common orbital parameters of Section 2.1.1 and Section 2.1.2 is omitted for simplicity.

unlike the 2-body problem model for binary stars systems, no general closed-form solution exists for its dynamics, as the resulting system is chaotic for most initial conditions. Thereby, numerical approximations of the solution are required. Unfortunately, the high computation time involved in solving this problem is prohibitive. To avoid such difficulties, the multiple stellar system can be modeled as an hierarchical system, which dynamic is a particular case of the n-body problem.

The hierarchical system approximation successively decomposes the system into two subgroups of stars whose dynamics follow the solution of the *two-body problem*. This results in a hierarchical structure of binary subsystems. As a consequence of the little interaction imposed by the hierarchical structure, the dynamics of each component follows an approximately stable Keplerian orbit around the center of mass of the system.

2.2.1 Triple Hierarchical System

The simpler hierarchical system is composed of three stars or components. Two different hierarchical structures composed of two binary systems are possible: An outer object orbiting an inner binary system and an outer binary system orbiting an inner object. An illustration of the two possible triple hierarchical system configurations is presented in Figure 2.3.

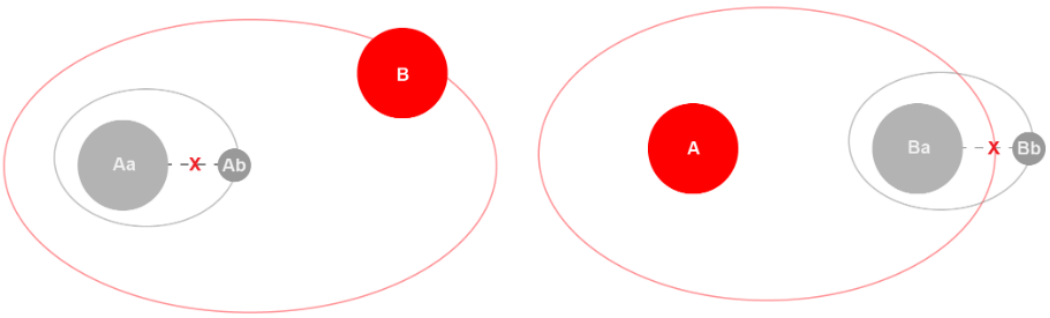

Figure 2.3: Left: Outer object B orbiting an inner binary system A_aA_b. Right: Outer binary system B_aB_b orbiting an inner object A. Reprinted from [Villegas et al., 2021].

2.2.1.1 Visual Hierarchical System

Following the hierarchical approximation mentioned above, two binary systems compose a triple system: an inner and an outer system. The positional observations of both binaries are relative to the most massive and brighter component. However, in some cases, the inner system can not be resolved, i.e., it is not possible to identify each of its components. Two different derivations must be done for each case to adequately describe the available positional observations: the resolved inner binary case and the unresolved inner binary case. These are described next.

2.2.1.1.1 Resolved Inner Binary

Considering the first hierarchical structure: let A_a and A_b be the components of the inner binary system (A_aA_b). The motion of the third component is described by a Keplerian orbit around the center of mass of the inner system A, forming the outer binary system AB.

According to the 2-body problem approximation, the force experimented by the companion

object A_b relative to the primary object A_a is given by:

$$\ddot{\vec{r}}_{A_aA_b} = \ddot{\vec{r}}_{A_b} - \ddot{\vec{r}}_{A_a} = -\frac{G(m_{A_a} + m_{A_b})}{r^2_{A_aA_b}}\hat{r}_{A_aA_b}. \tag{2.21}$$

In the same manner, the force experimented by the relative position of the external object B relative to the center of mass of the binary system A_aA_b is given by:

$$\ddot{\vec{r}}_{AB} = \ddot{\vec{r}}_{B} - \ddot{\vec{r}}_{A} = -\frac{G(m_{A} + m_{B})}{r^2_{AB}}\hat{r}_{AB}. \tag{2.22}$$

As explained in Section 2.1.1, the solution of equations (2.21) and (2.22) are Keplerian elliptical orbits characterized by the set of orbital parameters $\{P_{A_aA_b}, T_{A_aA_b}, e_{A_aA_b}, a_{A_aA_b}, \omega_{A_aA_b}, \Omega_{A_aA_b}, i_{A_aA_b}\}$ and $\{P_{AB}, T_{AB}, e_{AB}, a_{AB}, \omega_{AB}, \Omega_{AB}, i_{AB}\}$, respectively. The procedure to determine the position in those orbits at a certain time t follows the ephemerides formuale described in equations (2.6) for polar coordinates and in (2.7) for rectangular coordinates.

The observations of the inner and outer binary systems are generally measured relative to the brighter and most massive component of the system A_a (principal component). Hence, as the solution of the inner system $\vec{r}_{A_aA_b}$ is already relative to A_a, it is necessary to characterize the position vector $\vec{r}_{A_aB}$ to describe the observations of the outer binary system.

The vector $\vec{r}_{A_aB}$ can be expressed as:

$$\vec{r}_{A_aB} = \vec{r}_{A_aA} + \vec{r}_{AB}. \tag{2.23}$$

Decomposing the vector $\vec{r}_{A_aA}$ into its components and noting that $\vec{r}_A$ is the position of the center of mass of the inner binary system, we have that:

$$\begin{aligned}
\vec{r}_{A_aA} &= \vec{r}_A - \vec{r}_{A_a} \\
&= \frac{m_{A_a}\vec{r}_{A_a} + m_{A_b}\vec{r}_{A_b}}{m_{A_a} + m_{A_b}} - \vec{r}_{A_a} \\
&= (\vec{r}_{A_b} - \vec{r}_{A_a})\frac{m_{A_b}}{m_{A_a} + m_{A_b}} \\
&= \vec{r}_{A_aA_b}\left(\frac{q_{A_aA_b}}{1 + q_{A_aA_b}}\right),
\end{aligned} \tag{2.24}$$

with $q_{A_aA_b} = m_{A_b}/m_{A_a}$ denoting the mass ratio of the inner system. By replacing (2.24) into (2.23), the position of the outer object relative to the system principal component becomes:

$$\vec{r}_{A_aB} = \vec{r}_{AB} + \vec{r}_{A_aA_b}\left(\frac{q_{A_aA_b}}{1 + q_{A_aA_b}}\right), \tag{2.25}$$

and therefore, the observations of the outer component relative to the principal component of the inner system is characterized by the set of orbital parameters $\{P_{A_aA_b}, T_{A_aA_b}, e_{A_aA_b}, a_{A_aA_b}, \omega_{A_aA_b}, \Omega_{A_aA_b}, i_{A_aA_b}\} \cup \{q_{A_aA_b}\} \cup \{P_{AB}, T_{AB}, e_{AB}, a_{AB}, \omega_{AB}, \Omega_{AB}, i_{AB}\}$. It is important to note that while the positional vectors $\vec{r}_{A_aA_b}$ and $\vec{r}_{AB}$ follow a Keplerian orbit (since they are the solution of the 2-body problem of the binary systems A_aA_b and AB, respectively), the positional vector $\vec{r}_{A_aB}$ does not,

being a non-closed orbit. A relevant derived orbital parameter that characterizes the interaction between the inner and outer system, denoted as the mutual inclination Φ, is defined as follows:

$$\cos(\Phi) = \cos(i_{A_aA_b}) \cdot \cos(i_{AB}) + \sin(i_{A_aA_b}) \cdot \sin(i_{AB}) \cdot \cos(\Omega_{AB} - \Omega_{A_aA_b}). \tag{2.26}$$

Now considering the alternative hierarchical structure: let A be the principal component and B_a, B_b be the components of the outer binary system. The motion of the center of mass of the outer system B is described by a Keplerian orbit around the inner component A, forming the inner binary system AB.

Then, an analogous procedure can be developed to express the outer binary system observations in terms of positional vectors that are the solution of the 2-body problem characterized by its respective orbital parameters. Therefore, the positional vector that describes the outer system observations $\vec{r}_{AB_a}$ can be expressed as:

$$\begin{aligned} \vec{r}_{AB_a} &= \vec{r}_{AB} + \vec{r}_{BB_a} \\ &= \vec{r}_{AB} - \vec{r}_{B_aB} \\ &= \vec{r}_{AB} - \vec{r}_{B_aB_b} \left(\frac{q_{B_aB_b}}{1 + q_{B_aB_b}} \right), \end{aligned} \tag{2.27}$$

where the last equality follows the result shown in (2.24).

2.2.1.1.2 Unresolved Inner Binary

If the inner binary A_aA_b is unresolved, the individual components A_a and A_b are not identifiable. Thus, the observations of the outer object B are relative to the photocenter of the inner binary system A^*.

The photocenter of the inner binary system can be expressed analogously to the center of mass by:

$$\vec{r}_{A^*} = \frac{L_{A_a}\vec{r}_{A_a} + L_{A_b}\vec{r}_{A_b}}{L_{A_a} + L_{A_b}}, \tag{2.28}$$

with L_{A_a} and L_{A_b} being the luminosity of the components A_a and A_b, respectively. Repeating the procedure of developed in Section 2.2.1.1.1, the vector $\vec{r}_{A^*B}$ can be expressed as:

$$\vec{r}_{A^*B} = \vec{r}_{A^*A} + \vec{r}_{AB}, \tag{2.29}$$

while the vector $\vec{r}_{A^*A}$ can be decomposed as:

$$\begin{aligned} \vec{r}_{A^*A} &= \vec{r}_A - \vec{r}_{A^*} \\ &= \frac{m_{A_a}\vec{r}_{A_a} + m_{A_b}\vec{r}_{A_b}}{m_{A_a} + m_{A_b}} - \frac{L_{A_a}\vec{r}_{A_a} + L_{A_b}\vec{r}_{A_b}}{L_{A_a} + L_{A_b}} \\ &= (\vec{r}_{A_b} - \vec{r}_{A_a}) \frac{m_{A_b}L_{A_a} - m_{A_a}L_{A_b}}{(m_{A_a} + m_{A_b})(L_{A_a} + L_{A_b})} \\ &= \vec{r}_{A_aA_b} \frac{q_{A_aA_b}L^{-1}_{A_aA_b} - 1}{(1 + q_{A_aA_b})(L^{-1}_{A_aA_b} + 1)} \\ &= \vec{r}_{A_aA_b} \frac{q_{A_aA_b} - L_{A_aA_b}}{(1 + q_{A_aA_b})(1 + L_{A_aA_b})}, \end{aligned} \tag{2.30}$$

with $q_{A_aA_b} = m_{A_b}/m_{A_b}$ and $L_{A_aA_b} = L_{A_b}/L_{A_a}$ being the mass and luminosity ratios of the inner binary system, respectively. By replacing (2.30) into (2.29), the position of the outer object relative to the inner system photocenter becomes:

$$\vec{r}_{A^*B} = \vec{r}_{AB} + \vec{r}_{A_aA_b}\frac{q_{A_aA_b} - L_{A_aA_b}}{(1+q_{A_aA_b})(1+L_{A_aA_b})}. \tag{2.31}$$

The solution of equations (2.21) and (2.22) are Keplerian elliptical orbits which position at a certain time t can be determined following the ephemerides formulae described in equations (2.6) for polar coordinates and (2.7) for rectangular coordinates. Therefore, the observations of the outer component relative to the inner system photocenter are characterized by the set of orbital parameters $\{P_{A_aA_b}, T_{A_aA_b}, e_{A_aA_b}, a_{A_aA_b}, \omega_{A_aA_b}, \Omega_{A_aA_b}, i_{A_aA_b}\} \cup \{q_{A_aA_b}\} \cup \{L_{A_aA_b}\} \cup \{P_{AB}, T_{AB}, e_{AB}, a_{AB}, \omega_{AB}, \Omega_{AB}, i_{AB}\}$.

2.2.1.2 Spectroscopic Hierarchical System

Considering a hierarchical triple system structure composed of an inner binary system A_aA_b and an outer binary system AB, as described in Section 2.2, the radial velocities of the inner system A_aA_b relative to its center of mass A can be expressed as:

$$\begin{aligned} V_{A_a}(t) &= V_A(t) + K_{A_a}[e_{A_aA_b}\cos\omega_{A_aA_b} + \cos(\nu_{A_aA_b}(t) + \omega_{A_aA_b})], \\ V_{A_b}(t) &= V_A(t) - K_{A_b}[e_{A_aA_b}\cos\omega_{A_aA_b} + \cos(\nu_{A_aA_b}(t) + \omega_{A_aA_b})]. \end{aligned} \tag{2.32}$$

In a similar manner, the radial velocities of the inner system AB relative to its center of mass can be expressed as:

$$\begin{aligned} V_A(t) &= V_0 + K_A[e_{AB}\cos\omega_{AB} + \cos(\nu_{AB}(t) + \omega_{AB})], \\ V_B(t) &= V_0 - K_B[e_{AB}\cos\omega_{AB} + \cos(\nu_{AB}(t) + \omega_{AB})]. \end{aligned} \tag{2.33}$$

The center of mass of the outer system AB is equal to the center of mass of the whole system, and therefore, the corresponding radial velocity is denoted as V_0 in (2.33). This is considered as an orbital parameter (constant) that characterizes the spectroscopic hierarchical system. Finally, the radial velocities of the inner system relative to its center of mass A becomes:

$$\begin{aligned} V_{A_a}(t) &= V_0 + K_A[e_{AB}\cos\omega_{AB} + \cos(\nu_{AB}(t) + \omega_{AB})] + K_{A_a}[e_{A_aA_b}\cos\omega_{A_aA_b} + \cos(\nu_{A_aA_b}(t) + \omega_{A_aA_b})], \\ V_{A_b}(t) &= V_0 + K_A[e_{AB}\cos\omega_{AB} + \cos(\nu_{AB}(t) + \omega_{AB})] - K_{A_b}[e_{A_aA_b}\cos\omega_{A_aA_b} + \cos(\nu_{A_aA_b}(t) + \omega_{A_aA_b})]. \end{aligned} \tag{2.34}$$

Although a closed expression for the determination of V_A is provided in (2.33), this component is fictional because it corresponds to the center of mass of the real components of the inner system A_a and A_b, and therefore, no radial velocity observations of the component A are available.

Now considering the alternative hierarchical structure, i.e., an inner binary system B_aB_b that orbits a primary object A, the radial velocities of the inner system B_aB_b relative to its center of mass B can be expressed as:

$$\begin{aligned} V_{B_a}(t) &= V_B(t) + K_{B_a}[e_{B_aB_b}\cos\omega_{B_aB_b} + \cos(\nu_{B_aB_b}(t) + \omega_{B_aB_b})], \\ V_{B_b}(t) &= V_B(t) - K_{B_b}[e_{B_aB_b}\cos\omega_{B_aB_b} + \cos(\nu_{B_aB_b}(t) + \omega_{B_aB_b})], \end{aligned} \tag{2.35}$$

where $V_B(t)$ is defined identically as in (2.33).

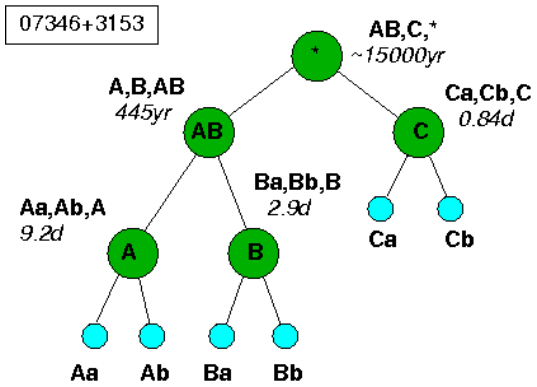

Figure 2.4: Mobile diagram of a 6-hierarchical stellar system. In this case WDS07346+3153 (=), from Tokovinin´s "Multiple Star Catalogue" (`http://www.ctio.noao.edu/~atokovin/stars/`).

2.2.2 Higher Multiplicities

The number of possible configurations of an hierarchical stellar systems increases with its multiplicity. To avoid confusions with the specific configuration of an hierarchical system, the architecture is specified in a mobile diagram.

The mobile diagram is a binary tree structure where each leaf (in cyan) represents a star of the hierarchical system, and the nodes (in green) represents the center of mass of the subsequent leafs, where the root of the tree represents the center of mass of the complete system. The dynamics of two child nodes rooted from the same parent node are described by the solution of the 2-body problem, i.e., composes a binary system. The most massive component of a binary system is always a left child node and vice-versa. An example of a mobile diagram that illustrates the configuration of a hierarchical stellar system is presented in Figure 2.4.

Chapter 3

Bayesian Inference

3.1 Bayesian Statistics

Bayesian statistics is a theory that interprets probability as a degree of belief of an event based on prior knowledge or personal belief of that event and treats all the modeling phenomena as random variables.

Let $x \in \mathcal{X}$ the observations of a phenomenon to be modeled and $\theta \in \Theta$ the parameters that characterize that model. The *Bayesian inference* is the procedure to characterize the probability distribution of a model conditioned to the observations, i.e., $p(\theta|x)$. According to the Bayes' theorem, the probability $p(\theta|x)$ can be expressed as:

$$p(\theta|x) = \frac{p(x|\theta)p(\theta)}{p(x)} \propto p(x|\theta)p(\theta). \tag{3.1}$$

In the above expression we have that:

- $p(\theta)$ is the prior knowledge or prior belief of the model, denoted as *prior distribution*.
- $p(\theta|x)$ is the probability of the model after an observation is provided, denoted as *posterior distribution*.
- $p(x|\theta)$ is the probability law of the model (θ) fits the observations, denoted as the *likelihood* or *goodness of fit* of the model.
- $p(x) = \int_\Theta p(x|\theta)p(\theta)d\theta$ is the probability of the observations, denoted as the *marginal likelihood* or *evidence*

The marginal likelihood $p(x)$ does not depends on a specific election of θ, and therefore, it can be treated as a normalization constant of the likelihood times prior, and hence, $p(\theta|x) \propto p(x|\theta)p(x)$.

The posterior distribution can be used to perform a *Bayesian prediction* of new samples $\tilde{x} \in \mathcal{X}$ through the determination of the probability distribution $p(\tilde{x}|x)$ as follows:

$$p(\tilde{x}|x) = \int_\Theta p(\tilde{x}|\theta)p(\theta|x)d\theta, \tag{3.2}$$

which is denoted as the *posterior predictive distribution*.

One of the main characteristics of the Bayesian modeling is that it allows incorporating all the sources of uncertainty: the *epistemic uncertainty*, which represents the inability to determine the true model that explains the observations and the *aleatoric uncertainty*, which represents the inherent variability of the model. The *epistemic uncertainty* is encapsulated by the prior $p(x)$, and it can be reduced through the addition of observational data of the modeled phenomena, while the *aleatoric uncertainty* is encapsulated by the likelihood $p(x|\theta)$ and is irreducible.

The determination of the posterior distribution $p(\theta|x)$ involves the computation of the prior distribution $p(\theta)$, which is given by expert knowledge, the likelihood $p(x|\theta)$, which is given by the observation evaluated in the model, and the evidence $p(x)$, which involves to compute the intractable integral $p(x) = \int_\Theta p(x|\theta)p(\theta)d\theta$, being the mayor difficulty of the Bayesian inference procedure.

There are two main approaches to perform the Bayesian inference: the *Variational Inference* methods, which avoids the intractable computation of the marginal likelihood by a proposal tractable distribution, minimizing its discrepancy concerning the intractable target distribution through an optimization procedure, and the *Monte Carlo Sampling* methods, which performs random realizations of the tractable part of the posterior target distribution (likelihood times prior) normalizing them. In general, the Monte Carlo sampling methods are more computationally expensive than the variational inference methods. However, they provide guarantees of asymptotic convergence to the target true distribution, while the variational inference methods do not.

Due to its theoretical guarantees, the present work focuses on Monte Carlo sampling methods for Bayesian inference, particularly in the Markov Chain Monte Carlo methods.

3.2 Markov Chain Monte Carlo

The Markov Chain Monte Carlo (MCMC) algorithms are sampling methods (based on the construction of a Markov chain) that offer an empirical distribution that converges to a target distribution. The samples are obtained by recording the chain states.

3.2.1 Markov Chain

To understand the theory beneath MCMC algorithms, some basic definitions and theorems on Markov Chains [Ross et al., 1996, Norris, 1998] must be introduced.

Definition 3.1 (Markov Chain) *A stochastic process $X = (X_n)_{n\geq 0}$ that takes values in a discrete space state I is defined as a Markov chain if it satisfies the Markov property:*

$$p(X_{n+1} = i_{n+1}|X_0 = i_0, X_1 = i_1, ..., X_n = i_n) = p(X_{n+1} = i_{n+1}|X_n = i_n), \ \ \forall i_0, ..., i_n \in I. \quad (3.3)$$

The definition of a Markov chain implies that the stochastic process is *memoryless*, i.e., given the present, the future does not depend on the past.

A Markov chain is characterized by an *initial distribution* μ and a *transition matrix* T.

Definition 3.2 (Initial Distribution) *The initial distribution of a Markov chain X is the probability vector μ over the space state I defined by $\mu_i = p(X_0 = i)$.*

Definition 3.3 (Homogeneity and Transition Matrix) *A Markov chain X is homogeneous if:*

$$p(X_{m+1} = j | X_m = i) = p(X_{n+1} = j | X_n = i), \quad \forall m, n \geq 0, i, j \in I. \tag{3.4}$$

In this case, the matrix $T = (T_{ij})_{i,j \in I}$, with $T_{ij} = p(X_1 = j | X_0 = i)$ is defined as the transition matrix of X.

The homogeneity implies that the transition probabilities between the states are invariant to any $n \geq 0$. It is easy to prove that the transitions matrix T is stochastic, i.e., $\forall i \in I$, $T_{ij} \geq 0$ and $\sum_{j \in I} T_{ji} = 1$.

Some important properties of the Markov chains are the notion of *irreducibility*, *aperiodicity* and *recurrence*.

Definition 3.4 (Irreducibility) *A Markov chain X is irreducible if:*

$$\forall i, j \in I, \exists n \geq 0 : p(X_n = j | X_0 = i) > 0. \tag{3.5}$$

Definition 3.5 (Aperiodicity) *A Markov chain X is aperiodic if:*

$$\forall i \in I, \exists n^* \geq 0, \forall n \geq n^* : p(X_n = i | X_0 = i) > 0. \tag{3.6}$$

Definition 3.6 (Recurrence) *A state $i \in I$ of a Markov chain is recurrent if:*

$$p(X_0 = i, \tau_i < \infty) = 1, \tag{3.7}$$

where τ_i is the return time of the state i defined as:

$$\tau_i \triangleq \inf\{n \geq 1 : X_n = i\}. \tag{3.8}$$

The *irreducibility* of a Markov chain implies that every state of the chain is connected, and the *aperiodicity* implies that for any state of the chain, there is a positive probability of returning to the same state in an arbitrary number of steps. The recurrence of a state i implies that the chain returns to the same state with probability one. A recurrent state can be classified into a positive recurrent or a null recurrent state.

Definition 3.7 (Positive/Null Recurrence) *A recurrent state $i \in I$ is said positive recurrent if the expected return time is finite, i.e., $\mathbb{E}_i(\tau_i) < \infty$. Otherwise, it is said to be null recurrent.*

A fundamental requirement of MCMC algorithms that allows sampling from a target distribution is that the corresponding Markov chain converges to a unique *stationary distribution*.

Definition 3.8 (Stationary Measure/Distribution) *The probability vector μ is a stationary measure over a transition matrix T if:*

$$\sum_i \mu_i T_{ij} = \mu_j, \ \ \forall j \in I. \tag{3.9}$$

If the stationary measure μ additionally satisfies that:

$$\sum_j \mu_i = 1, \ \ \forall j \in I, \tag{3.10}$$

then is denoted as a stationary distribution π.

Theorem 3.9 (Existence of Stationary Distribution) *Let X be an irreducible Markov chain. The following statements are equivalent:*

1. *A recurrent positive state exists.*
2. *All the states are recurrent positives.*
3. *A stationary distribution π exists. Moreover, the stationary distribution is determined by:*

$$\pi(i) = \frac{1}{\mathbb{E}_i(\tau_i)} \tag{3.11}$$

Theorem 3.10 (Unicity of Stationary Distribution) *Let X be an irreducible Markov chain. If X has a stationary distribution, all the stationary measures are multiples of it; therefore, the stationary distribution is unique.*

Theorems (3.9) and (3.10) ensure the existence and unicity of a stationary distribution π of a irreducible Markov chain with a recurrent positive state. However, a result that ensures the convergence of the Markov chain to that stationary distribution is required.

Theorem 3.11 (Ergodicity of Markov Chains) *Let X be an irreducible and aperiodic Markov chain with a stationary distribution π. Therefore, for any initial distribution μ, the chain satisfies that:*

$$\lim_{n\to\infty} p_\mu(X_n = i) = \pi(i), \ \ \forall i \in I. \tag{3.12}$$

Moreover, if the states space I is finite, $\exists c > 0, \exists \alpha \in (0,1)$ that:

$$\sup_{i,j\in I} |p(X_n = j | X_0 = i) - \pi(j)| \leq c\alpha^n. \tag{3.13}$$

Therefore, the irreducibility, aperiodicity, and the existence of a recurrent positive state ensure that a Markov chain converges to a unique stationary distribution.

Although Theorem (3.9) provides an expression for the stationary distribution π, the involved calculus is generally a non-trivial task. The following condition can help to find the stationary distribution of a Markov chain.

Theorem 3.12 (Detailed Balance) *A measure λ and a stochastic matrix $T = (p_{ij})_{i,j\in I}$ satisfies the detailed balance condition if:*

$$\lambda_i T_{ij} = \lambda_j T_{ji}. \tag{3.14}$$

It is easy to prove that if λ satisfies the detailed balance condition for T, then λ is a stationary measure of T, and using the Theorem (3.10), the stationary distribution π can be founded (if it exists). The detailed balance is a sufficient but not a necessary condition for stationary measures.

All the definitions and theorems presented in the discrete state space I directly extend for the continuous state space case, where the transition matrix T becomes an integral kernel K.

3.2.2 MCMC Algorithms

Although there is a wide variety of Markov chain Monte Carlo methods, the current section covers the four most popular algorithms.

3.2.2.1 Metropolis-Hastings

The Metropolis-Hastings algorithm [Metropolis et al., 1953, Hastings, 1970] aims to sample from a target distribution P by designing a Markov chain that converges to an unique stationary distribution equals to P.

The design of the Markov chain involves defining the transition probabilities that fulfill the desired conditions. Let X be a homogeneous Markov chain that takes values in a space state I characterized by a transition matrix $T = (T_{ij})_{i,j\in I}$ and Q an arbitrary proposal distribution. The transition probabilities of X are defined as follows:

$$T_{ij} = Q(j|i)\mathcal{A}(i,j) + \mathbf{1}_{i=j} \underbrace{\sum_{k\neq i} Q(i|k)\cdot(1-\mathcal{A}(i,k))}_{r(i)}, \tag{3.15}$$

where $\mathcal{A}(i,j) \triangleq \min\left\{1, \frac{P(j)Q(i|j)}{P(i)Q(j|i)}\right\}$ is the acceptance probability of transitioning from i to j and $r(i)$ is the probability of rejecting the transition. If Q is symmetric, i.e., $Q(j|i) = Q(i|j)$, the acceptance probability reduces to $\mathcal{A}(i,j) = \min\{1, P(j)/P(i)\}$. Note that the transition probabilities T_{ij} requires to compute the probability ratio $P(j)/P(i)$, and hence, any distribution equals to the target distribution up to a constant is feasible.

It is clear that the Markov chain X satisfies the Markov property, is irreducible if $supp(P) \subseteq supp(Q)$ and is aperiodic since T_{ij} always allows rejection. Moreover, the designed Markov chain satisfies the detailed balance condition (Theorem 3.12) for $\lambda = P$:

$$\begin{aligned} P(i)T_{ij} &= P(i)Q(j|i)\min\left\{1, \frac{P(j)Q(i|j)}{P(i)Q(j|i)}\right\} \\ &= \min\{P(i)Q(j|i), P(j)Q(i|j)\} \\ &= P(j)Q(i|j)\min\left\{\frac{P(i)Q(j|i)}{P(j)Q(i|j)}, 1\right\} \\ &= P(j)T_{ji}. \end{aligned} \tag{3.16}$$

Therefore, P is a stationary measure of X, and since X is irreducible, p is the unique stationary distribution of X. Finally, the Theorem 3.11 allows to conclude that X converges to the stationary distribution P. The complete Metropolis-Hastings sampling procedure is described in Algorithm 1.

Algorithm 1: Metropolis-Hastings.

Parameters: N
Initialize $x^{(0)}$
for $0 \leq i < N$ **do**
  Sample $u \sim \mathcal{U}[0,1]$
  Sample $x^* \sim q(x^*|x^{(i)})$
  if $u < \mathcal{A}(x^{(i)}, x^*) = \min\left\{\frac{p(x^*)q(x^{(i)}|x^*)}{p(x^{(i)})q(x^*|x^{(i)})}\right\}$ **then**
    $x^{(i+1)} = x^*$
  else
    $x^{(i+1)} = x^{(i)}$

3.2.2.2 Gibbs Sampler

Since the Metropolis-Hastings allows sampling from a target distribution P by evaluating a version of it up to a constant, this is generally difficult to compute in a multivariate scenario. However, the evaluation of the joint distribution up to a constant can be avoided if its marginal distributions are known.

The Gibbs sampler [Geman and Geman, 1984] aims to obtain samples from a multivariate target distribution $P(x_1, ..., x_d)$ by sampling from its marginal distributions $P(x_j|x_1, ..., x_{j-1}, x_{j+1}, ..., x_d)$. Let X an homogeneous Markov chain that takes values in a state space $I \in \mathbb{R}^d$ characterized by a transition matrix $T = (T_{\mathbf{xy}})_{\mathbf{x},\mathbf{y}\in I}$ and $\mathbf{x} \sim_j \mathbf{y}$ an equivalence relation which states that $\mathbf{x}_i = \mathbf{y}_i, \forall i \neq j$. The transition probabilities of X are defined as follows:

$$T_{\mathbf{xy}} = \mathbf{1}_{\mathbf{x}\sim_j\mathbf{y}} \cdot \frac{1}{d} \cdot \frac{P(\mathbf{y})}{\sum_{\mathbf{z}\sim_j\mathbf{x}} P(\mathbf{z})}. \tag{3.17}$$

By transitivity of the equivalence relation $\mathbf{x} \sim_j \mathbf{y}$, the designed Markov chain satisfies the detailed balance condition for $\lambda = P$:

$$\begin{aligned} P(\mathbf{x})T_{\mathbf{xy}} &= \mathbf{1}_{\mathbf{x}\sim_j\mathbf{y}} \cdot \frac{1}{d} \cdot \frac{P(\mathbf{x})P(\mathbf{y})}{\sum_{\mathbf{z}\sim_j\mathbf{x}} P(\mathbf{z})} \\ &= \mathbf{1}_{\mathbf{y}\sim_j\mathbf{x}} \cdot \frac{1}{d} \cdot \frac{P(\mathbf{y})P(\mathbf{x})}{\sum_{\mathbf{z}\sim_j\mathbf{y}} P(\mathbf{z})} \\ &= P(\mathbf{y})T_{\mathbf{yx}}. \end{aligned} \tag{3.18}$$

The Markov chain X is also irreducible and aperiodic, and therefore, it converges to the stationary distribution P.

By considering that the marginal distribution of $P(\mathbf{x})$ can be expressed as:

$$P(x_j|x_1, ..., x_{j-1}, x_{j+1}, ..., x_d) = \frac{P(x_1, ..., x_d)}{P(x_1, ..., x_{j-1}, x_{j+1}, ..., x_d)}, \tag{3.19}$$

the definition of the transition probabilities $T_{\mathbf{xy}}$ implies that the samples of $P(\mathbf{x})$ can be obtained by sampling from the marginal distributions $P(x_j|x_1, ..., x_{j-1}, x_{j+1}, ..., x_d)$, picking $j \in \{1, ..., d\}$ at random. The Gibbs sampler can be viewed as an special of the Metropolis-Hastings algorithm with a proposal distribution $Q(\mathbf{y}|\mathbf{x}) = \mathbf{1}_{\mathbf{x}\sim_j\mathbf{y}} P(x_j|x_1, ..., x_{j-1}, x_{j+1}, ..., x_d)$. In that case, the acceptance probability is always $\mathcal{A}(\mathbf{x}, \mathbf{y}) = 1$. The complete Gibbs sampling procedure is described in Algorithm 2.

Algorithm 2: Gibbs sampler.

Parameters: N
Initialize $x^{(0)}$
for $0 \leq i < N$ **do**

$$
\begin{aligned}
\mathrm{x}_1^{(i+1)} &\sim p(x_1|x_2^{(i)}, x_3^{(i)}, ..., x_d^{(i)}) \\
\mathrm{x}_2^{(i+1)} &\sim p(x_2|x_1^{(i+1)}, x_3^{(i)}, ..., x_d^{(i)}) \\
&\vdots \\
\mathrm{x}_j^{(i+1)} &\sim p(x_j|x_1^{(i+1)}, ..., x_{j-1}^{(i+1)}, x_{j+1}^{(i)}, ..., x_d^{(i)}) \\
&\vdots \\
\mathrm{x}_d^{(i+1)} &\sim p(x_d|x_1^{(i+1)}, x_2^{(i+1)}, ..., x_{d-1}^{(i+1)})
\end{aligned}
$$

3.2.2.3 Hamiltonian Monte Carlo

The Hamiltonian Monte Carlo [Neal et al., 2011], also known as Hybrid Monte Carlo, is an instance of the Metropolis-Hastings algorithm that makes use of the geometry of the target probability distribution to guide the transitions of the Markov chain. This allows performing the sampling efficiently, avoiding random walk behaviors and sensitivity to correlated parameters. These features allow convergence to high-dimensional target distributions much more quickly than simpler methods like the random walk Metropolis-Hastings or the Gibbs sampler.

The core of the Hamiltonian Monte Carlo method is to sample from the zone of the parameter space that highly contributes to the computation of the exception of a target distribution $P(x)$ given a parametrization (denoted by f) of the parameter space $\mathcal{X}$:

$$\mathbb{E}_P(f) = \int_{\mathcal{X}} P(x)f(x)dx, \tag{3.20}$$

denoted as the *typical set*.

For that purpose, the transitions of the Markov chain must be guided by a vector field in the direction of the typical set by exploiting the differential structure of the target distribution. Hence, the vector field is generated by using the gradient of the target distribution and auxiliary momentum parameters that compensate the attractive force of the gradient to the target distribution mode, preserving a dynamical equilibrium that allows aligning the generated vector field with the typical set.

A conservative dynamic in physical systems requires that any compression or expansion in the position space be compensated with a respective expansion or compression in the momentum space, preserving the volume in the joint space of position and momentum. To ensure the conservative dynamic behavior, the transition probabilities of the chain follow the Hamiltonian dynamics. The

Hamiltonian dynamical system is described by a function over the position x and momentum p variables, known as the Hamiltonian function $H(x, p)$.

Let $x_n \in \mathbb{R}^d$ the vector of parameters of the space state and $P(x)$ the target distribution, each dimension of the space state is complemented by a fictitious momentum variable x:

$$x_n \to (x_n, p_n), \tag{3.21}$$

where $p_n \in \mathbb{R}^d$. The combined space of the parameters $(x_n, p_n) \in \mathbb{R}^{2d}$ is denoted as the phase space and the respective induced distribution $P(x, p)$ is denoted as the canonical distribution.

To mimic the conservative dynamic behavior of the space variables and the momentum variables, the canonical distribution is written in terms of the Hamiltonian function[1]:

$$P(x, p) = e^{-H(x,p)}, \tag{3.22}$$

which implies that:

$$H(x, p) = -\log P(x, p), \tag{3.23}$$

Hence, the Hamiltonian function captures the probabilistic structure of the phase space, and consequently, the geometry of its typical set.

Marginalization of the canonical distribution $P(x, p)$ in terms of the state variable p induces the following decomposition of the Hamiltonian function $H(x, p)$:

$$\begin{aligned} H(x, p) &= -\log P(p|x) - \log P(x) \\ &\equiv K(p, x) + V(x). \end{aligned} \tag{3.24}$$

The decomposition can be interpreted as a kinetic energy $K(p, x)$ function, dependent on both space and momentum variables, and potential energy function $V(x)$, dependent on the momentum variables only. The potential function is simply the negative logarithm of the target distribution, while kinetic energy is usually expressed as a quadratic term on p:

$$K(p) = \frac{1}{2} p^T \cdot M^{-1} \cdot p, \tag{3.25}$$

where M is a symmetric, positive-definite matrix denoted as mass-matrix. The mass matrix is typically a scalar multiple of the identity matrix but can explicitly depends on x as in (3.24).

Therefore, the vector field oriented in the direction of the typical set can be defined through the Hamiltonian equations:

$$\begin{aligned} \frac{dx}{dt} &= +\frac{\partial H}{\partial p} = \frac{\partial K}{\partial p} \\ \frac{dp}{dt} &= -\frac{\partial H}{\partial x} = -\frac{\partial K}{\partial x} - \frac{\partial V}{\partial x}. \end{aligned} \tag{3.26}$$

Following the vector field (determined by the Hamiltonian equations for a time t) generate trajectories $\phi_t(x, p)$ that moves along the typical set. To compute these trajectories, the solution

[1]It follows the Boltzmann canonical distribution $P(x) = z^{-1} e^{-E(x)/t}$, with z a normalization constant and t the temperature variable fixed to one.

of (3.26) is obtained by numerical methods. In particular, the trajectory $\phi_T(x,p)$ of a variable i can be approximated by the *leap-frog integration* method iterating the following expressions:

$$\begin{aligned} p_i(t+\varepsilon/2) &= p_i(t) - \frac{\varepsilon}{2}\frac{\partial V}{\partial x_i}(x_i(t)) \\ x_i(t+\varepsilon) &= x_i(t) + \frac{\varepsilon}{m_i} p_i(t+\varepsilon/2) \\ p_i(t+\varepsilon) &= p_i(t+\varepsilon/2) - \frac{\varepsilon}{2}\frac{\partial V}{\partial x_i}(x_i(t+\varepsilon)), \quad \forall i \in 1,...,L. \end{aligned} \tag{3.27}$$

L is the number of steps, $\varepsilon \in \mathbb{R}$ is the step size, and $T = \lfloor L/\varepsilon \rfloor$ the integration time. The adequate selection of the algorithm hyper-parameter ε and L is crucial for a good sampling performance.

In summary, the Markov chain that samples from the target distribution $P(x)$ will follow the Metropolis-Hastings algorithm defined on the phase space (x,p) with transition probabilities $T_{(x_0,p_0),(x_L,p_L)}$ determined by the solution of the Hamiltonian equations. Its computation follows the leap-frog integration method for a fixed number of steps L and a step size ε. The momentum variables are sampled from a proposal marginal distribution $P(p|x)$, and the final samples of $P(x)$ are obtained by projecting the samples of the phase space on the state space $(x,p) \to x$, i.e., ignoring the momentum variables.

The transition probabilities of the current Markov Chain must be modified since the transition ratio in the Metropolis-Hastings acceptance probability $T_{(x_L,p_L),(x_0,p_0)}/T_{(x_0,p_0),(x_L,p_L)} = 0/1 = 0$, because the leap-frog integration does not allow reverse trajectories. Thus, the transition probabilities are modified to be reversible by augmenting the numerical integration with a negation step that flips the sign of momentum $(x,p) \to (x,-p)$. Thereby, the Metropolis-Hastings acceptance rate is:

$$\begin{aligned} \mathcal{A}((x_L,-p_L),(x_0,p_p)) &= \min\left\{1, \frac{T_{(x_L,-p_L),(x_0,p_0)} P(x_L,-p_L)}{T_{(x_0,p_0),(x_L,-p_L)} P(x_0,p_0)}\right\} \\ &= \min\left\{1, \frac{\delta(x_L - x_L)\delta(-p_l + p_L) P(x_L,-p_L)}{\delta(x_0 - x_0)\delta(p_0 - p_0) P(x_0,p_0)}\right\} \\ &= \min\left\{1, \frac{P(x_L,-p_L)}{P(x_0,p_0)}\right\} \\ &= \min\left\{1, e^{-H(x_L,-p_L)+H(x_0,p_0)}\right\}. \end{aligned} \tag{3.28}$$

The complete Hamiltonian Monte Carlo sampling procedure is described in Algorithm 3.

Algorithm 3: Hamiltonian Monte Carlo.

Parameters: N, ε, T, M
Initialise $x^{(0)}$
for $0 \leq i < N$ **do**
  Sample $u \sim \mathcal{U}([0,1]), p \sim \mathcal{N}(0, M)$
  $x_0 \leftarrow x^{(i)}, p_0 \leftarrow p, L \leftarrow \lfloor \frac{T}{\varepsilon} \rfloor$
  for $0 \leq l < L$ **do**
    $p_{l+\frac{1}{2}} \leftarrow p_l - \frac{\varepsilon}{2}\frac{\partial V}{\partial x}(x_l)$
    $x_{l+1} \leftarrow x_l + \varepsilon p_{l+\frac{1}{2}}$
    $p_{l+1} \leftarrow p_{l+\frac{1}{2}} - \frac{\varepsilon}{2}\frac{\partial V}{\partial x}(x_l)$
  if $u < \min\left(1, e^{-H(x_L, -p_L)+H(x_0,p_0)}\right)$ **then**
    $x^{(i+1)} \leftarrow x_L$
  else
    $x^{(i+1)} \leftarrow x^{(i)}$

3.2.2.4 No-U-Turn Sampler

The No-U-Turn Sampler algorithm [Hoffman et al., 2014] is an extension of the Hamiltonian Monte Carlo algorithm that adaptively sets the number of steps L of the trajectories. This adaptive selection facilitates the use of the sampling tool by avoiding a low-performance selection of the Hamiltonian Monte Carlo user-defined hyper-parameters. The selection of the number of steps L follows the computation of the Hamiltonian Monte Carlo algorithm's forward and backward exploration trajectories until an end condition is met. A new sample is obtained by a random selection of the generated trajectories.

To generate the exploration trajectories, a binary tree is constructed iteratively. Let $(x_n(0), p_n(0))$ be an initial particle composed by the a position and a momentum of the n-th iteration of the Markov chain, (x_n^+, p_n^+) be a forward in time particle and (x_n^-, p_n^-) be a backward in time particle. In each iteration j, the binary tree selects, in a uniform random way, to move the $(j-1)$-particle forward or backward in time, with 2^j leap-frog integration steps. Figure 3.1 illustrates the No-U-turn sampler trajectories construction, where an initial particle (in black) is moved forward-backward-backward-forward in time, respectively.

The iterative procedure continues until the following condition (namely the U-Turn condition) is met:

$$(x_n^+ - x_n^-) \cdot p_n^- < 0 \ \vee \ (x_n^+ - x_n^-) \cdot p_n^+ < 0, \tag{3.29}$$

or when the Hamiltonian trajectory generated by the leap-frog integration becomes imprecise in the sense that:

$$e^{-H(x_n^+, p_n^+)+\Delta_{max}} < U_n \ \vee \ e^{-H(x_n^-, p_n^-)+\Delta_{max}} < U_n, \tag{3.30}$$

where $U_n \sim \mathcal{U}\left(0, e^{-H(q_n(0), p_n(0))}\right)$ is a slice random variable sample and Δ_{max} is a maximum energy hyper-parameter. The idea behind the No-U-Turn condition is to avoid the generation of redundant trajectories by stopping the exploration when the trajectory begins to turn back to previously explored zones.

Finally, the new sample (x_{n+1}, p_{n+1}) is selected by an uniform sampling of the generated trajec-

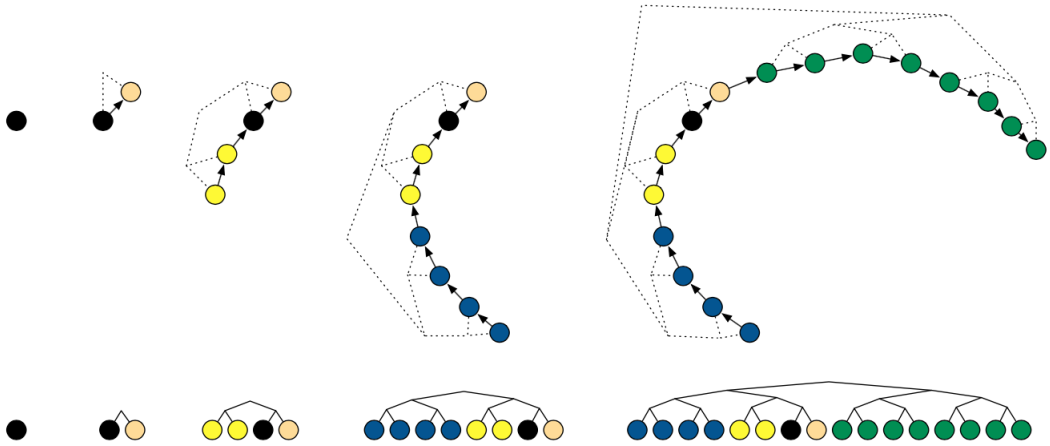

Figure 3.1: Illustration of No-U-Turn sampler trajectories construction procedure. Reprinted from [Hoffman et al., 2014].

tory that satisfies the precision condition $U_n < e^{-H(x_{n+1},p_{n+1})}$. The complete No-U-Turn sampler sampling procedure is described in Algorithm 4.

Algorithm 4: No-U-Turn-Sampler.

Parameters: $x^0, \varepsilon, \mathcal{L}, M$; // $\mathcal{L} = \log P$

for $m = 1$ **to** M **do**

Sample $p^0 \sim \mathcal{N}(0, I), u \sim \mathcal{U}([0, \exp(\mathcal{L}(x^{m-1} - \frac{1}{2}p^0 \cdot p^0))])$

Initialise $x^- = x^{m-1}, x^+ = x^{m-1}, p^- = p^0, p^+ = p^0, x^m = x^{m-1}$

Initialise $j = 1, n = 1, s = 1$

while *s=1* **do**

Sample $v_j \sim \mathcal{U}(\{-1, 1\})$

if $v_j = -1$ **then**

$x^-, p^-, -, -, x', n', s' \leftarrow$ BUILDTREE$(x^-, p^-, u, v_j, j, \varepsilon)$

else

$-, -, x^+, p^+, x', n', s' \leftarrow$ BUILDTREE$(x^+, p^+, u, v_j, j, \varepsilon)$

if $s' = 1$ **then**

With probability $\min\left(1, \frac{n'}{n}\right)$: $x^m \leftarrow x'$

$n \leftarrow n + n', s \leftarrow s'\mathbb{I}[(x^+ - x^-) \cdot p^- \geq 0]\mathbb{I}[(x^+ - x^-) \cdot p^+ \geq 0], j \leftarrow j + 1$

BUILDTREE$(x, p, u, v, j, \varepsilon)$*:*

if $j = 0$ **then**

$x' \leftarrow x + \varepsilon(p + \frac{1}{2}\varepsilon\nabla_x\mathcal{L}(x))$

$p' \leftarrow p' + \frac{1}{2}\varepsilon\nabla_x\mathcal{L}(x')$

$n' \leftarrow \mathbb{I}[u \leq \exp\{\mathcal{L}(x') - \frac{1}{2}p' \cdot p'\}]$

$s' \leftarrow \mathbb{I}[u < \exp\{\Delta_{max} + \mathcal{L}(x') - \frac{1}{2}p' \cdot p'\}]$

return $x', p', x', p', x', n', s'$

else

$x^-, p^-, x^+, p^+, x', n', s' \leftarrow$ BUILDTREE$(x, p, u, v, j - 1, \varepsilon)$

if $s' = 1$ **then**

if $v = -1$ **then**

$x^-, p^-, -, -, x'', n'', s'' \leftarrow$ BUILDTREE$(x^-, p^-, u, v, j - 1, \varepsilon)$

else

$-, -, x^+, p^+, x'', n'', s'' \leftarrow$ BUILDTREE$(x^+, p^+, u, v, j - 1, \varepsilon)$

With probability $\frac{n''}{n' + n''}$: $x' \leftarrow x''$

$s' \leftarrow s''\mathbb{I}[(x^+ - x^-) \cdot p^- \geq 0]\mathbb{I}[(x^+ - x^-) \cdot p^+ \geq 0]$

$n' \leftarrow n' + n''$

return $x^-, p^-, x^+, p^+, x', n', s'$

Chapter 4

Related Work

This chapter exposes a brief review of some of the work on orbital parameters estimation of binary and hierarchical stellar systems, encompassing both optimization-based and Bayesian-based techniques. This section also mentions recent papers related to the application of Bayesian tools on similar celestial objects.

4.1 Optimization-Based Orbital Parameters Estimation

The problem of estimating the orbital parameters of binary stellar systems is nothing new. The first proposed method that solved the problem of astrometric orbital fitting on visual binaries consists of analytical formulations (from physical models). This requires a set of three complete and highly precise observations of relative position on the apparent orbit, of the form (t, ρ, θ), and a double areal constant obtained from additional data [Thiele, 1883], an additional incomplete observation of the form (t, θ) [Cid Palacios, 1958] or an auxiliary angular variable that maps a set of feasible apparent orbits [Docobo, 1985]. All of these methods require highly precise observations, ignoring the different levels of precision and uncertainties of the measurements.

To obtain robust solutions to the orbital fitting problem considering multiple observations (with different levels of precision), the optimization-based approach was proposed. Some of these optimization-based methods minimize a sum of weighted square errors between the physical model estimates and the observations by using the Levenberg-Marquardt algorithm [Tokovinin, 1992], simulated annealing [Pourbaix, 1994], or the downhill simplex method [MacKnight and Horch, 2004], among others. All the works mentioned above focus on fitting positional observations of the orbit, completely ignoring another important source of information obtained through spectroscopy: the radial velocities of each system component.

The first attempts to solve the orbital parameters using positional and radial velocities observations of its components were made by fitting each source of information separately, using the estimates obtained by fitting one of the sources of observations to fix some orbital parameters. The determination of the remaining orbital parameters was made through the fitting of the other source of observations [Docobo et al., 1992, Hummel et al., 1994]. Unfortunately, the separate fitting of the astrometric and spectroscopic sources of observations yields a sub-optimal determination of the

complete set of orbital parameters. Indeed, there is no guarantee that the solutions obtained by both fittings were consistent. To avoid this issue, [Morbey, 1975] address the problem of determining the orbital parameters of a visual-spectroscopic binary system by fitting each source of observations jointly. They use a maximum likelihood estimation and Lagrange multipliers method. Lately, the methods developed for the fitting of astrometric observations were extended to fit both the astrometric and radial velocities observations simultaneously [Pourbaix, 1998]. Similar approaches were performed to the estimation of the individual masses in single-lined spectroscopic binaries. In this context, a well-determined supplementary observation of the system parallax (using other techniques besides relative astrometry of the binary pair and spectroscopy for radial velocity) are used as a fixed value within the estimation of the orbital parameters [Docobo et al., 2018] or a supplementary parallax observation is used as an additional observation to perform the estimation in a joint manner [Muterspaugh et al., 2010].

The same guidelines were followed for the estimation of the orbital parameter of multiple stellar systems. By using the hierarchical approximation, the system is decomposed into separate interdependent binaries. Hence, the orbital parameters of the whole system can be determined by fitting each binary's astrometric and spectroscopic observations. Some works perform the estimation of the orbital parameters of hierarchical systems by determining the orbital parameters of each binary separately (using the estimates of one binary as fixed parameters for the estimation of the orbital parameters of the other [Docobo et al., 2008, Köhler et al., 2012]). In contrast, other work performs the estimation of the orbital parameter of the whole system jointly. On this last family, [Tokovinin and Latham, 2017] proposed to perform the joint estimation in hierarchical triple system by minimizing the sum of weighted square errors of the observations, alternating between the fitting of the astrometric and spectroscopic observations from the inner and outer binary systems, considering that, often, some observations are more reliable than others (e.g., the radial velocity of the inner system over its astrometry, and vice-versa for the outer pair).

4.2 Bayesian-Based Orbital Parameters Estimation

One of the main drawbacks of the optimization-based methods presented in Section 4.1 is that obtained solution is entirely deterministic. This strategy does not provide a reliable characterization of the estimation uncertainty. In contrast, the Bayesian-based methods are a suitable alternative to characterize the uncertainty of the estimations.

The Bayesian approach was widely used in exoplanet orbit estimation, characterizing the posterior distribution of the orbital parameters through Markov Chain Monte Carlo (MCMC) sampling. Many variants of the MCMC algorithm have been explored for the characterization of the estimates uncertainty in exoplanets, such as the Metropolis-Hastings within Gibbs sampler [Ford, 2005], the Parallel Tempering sampler [Gregory, 2005, Gregory, 2011], the Affine Invariant MCMC Ensemble sampler [Hou et al., 2012], the Differential Evolution Markov Chain sampler [Nelson et al., 2013] and the Hamiltonian Monte Carlo [Hajian, 2007], among others. One of the most popular MCMC samplers in the statistical community, due to its efficiency on high-dimensional and complex-correlated scenarios, the No-U-Turn sampler, was explored into the exoplanets context [Ji et al., 2017, Shabram et al., 2020].

The Bayesian approach was later adapted to the astrometric orbital estimation in visual binaries [Burgasser et al., 2012, Sahlmann et al., 2013, Lucy, 2014], providing a robust estimation of orbital

parameters. The Bayesian estimation of orbital parameters considering both the positional and the radial velocities sources of observations jointly was later addressed by [Mendez et al., 2017] uses the Metropolis-Hastings within Gibbs sampler method. This method provides the posterior distribution of double-lined visual-spectroscopic binaries. A similar Bayesian-based approach for the joint orbital parameters estimation on these types of binary system was also developed by [Lucy, 2018], and the geometric-based Hamiltonian Monte Carlo was used for the determination of the orbital parameters in binary neutron star [Bouffanais and Porter, 2019]. However, to our knowledge, the Bayesian approach has not been explored in single-lined spectroscopic binaries to estimate the individual masses of the system. In this context, the incorporation of suitable priors on specific parameters of the system has the potential to characterize the posterior distributions of the individual masses and all the orbital parameters of the system. These priors could significantly enrich the analysis of those types of stellar systems.

Recently, the Bayesian-based inference approach was also applied in the context of hierarchical triple stellar systems. For example, [Czekala et al., 2017] proposed the Affine Invariant Markov chain Monte Carlo Ensemble sampler method to compute the posterior distribution of a hierarchical triple system. [Villegas et al., 2021] adapted the alternating fitting approach between the inner and outer astrometric and spectroscopic observations of hierarchical triple systems, proposed by [Tokovinin and Latham, 2017], into the Bayesian framework by proposing a Bayesian hierarchical modeling using the Metropolis-Hastings within Gibbs sampler. Importantly, none of the previously mentioned approaches were used in the context of estimating the parameters of hierarchical systems with higher multiplicities. This is a high-dimensional problem, where the diversity of possible architectures and the lack of sources of observations make the estimation problem very challenging.

In the following chapters, we apply the Bayesian sampling technique No-U-Turn sampler to perform an inference on the orbital parameters of single-lined spectroscopic binaries with a visual orbit (Chapter 5), hierarchical stellar systems (Chapter 6), and to propose a novel methodology to estimate the optimal time of measurement in binary systems (Chapter 7).

Chapter 5

Bayesian Inference in Single-lined Spectroscopic Binaries with a Visual Orbit

The double-lined spectroscopic binaries with a visual orbit are the best observational cases since observations of position and radial velocities of both components are available. This information allows determining the individual masses of the systems. The mass of a star is the most crucial orbital parameter since it allows studying its composition and evolution. Unfortunately, in most binaries, only the spectral lines of the brighter component are visible, and therefore, the observations of the radial velocities of the companion object are absent. This partial information scenario prevents the estimation of the individual masses.

The current chapter introduces a Bayesian modeling for the inference of orbital parameters in single-lined spectroscopic binary systems. Section 5.1 presents an empirical study over the quality of the inference in this type of system by comparing the estimated posterior distribution and its projection on the observation space with the full-information scenario, i.e., the double-lined spectroscopic binaries counterpart. Section 5.2 addresses the problem of the determination of the individual masses in the Bayesian inference through the incorporation of additional information of the system, and Section 5.3 applies the methodology on twelve new single-lined spectroscopic binaries with a visual orbit, providing a complete characterization of the uncertainty of the estimates.

5.1 Quality of the Inference

The proposed Bayesian inference strategy for double and single-lined spectroscopic binaries with a visual orbit is introduced, including the sampling tool and the re-parametrizations of the model used for an efficiency improvement of the inference routine.

On the experimental side, the inference of eight double and single-lined spectroscopic binaries with a visual orbit is performed. We compare the estimated posterior distributions and their projection on the observations space, providing analysis and quantification of the uncertainty gain due to the absence of radial velocities of the companion object.

5.1.1 Bayesian Model

The Bayesian model used to perform the inference in double-lined spectroscopic binary systems, denoted as $SB2$ hereinafter, and single-lined spectroscopic binary systems, denoted as $SB1$ hereinafter, it is nothing more than a suitable re-parametrization of the Keplerian orbital model introduced in Section 2.1.3. For the determination of positional and radial velocities observations, we assume that the observations follow a Gaussian distribution and consider uniform priors for the model's parameters.

Let $\{t_i, X_i, Y_i\}_{i=1}^{n}$ a set of n positional observations of the companion star relative to the primary star of a binary stellar system in rectangular coordinates and $\{t_i, V_{1i}\}_{i=1}^{n_1}, \{t_i, V_{2i}\}_{i=1}^{n_2}$ a sets of n_1 and n_2 radial velocity observations of the primary and companion stars, respectively. Each observation distributes as an independent normal distribution centered in the model estimate with standard deviation equal to the observational error σ_i:

$$X_i \sim \mathcal{N}(\hat{X}_\theta(t_i), \sigma_i^2),\ Y_i \sim \mathcal{N}(\hat{Y}_\theta(t_i), \sigma_i^2),\ V_{1i} \sim \mathcal{N}(\hat{V}_{1\theta}(t_i), \sigma_i^2),\ V_{2i} \sim \mathcal{N}(\hat{V}_{2\theta}(t_i), \sigma_i^2), \tag{5.1}$$

where $(\hat{X}_\theta(t_i), \hat{Y}_\theta(t_i))$ is the estimated position in the orbit at a time t_i, which follows the equation (2.7), and $\hat{V}_{1\theta}(t_i), \hat{V}_{2\theta}(t_i)$ is the estimated radial velocities of each star at a time t_i, which follows the equations (2.17).

The positional and radial velocities estimates are determined through the visual-spectroscopic binary system model presented in Section 2.1.3. Therefore, the set of orbital parameters that characterizes the estimates is $\vartheta_{SB2} = \{P, T, e, a, \omega, \Omega, i, V_0, \pi, q\}$ for the double-lined spectroscopic binaries with a visual orbit, and $\vartheta_{SB1} = \{P, T, e, a, \omega, \Omega, i, V_0, f/\pi\}$ for the single-lined spectroscopic binaries with a visual orbit, where f is the so called fractional-mass of the system. For this last system, the parameter $f/\pi = q/(1+q) \cdot 1/\pi$ condenses the pair of parameters π, q in (2.17), since they are not determinable due to the absence of $\{V_{2i}\}_{i=1}^{n_2}$ observations. The auxiliary parameter f/π has units of parsecs since it is inversely proportional to the parallax π which has units of seconds of arc. The range of f/π is $(0, d_{max}/2]$, considering that $q \in (0, 1]$ and $\pi > 0$, with d_{max} the maximum distance of observation determined by the measurement instrument.

Denoting the set of all observations as $\mathcal{D} = \{t_i, X_i, Y_i\}_{i=1}^{n} \cup \{t_i, V_{1i}\}_{i=1}^{n_1} \cup \{t_i, V_{2i}\}_{i=1}^{n_2}$, the log-likelihood is expressed as:

$$\begin{aligned} \log p(\mathcal{D}|\theta) = &\sum_{i=1}^{n} \log \mathcal{N}(X_i|\hat{X}_\theta(t_i), \sigma_i^2) + \sum_{i=1}^{n} \log \mathcal{N}(Y_i|\hat{Y}_\theta(t_i), \sigma_i^2) \\ &+ \sum_{i=1}^{n_1} \log \mathcal{N}(V_{1i}|\hat{V}_{1\theta}(t_i), \sigma_i^2) + \sum_{i=1}^{n_2} \log \mathcal{N}(V_{2i}|\hat{V}_{2\theta}(t_i), \sigma_i^2). \end{aligned} \tag{5.2}$$

The prior distribution of each orbital parameter is modeled as independent uniform priors on their valid physical range (defined in Section 2.1). Therefore, the prior distribution of the complete set of orbital parameters is expressed as:

$$\log p(\theta) = \sum_{i=1}^{|\theta|} \log U(\min \Theta_i, \max \Theta_i), \tag{5.3}$$

with $\theta_i \in \Theta_i, \forall i \in \{1, ..., |\theta|\}$ and Θ_i the valid physical range of the orbital parameter θ_i.

According to the Bayes theorem (3.1), the posterior distribution is proportional to the likelihood times the prior, i.e., $p(\theta|\mathcal{D}) \propto p(\mathcal{D}|\theta)p(\theta)$, and therefore, the complete posterior distribution can be obtained through any sampling technique. In the following, we compute the posterior distribution using the No-U-Turn sampler algorithm introduced in Section 3.2.2.4. This is a state-of-art Markov chain Monte Carlo (MCMC) method widely adopted by the statistical community in recent years due to its computational efficiency, effectiveness in high-dimensional problems, and theoretical guarantees [Hoffman et al., 2014]. The implementation details of the probabilistic model are presented in Appendix A.

Only the reparametrization of the time of periastron passage T proposed by [Lucy, 2014] is adopted in this work. The author suggests to sample from $T' = (T - t_0)/P$ instead of T, since it is beneficial to sample from a well-constrained parameter space, restricting the range of the time of periastron passage to $[0, 1)$. Reparametrizations that involves a dimensionality reduction of the parameters space (*e.g.* [Mendez et al., 2017]) or transformations of well-constrained parameters (*e.g.* [Ford, 2005]) were avoided since it is shown to have a negative impact on the correlation of the parameters space, considerably hindering the exploration of the parameters space through first-order gradient information. This increases the computational cost of the gradient calculation required by the No-U-Turn sampling routine.

Finally, as the orbital parameter space of binary stellar systems (and especially in hierarchical stellar systems) is highly correlated, many authors recommend choosing a starting point that lies in the typical set of the parameter space (defined in (3.20)), to avoid miss-convergence issues of the sampling process. This work uses the quasi-Newton optimization method *L-BFGS* [Liu and Nocedal, 1989] to find a good starting point that alleviates convergence issues of the Bayesian inference since it allows to perform the optimization using any prior distribution on the parameters. In contrast, other commonly used optimization methods in the astronomical field are restricted to least-squares problems (*e.g.* the Levenberg-Marquardt algorithm [Moré, 1978]), restricting considerably the family of prior distributions that can be used.

5.1.2 Experiments

The inference of eight double-lined spectroscopic binaries with a visual orbit is compared with its single-lined counterpart by omitting the radial velocity observations of the companion object. The obtained estimates and their uncertainties are compared in the parameters space by visualizing the posterior marginal distributions. We also observe the posterior distribution on the observations space through the projection of 1000 randomly selected samples of the posterior distribution on the observation space, drawing trajectories from the first observation time t_0 to the first completion of the orbit $t_0 + P$. The maximum a posteriori MAP (the most probable sample of the posterior distribution) and the 95% high posterior density interval HPDI (the narrowest interval that contains the 95% of the posterior distribution, including the mode) are summarized in Table 5.1.

The inference process is performed by simulating 10000 samples of the respective posterior distributions (discarding the first half for warm-up) on 4 independent Markov chains using the No-U-Turn sampler algorithm. Each chain starts from an initial point determined by the results of the quasi-Newton optimization method $L - BFGS$.

To avoid any redundancy in the analysis, we select three of the eight most representative systems: CHR111, YSC132AaAb, and HIP117186.

5.1.2.1 CHR111

The system CHR111 (HIP111170) is a $SB2$ binary presented and solved in [Mendez et al., 2017]. The available data consists of astrometric observations mainly concentrated in apoastron passage with few observations of the rest of the orbit but abundant observations of both components' radial velocities. The observations and their uncertainties are visualized in Figure 5.2.

The estimated posterior distribution presented in Figure 5.1 has a Gaussian shape with almost no differences in mean value and dispersion between the $SB2$ and $SB1$ cases. The $SB2$ case is slightly more constrained than the $SB1$. This reflects the expected precision gain due to incorporating both radial velocities observations instead of only the primary object radial velocities.

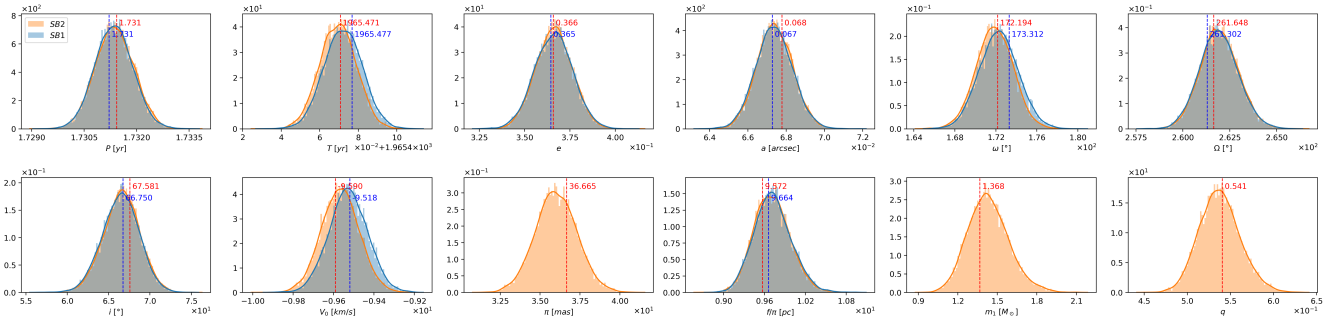

Figure 5.1: Marginal posterior distributions and MAP estimates of the orbital parameters of the CHR111 binary system for the $SB2$ and $SB1$ cases.

The projection of the estimated posterior distribution in the observation space is presented in Figure 5.2. The first column presents the MAP estimate in the observation space. The second and third columns show the projection of 1000 uniformly selected samples of the posterior distribution of $SB2$ and $SB1$ cases, respectively. A slight difference is observed between the MAP orbit of the $SB2$ and $SB1$ cases, but it is almost negligible. There is no difference between the MAP posterior projections in the radial velocities space between both cases. The posterior projection in the orbit space of the $SB2$ case shows less uncertainty in the apoastron, which is the zone that has more samples. Similar uncertainty is observed in the opposite zone -the periastron- where only two observations are available. The zones of the orbit with a higher uncertainty are between the periastron and the apoastron, which coincides with the lack of observations. The posterior projection in the radial velocities space of the $SB2$ case shows almost no uncertainty, attributed to the higher number of samples of radial velocities of both components. However, a slight uncertainty increase of the radial velocity curves of both components is observed in their maximum and minimum points. The projected posterior distributions of the $SB1$ case in both the orbit and radial velocities spaces present no appreciable differences compared to the $SB2$ case, which coincides with the coincidences of their respective posterior distributions.

5.1.2.2 YSC132AaAb

The system YSC132AaAb (HIP89000) is a $SB2$ binary presented and solved in [Mendez et al., 2017]. The available data consists of low and precise astrometric observations mainly concentrated in the apoastron passage but with abundant and precise observations of radial velocities of both components. The observations and their uncertainties are visualized in Figure 5.4.

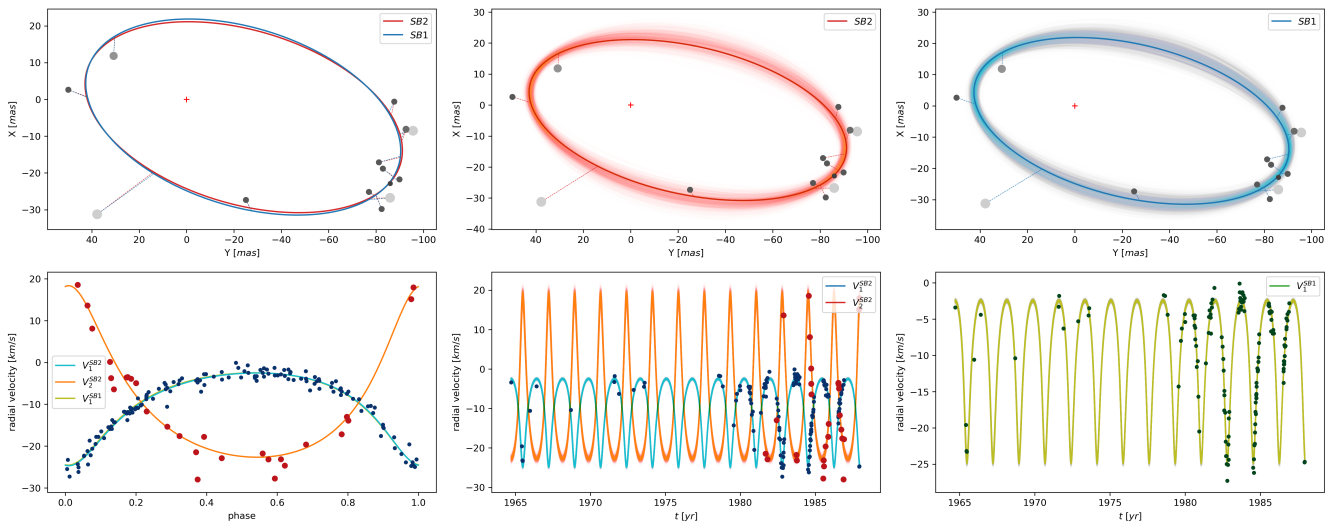

Figure 5.2: Estimated orbits and radial velocity curves of the CHR111 binary system. First column: MAP point estimate projection of the posterior distribution for the $SB2$ and $SB1$ cases. Second column: Projected posterior distribution of the $SB2$ case. Third column: Projected posterior distribution of the $SB1$ case.

The estimated posterior distribution presented in Figure 5.3 shows that almost all the posterior marginal distribution presents a Gaussian shape. The exception to this rule is the parameters $i, f/\pi, m_1$, whose distributions show a high positive skew. Slight differences in the mean values and significant differences in the dispersions of the posterior distribution are observed between the $SB2$ and the $SB1$ case. As expected, the $SB2$ case offers less posterior uncertainty than the $SB1$ case in all the orbital parameters, reflecting the significant impact of incorporating observations of both radial velocities instead of only the primary object radial velocities.

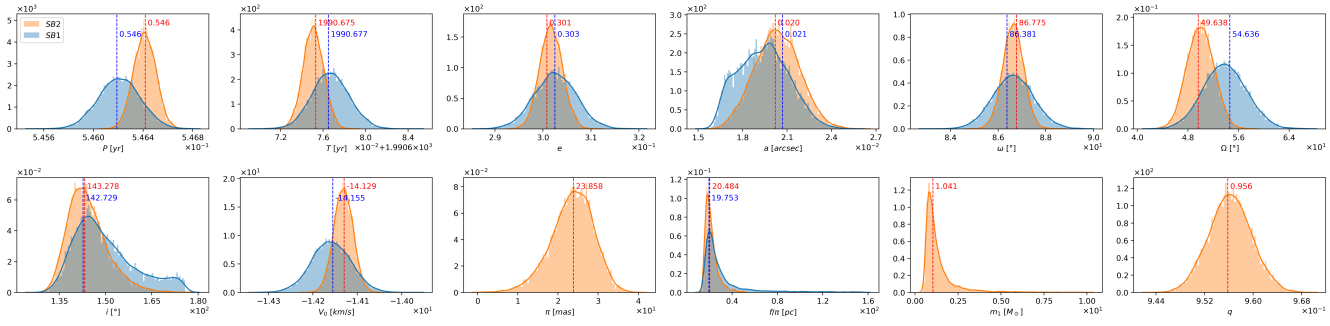

Figure 5.3: Marginal posterior distributions and MAP estimates of the orbital parameters of the YSC132AaAb binary system for the $SB2$ and $SB1$ cases.

The projection of the estimated posterior distribution in the observation space is presented in Figure 5.4. The first column presents the MAP estimate in the observation space. The second and third columns show the projection of 1000 uniformly selected samples of the posterior distribution of $SB2$ and $SB1$ cases, respectively. A slight difference is observed between MAP projection on the orbit between the $SB2$ and $SB1$ cases. There is no difference between the MAP posterior projection in the radial velocities space between both cases. The posterior projection in the orbit space of the $SB2$ case shows less uncertainty in the apoastron, which is the zone that has more

samples. Similar uncertainty is observed in the opposite zone -the periastron- even considering that the periastron presents only two observations. The zones of the orbit with a higher uncertainty are between the periastron and the apoastron, which coincides with the lack of observations. The posterior projection in the radial velocities space of the $SB2$ case presents almost no uncertainty along all the curves, which coincides with the higher number of samples of radial velocities of both components. The projected posterior distribution of the $SB1$ case in the radial velocities space presents no differences compared to the $SB2$ case. However, the posterior distribution projection on the orbit space is higher for the $SB1$ case than the $SB2$ case, coinciding with the appreciable differences in their respective posterior distributions.

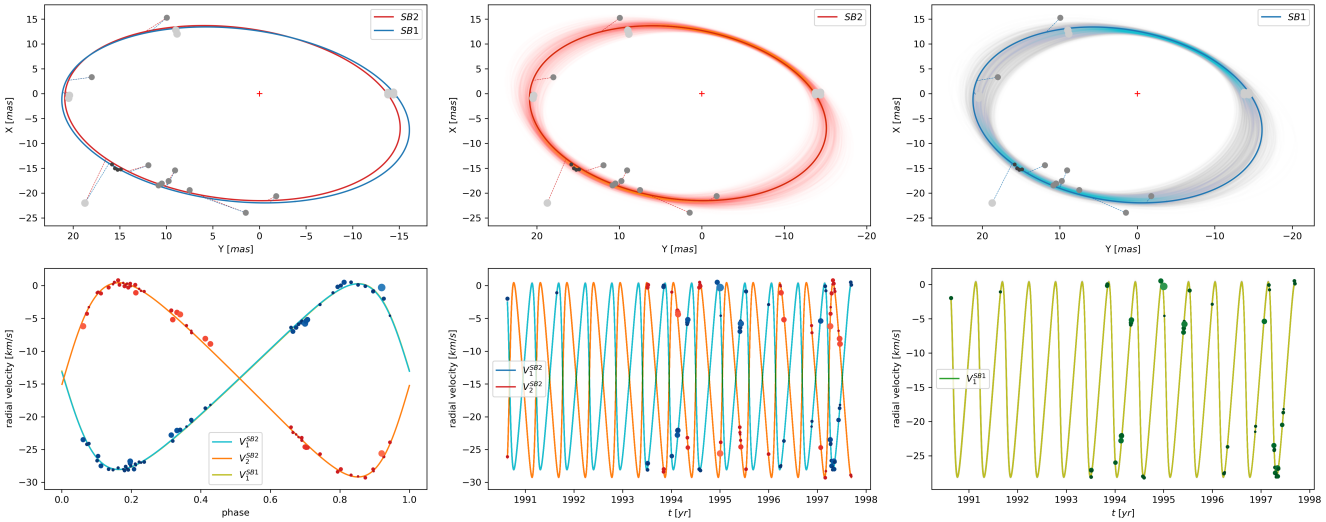

Figure 5.4: Estimated orbits and radial velocity curves of the YSC132AaAb binary system. First column: MAP point estimate projection of the posterior distribution for the $SB2$ and $SB1$ cases. Second column: Projected posterior distribution of the $SB2$ case. Third column: Projected posterior distribution of the $SB1$ case.

5.1.2.3 HIP117186

The system HIP117186 is a $SB2$ binary presented and solved in [Halbwachs et al., 2016]. The available data consists of highly precise astrometric observations dispersed along all the orbit with abundant and precise observations of radial velocities of both components. The observations and their uncertainties are visualized in Figure 5.6.

The estimated posterior distribution presented in Figure 5.5 shows that almost all the posterior marginal distribution presents a Gaussian shape, where slight differences in the mean values and significant differences in the dispersions of the posterior distribution are observed between the $SB2$ and the $SB1$ case. As expected, the $SB2$ case offers less posterior uncertainty than the $SB1$ case in almost all the orbital parameters. The exception to this rule are the angular parameters Ω and i (usually mostly constrained by astrometric observations on visual binaries), where the dispersion between both cases is almost the same. The evident differences in the dispersion of the posterior distribution between the $SB2$ and $SB1$ cases reflect the impact on the uncertainty of the estimates of the incorporation of observations of both radial velocities instead of only using the primary object radial velocities.

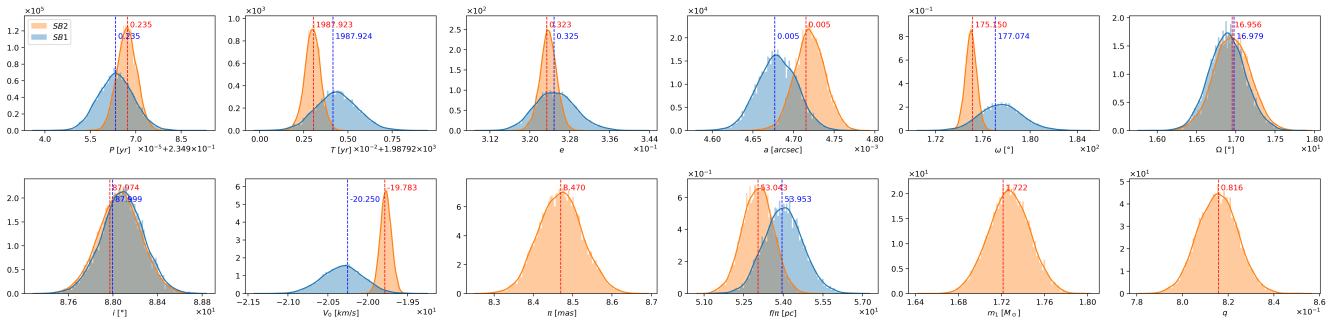

Figure 5.5: Marginal posterior distributions and MAP estimates of the orbital parameters of the HIP117186 binary system for the $SB2$ and $SB1$ cases.

The projection of the estimated posterior distribution in the observation space is presented in Figure 5.6. The first column presents the MAP estimate in the observation space, and the second and third columns show the projection of 1000 uniformly selected samples of the posterior distribution of $SB2$ and $SB1$ cases, respectively. A slight difference is observed between MAP projection on the orbit between the $SB2$ and $SB1$ cases. There is no difference between the MAP posterior projection of the MAP in the radial velocities space between both cases. The posterior projection in the orbit space of the $SB2$ case shows almost no uncertainty in the zones with observations and a slight increase of uncertainty in the other zones. The uncertainty in all the orbit space is almost negligible, attributed to the extremely high precision of the positional observations, even considering that only seven observations were taken. The posterior projection in the radial velocities space of the $SB2$ case shows almost no uncertainty along all the curves, which coincides with the higher number of samples of radial velocities of both components. Finally, the projected posterior distribution in the observations space of the $SB1$ case presents no appreciable differences compared to the $SB2$ case.

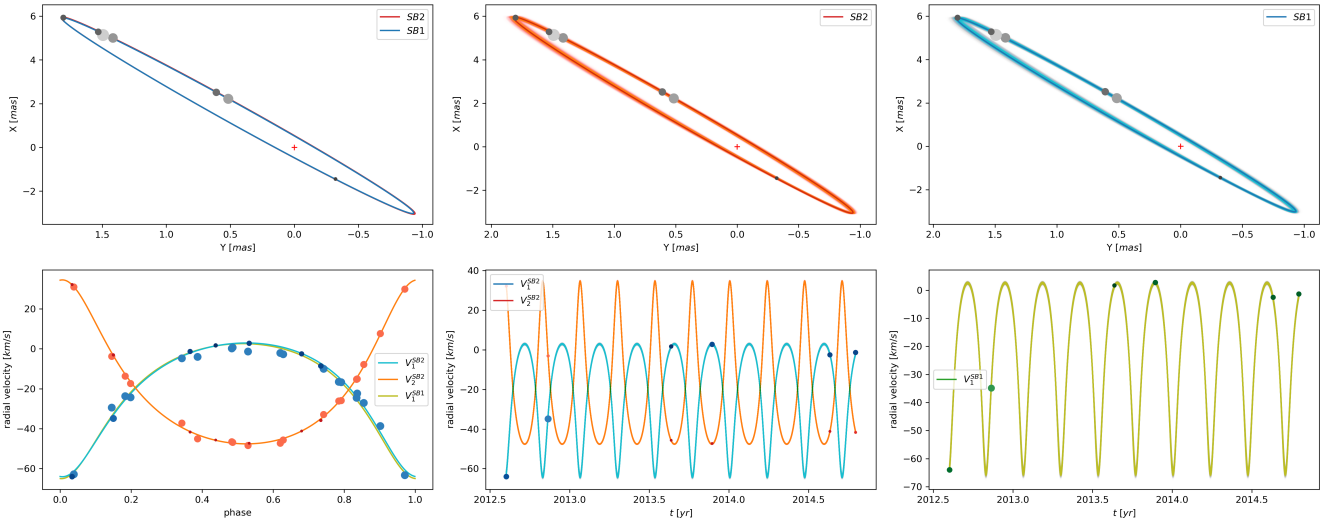

Figure 5.6: Estimated orbits and radial velocity curves of the HIP117186 binary system. First column: MAP point estimate projection of the posterior distribution for the $SB2$ and $SB1$ cases. Second column: Projected posterior distribution of the $SB2$ case. Third column: Projected posterior distribution of the $SB1$ case.

Table 5.1: MAP estimates and 95% HDPIs of the marginal posterior distributions of the orbital parameters incorporating priors of m_1 and π of $SB2$ binary systems, considering the $SB2$ and $SB1$ cases.

System	Case	Author	P [yr]	T [yr]	e	a [']	Ω [°]	ω [°]	i [°]	V_0 [km/s]	π [mas]	f/π [pc]	m_1 [$M_\odot$]	q
CHR111	$SB2$	(a)	1.731	1,965.475	0.367	0.066	172.100	261.393	67.141	-9.573	35.542	9.842	1.389	0.538
	$SB2$	(b)	1.731	1,965.471	0.366	0.068	172.194	261.648	67.581	-9.590	36.665	9.572	1.368	0.541
			[1.73, 1.732]	[1965.452, 1965.488]	[0.348, 0.386]	[0.066, 0.069]	[168.423, 175.407]	[260.027, 263.951]	[62.243, 70.742]	[-9.742, -9.377]	[33.559, 38.62]	[9.214, 10.242]	[1.132, 1.739]	[0.493, 0.584]
	$SB1$	(b)	1.731	1,965.477	0.365	0.067	173.312	261.302	66.750	-9.518	-	9.664	-	-
			[1.73, 1.732]	[1965.453, 1965.492]	[0.346, 0.387]	[0.066, 0.069]	[168.692, 176.002]	[259.93, 263.951]	[62.154, 70.758]	[-9.711, -9.341]	-	[9.213, 10.262]	-	-
YSC132AaAb	$SB2$	(a)	0.546	1,990.675	0.302	0.019	86.650	51.189	146.167	-14.131	21.308	22.938	1.214	0.956
	$SB2$	(b)	0.546	1,990.675	0.301	0.020	86.775	49.638	143.278	-14.129	23.858	20.484	1.041	0.956
			[0.546, 0.547]	[1990.673, 1990.677]	[0.297, 0.307]	[0.017, 0.023]	[85.787, 87.484]	[45.879, 53.981]	[133.337, 157.681]	[-14.173, -14.09]	[13.194, 33.457]	[13.426, 34.668]	[0.495, 3.494]	[0.949, 0.963]
	$SB1$	(b)	0.546	1,990.677	0.303	0.021	86.381	54.636	142.729	-14.155	-	19.753	-	-
			[0.546, 0.547]	[1990.674, 1990.68]	[0.294, 0.311]	[0.016, 0.022]	[84.949, 88.248]	[47.403, 60.572]	[135.808, 173.393]	[-14.247, -14.077]	-	[13.071, 89.385]	-	-
HIP14157	$SB2$	(a)	0.119	1,999.844	0.759	0.006	174.690	19.141	92.240	30.743	19.557	24.193	0.982	0.898
	$SB2$	(b)	0.119	1,993.795	0.761	0.006	175.114	19.437	91.532	30.682	19.436	24.367	1.003	0.900
			[0.119, 0.119]	[1993.795, 1993.796]	[0.758, 0.763]	[0.006, 0.006]	[174.829, 175.564]	[19.092, 19.924]	[90.104, 92.968]	[30.573, 30.78]	[19.279, 19.604]	[24.145, 24.616]	[0.982, 1.018]	[0.891, 0.91]
	$SB1$	(b)	0.119	1,993.795	0.761	0.006	175.044	19.530	91.454	30.695	-	24.401	-	-
			[0.119, 0.119]	[1993.795, 1993.796]	[0.758, 0.765]	[0.006, 0.006]	[174.479, 175.503]	[19.089, 19.905]	[90.131, 92.966]	[30.569, 30.841]	-	[24.146, 24.652]	-	-
HIP20601	$SB2$	(a)	0.428	2,013.942	0.851	0.011	202.026	340.526	103.138	41.623	16.702	25.486	0.980	0.741
	$SB2$	(b)	0.428	1,971.983	0.852	0.011	-157.740	340.812	103.636	42.097	16.710	25.677	0.984	0.752
			[0.428, 0.428]	[1971.983, 1971.983]	[0.849, 0.854]	[0.011, 0.012]	[-158.465, -156.891]	[339.649, 341.814]	[101.114, 105.151]	[41.94, 42.251]	[16.429, 16.991]	[25.178, 26.099]	[0.947, 1.013]	[0.742, 0.759]
	$SB1$	(b)	0.428	1,971.983	0.852	0.011	-157.572	340.506	102.627	42.225	-	25.623	-	-
			[0.428, 0.428]	[1971.983, 1971.983]	[0.849, 0.855]	[0.011, 0.012]	[-158.361, -156.482]	[339.57, 341.959]	[100.916, 105.761]	[41.922, 42.433]	-	[25.163, 26.184]	-	-
HIP117186	$SB2$	(a)	0.235	2,013.301	0.327	0.005	176.070	16.928	88.054	-19.890	8.445	53.509	1.686	0.824
	$SB2$	(b)	0.235	1,987.923	0.323	0.005	175.150	16.956	87.974	-19.783	8.470	53.043	1.722	0.816
			[0.235, 0.235]	[1987.922, 1987.924]	[0.32, 0.327]	[0.005, 0.005]	[174.106, 175.971]	[16.5, 17.415]	[87.683, 88.42]	[-19.903, -19.632]	[8.366, 8.58]	[51.939, 54.212]	[1.69, 1.765]	[0.799, 0.833]
	$SB1$	(b)	0.235	1,987.924	0.325	0.005	177.074	16.979	87.999	-20.250	-	53.953	-	-
			[0.235, 0.235]	[1987.922, 1987.927]	[0.317, 0.333]	[0.005, 0.005]	[-179.152, 179.999]	[16.446, 17.354]	[87.717, 88.459]	[-20.83, -19.843]	-	[52.613, 55.459]	-	-
HD6840	$SB2$	(a)	7.454	2,002.952	0.743	0.083	215.600	151.300	50.900	-9.570	16.400	29.048	1.222	0.910
	$SB2$	(b)	7.455	1,995.494	0.744	0.082	-144.994	152.001	50.669	-9.526	16.201	29.379	1.244	0.908
			[7.45, 7.458]	[1995.49, 1995.502]	[0.741, 0.746]	[0.082, 0.083]	[-145.468, -144.197]	[151.314, 152.497]	[49.926, 51.606]	[-9.596, -9.477]	[15.952, 16.572]	[28.751, 29.857]	[1.176, 1.294]	[0.894, 0.925]
	$SB1$	(b)	7.454	1,995.495	0.743	0.083	-145.022	152.032	50.782	-9.500	-	29.345	-	-
			[7.45, 7.458]	[1995.488, 1995.5]	[0.741, 0.746]	[0.082, 0.083]	[-145.95, -144.551]	[151.58, 152.828]	[49.865, 51.528]	[-9.551, -9.417]	-	[28.914, 30.048]	-	-
HIP108917	$SB2$	(a)	2.241	1,970.992	0.496	0.080	92.870	268.341	74.479	-10.743	32.170	8.246	2.246	0.361
	$SB2$	(b)	2.245	1,968.750	0.463	0.072	90.343	273.135	67.582	-10.796	37.243	7.239	1.043	0.369
			[2.245, 2.246]	[1968.744, 1968.756]	[0.452, 0.469]	[0.071, 0.072]	[89.919, 90.644]	[272.559, 273.573]	[67.044, 67.983]	[-11.394, -10.035]	[34.622, 40.09]	[6.073, 8.181]	[0.807, 1.266]	[0.3, 0.425]
	$SB1$	(b)	2.245	1,968.751	0.461	0.072	90.306	273.266	67.505	-10.791	-	7.130	-	-
			[2.245, 2.246]	[1968.745, 1968.759]	[0.452, 0.468]	[0.071, 0.072]	[89.94, 90.695]	[272.696, 273.69]	[67.005, 67.956]	[-11.745, -10.193]	-	[6.002, 8.091]	-	-
HIP677	$SB2$	(a)	0.265	1,988.583	0.535	0.024	79	104.400	105.600	-11	33.620	10.023	3.441	0.508
	$SB2$	(b)	0.265	1,988.847	0.535	0.024	78.339	104.952	106.172	-12.465	32.796	10.500	3.697	0.525
			[0.265, 0.265]	[1988.842, 1988.851]	[0.47, 0.589]	[0.022, 0.025]	[73.74, 81.112]	[101.653, 108.823]	[102.45, 109.408]	[-14.043, -9.567]	[29.447, 35.93]	[9.113, 12.139]	[2.983, 4.461]	[0.449, 0.6]
	$SB1$	(b)	0.265	1,988.848	0.529	0.024	77.504	104.112	106.075	-10.194	-	10.952	-	-
			[0.264, 0.265]	[1988.843, 1988.852]	[0.461, 0.599]	[0.022, 0.026]	[73.201, 81.426]	[100.809, 108.109]	[102.329, 109.601]	[-13.237, -7.576]	-	[9.117, 12.265]	-	-

(a) Results reported by other authors. (b) Results obtained in this work.

5.1.3 Concluding Remarks

The experiments presented in this section show an uncertainty reduction of the estimated posterior distributions when radial velocities observations of both components are available instead of only one component, as well as a slight shift on the MAP value of the posterior distributions in some orbital parameters. The orientation and magnitude of the shift between the posterior distribution of the $SB1$ and $SB2$ cases and the magnitude of the uncertainty reduction do not follow an evident tendency along the dimensions of the posterior distribution, neither along the different systems studied. The shift and dispersion magnitudes differences between the posterior distribution of $SB2$ and $SB1$ cases depend on the system itself and the quality and quantity of the observations available. All the studied systems present an almost negligible difference in the MAP values distance and dispersion magnitude of the posterior distribution of both cases. Finally, the MAP estimates of both the $SB2$ and $SB1$ cases in all the systems studied were very similar to the estimations reported by other authors using different methodologies.

The projected uncertainty in the observations space is lower in the zone with observations and higher in the zones without observations. The joint estimation of the orbit and radial velocity curves allows sharing the knowledge provided by both sources of information, reducing the uncertainty of the estimates in the observations space significantly even if one source of information is highly uncertain. This was observed in the system HIP108917 where the projected radial velocity curves are lowly uncertain even when the respective observations are highly uncertain. The projected orbits and radial velocity curves of the $SB2$ and $SB1$ cases show almost no differences in all the studied systems and the MAP point estimate projections of the posterior distributions. The only appreciable difference between the uncertainty projection on the observations space was in the orbit of the YSC132AaAb system, where the $SB1$ case was slightly higher uncertain than the $SB2$ counterpart case. This is attributed due to the higher uncertainty of its positional observations compared to the other studied systems.

As all the studied systems are well determined through abundant and precise observations, the differ-

ences between the posterior distributions and their projections on the observations space was negligible, meaning that the information provided by the radial velocity observations of the companion object is redundant, being irrelevant in the inference process, and hence, in the estimation of the orbital parameters. However, when the observations are not abundant or precise, the use of the radial velocity observations of both components visibly reduces the posterior uncertainty instead of using only one of them.

5.2 Determining the Mass Ratio

The single-lined spectroscopic binary with a visual orbit is an abundant observational object. Unfortunately, the lack of observations of the radial velocity of the companion object does not allow estimating the mass ratio of the system and its individual masses. This limits the use of this type of data in astronomical studies. However, suitable additional information about the system can be incorporated to determine the individual masses, such as the system parallax and the mass of the primary object.

This section presents a Bayesian modeling that incorporates information of the parallax and the primary object mass of the system in the sampling routine. This allows performing a joint Bayesian inference on all the orbital parameters, including the mass ratio, which allows determining the individual masses of the system. A comparative study of the estimated posterior distribution of single-lined visual-spectroscopic binary systems is presented for three cases: incorporating prior information on the parallax, incorporating prior information on the mass of the primary object and incorporating prior information on both the parallax and the mass of the primary object. These results are also compared to the inference of single-lined spectroscopic binary systems without incorporating additional information on the system. The experiments are performed on the same eight binary systems previously studied in Section 5.1.

5.2.1 Bayesian Model

As presented in Section 5.1, the Bayesian model for the single-lined spectroscopic binary systems $SB1$ is characterized by the set of orbital parameters $\vartheta_{SB1} = \{P, T, e, a, \omega, \Omega, i, V_0, f/\pi\}$, where the parameter f/π replaces the individual parameters π and q, since they are non-identifiable in absence of radial velocities observations of the companion object, i.e., exists different values of π, q that maps on the same value on the posterior distribution, preventing to determine their values. The non-identifiability of the mass ratio implies that the individual masses of this type of binary system can neither be determined. As the individual masses are relevant for studying these systems, many authors address the non-identifiability problem on $SB1$ systems by incorporating information about the parallax parameter from external measurements (independent from positional and radial velocities observations). This information is added in a posteriori manner, i.e., once the initial set of orbital parameters ϑ_{SB1} is estimated from the observations. An essential disadvantage of this approach is that it ignores the influence of the external information on the estimation, which leads to a sub-optimal solution. Moreover, if the additional information about parallax is highly uncertain or biased, the estimated mass ratio given the previously estimated orbital parameters and the parallax can be out of the valid physical range $q \in (0, 1]$.

In order to alleviate the non-identifiability problem of the parameter q in this setting, two different approaches are presented: one based on the incorporation of prior information about the parallax $p(\pi)$ and the other based on the incorporation of an additional observation about the primary object mass in the likelihood $p(m_1|\theta)$.

The first proposed approach makes use of the $SB2$ orbital model described in Section 2.1.3, characterized by the set of parameters $\vartheta_{SB2} = \{P, T, e, a, \omega, \Omega, i, V_0, \pi, q\}$, with the incorporation of an informative prior distribution on the parallax $p(\pi)$. Specifically, $p(\pi)$ is modeled as a normal distri-

bution with mean and standard deviation determined respectively by the measurement π_{obs} and error σ_π of the astronomical parallax. These are precisely determined for most of the observed systems [Wenger et al., 2000, Prusti et al., 2016, Brown et al., 2018]. The addition of the prior $p(\pi) = \mathcal{N}(\pi_{obs}, \sigma_\pi^2)$ makes the model soft-identifiable, i.e., the indeterminable parameters turns determinable through the incorporation of suitable priors. This allow the determination the mass ratio q of the system.

Alternatively, since only the spectral line of the primary object of the $SB1$ system is visible, these observations can be used to estimate the mass of the primary object m_1 through empirical relations [Habets and Heintze, 1981, Straižys and Kuriliene, 1981, Aller et al., 1996], providing additional external information that allows alleviating the non-identifiability problem of the mass ratio as well. In fact, the primary object mass m_1 relates to the set of orbital parameters of ϑ_{SB2} through the third law of Kepler:

$$\hat{m}_1 = \left(\frac{a}{\pi}\right)^3 \cdot \frac{1}{P^2(1+q)}. \tag{5.4}$$

Therefore, by setting that the m_1 observation is independent from the other system observations and that distributes as Gaussian with it mean and standard deviation determined respectively by $\hat{m}_1$ and the measurement error σ_{m_1}, the distribution $p(m_1|\theta) = \mathcal{N}(\hat{m}_1, \sigma_{m_1}^2)$ can be directly incorporated in the likelihood term in (5.2).

Finally, those two approaches can be incorporated simultaneously in the inference routine, i.e., incorporating a prior $p(\pi)$ and adding the term $p(m_1|\theta)$ in the likelihood computation.

5.2.2 Experiments

The inference of eight well-determined double-lined spectroscopic binaries with a visual orbit (performed in Section 5.1) is compared with its counterpart omitting the radial velocity observations of the companion object. For this comparison we use the three mentioned different approaches to determine the mass ratio: the incorporation of a prior distribution on the parallax $p(\pi)$, denoted as $SB1 + p(\pi)$ hereinafter, the incorporation of an observation of the primary object mass in the likelihood term $p(m_1|\theta)$, denoted as $SB1 + p(m_1|\theta)$ hereinafter, and incorporating both sources of additional information, denoted as $SB1 + p(\pi) + p(m_1|\theta)$ hereinafter. The estimates and their uncertainties are compared in the parameters space through visualization of the posterior marginal distributions and the observations space through the projection of 1000 randomly selected samples of the posterior distribution on the observation space. For the last analysis, we draw trajectories from the first observation time t_0 to the first completion of the orbit $t_0 + P$. The maximum a posteriori estimation (MAP) and the 95% confidence interval around the MAP solution are summarized in Table 5.2. The MAP estimation error, high densities intervals lengths and estimated Kullback-Leibler divergence (a measure of similarity between probability distributions) between the marginal posterior distributions of the mass ratio q between the full-information $SB2$ case and the $SB1$ cases with additional information $SB1 + p(\pi)$, $SB1 + p(m_1|\theta)$ and $SB1 + p(\pi) + p(m_1|\theta)$ are presented in Table 5.3. The Kullback-Leibler divergence is estimated through the k-nearest neighbor method [Wang et al., 2009].

The inference process is performed through the simulation of 10000 samples of the respective posterior distributions (discarding the first half for warm-up) on 4 independent Markov chains using the No-U-Turn sampler algorithm as presented in Section 5.1.

For clarity, the analysis is focused only on three of the eight systems: CHR111, YSC132AaAb, and HIP117186.

5.2.2.1 CHR111

Figure 5.7 shows that the posterior distribution are almost equal for all the orbital parameters except of the parameters π, m_1 and q, which are identified through the incorporation of the additional information $p(\pi)$ or $p(m_1|\theta)$. It is noted that, for these last trio of parameters, their distributions are shifted with respect to the distributions of the case $SB2$, i.e., they show a slight discrepancy with respect to the full information scenario $SB2$. For the $SB1 + p(\pi)$ case, the posterior distribution of π is equal to the prior $p(\pi)$, while for the $SB1 + p(m_1|\theta)$ case, the uncertainty of the posterior distribution of m_1 is equal to the likelihood $p(m_1|\theta)$, which follows the soft-identifiability of both models on the corresponding parameters. All the cases with additional information offer a significant reduction on the uncertainty of the posterior distribution of π, m_1, q relative to the full-information scenario $SB2$. The mean values of the posterior distribution of π, m_1, q is almost the same for the cases with additional information, but are slightly biased with respect to the full information scenario $SB2$. The mixed case $SB1+p(\pi)+p(m_1|\theta)$ presents the lowest uncertainty on π, m_1, q, followed by the $SB1+p(m_1|\theta)$ and $SB1+p(\pi)$ cases. $SB1+p(\pi)+p(m_1|\theta)$ is the only case that presents a variation on mean and variance of the posterior distribution of the semi-major axis a with respect to all other cases. Finally, it can be observed that in the mixed case the posterior distribution of π is in between the posterior distributions of the $SB1+p(\pi)$ and $SB1+p(m_1|\theta)$ cases, and the posterior distributions of m_1 and q are almost equal to the posterior distribution of the $SB1+p(m_1|\theta)$ and different to the $SB1 + p(\pi)$ case.

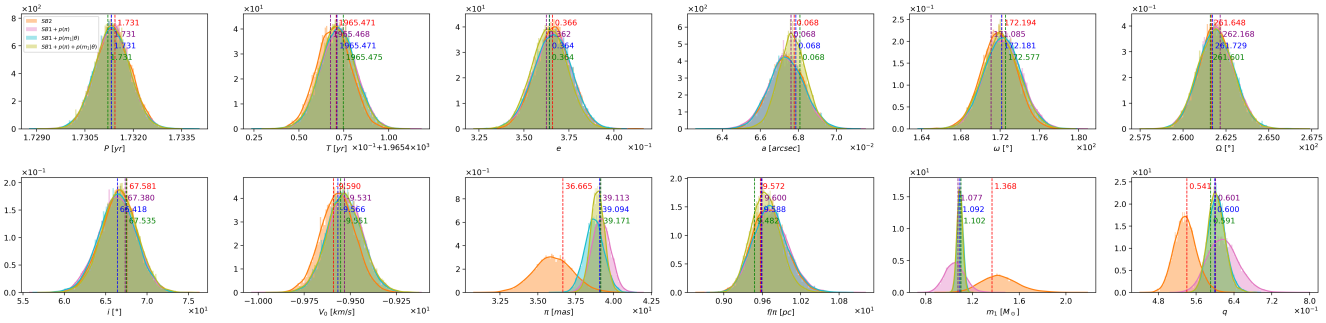

Figure 5.7: Marginal posterior distributions and MAP estimates of the orbital parameters of the CHR111 binary system for the $SB1$, $SB1+p(\pi)$, $SB1+p(m_1|\theta)$ and $SB1+p(\pi)+p(m_1|\theta)$ cases.

The projection of the estimated posterior distributions in the observation space is presented in Figure 5.8, where no significant differences are observed in the MAP and uncertainties projections between all the four cases.

5.2.2.2 YSC132AaAb

Figure 5.9 shows that the posterior distributions of the cases with additional information are almost equal except of the parameters π, m_1 and q, which are identified through the incorporation of the additional information $p(\pi)$ or $p(m_1|\theta)$, and the orbital parameters a and i. The posterior distribution of the other orbital parameters are equal to the posterior distributions of the $SB1$ case presented in the previous section. For the $SB1 + p(\pi)$ case, the posterior distribution of π is equal to the prior $p(\pi)$, while for the $SB1 + p(m_1|\theta)$ case, the uncertainty of the posterior distribution of m_1 is equal to the likelihood $p(m_1|\theta)$. All the cases with additional information present a significant reduction on the uncertainty of the posterior distribution of π, m_1, a, i with respect to the full-information scenario $SB2$. In contrast, they show an increase of the uncertainty of the posterior distribution of q. The mixed case $SB1+p(\pi)+p(m_1|\theta)$ presents the lowest uncertainty on π, m_1, q, followed by the $SB1 + p(m_1|\theta)$ and $SB1 + p(\pi)$ cases. The posterior distribution of the angular parameters a, i of the $SB1 + p(\pi)$ and $SB1 + p(m_1|\theta)$ cases are pretty similar but significantly different to the mixed case $SB1 + p(\pi) + p(m_1|\theta)$. It can be observed that, in the mixed

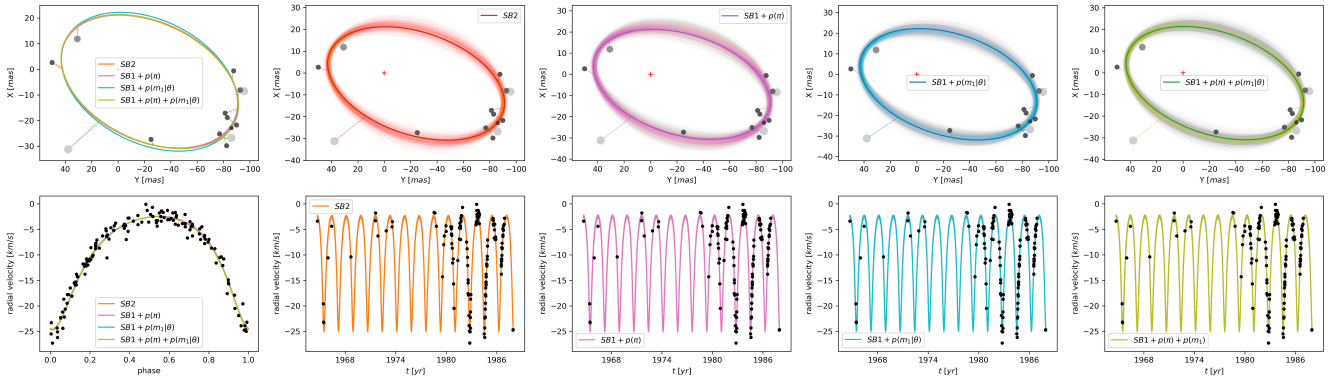

Figure 5.8: Estimated orbits and radial velocity curves of the CHR111 binary system. First column: MAP point estimate projection of the posterior distribution for the $SB2$, $SB1 + p(\pi)$, $SB1 + p(m_1|\theta)$ and $SB1 + p(\pi) + p(m_1|\theta)$ cases. Second column: Projected posterior distribution of the $SB2$ case. Third column: Projected posterior distribution of $SB1 + p(\pi)$ case. Fourth column: Projected posterior distribution of the $SB1 + p(m_1|\theta)$ case. Fifth column: Projected posterior distribution of the $SB1 + p(\pi) + p(m_1|\theta)$ case.

case, the posterior distribution of π is almost equal to the posterior distributions of the $SB1 + p(\pi)$ case, the posterior distribution of m_1 is almost equal to the posterior distributions of the $SB1 + p(m_1|\theta)$ case and the posterior distribution of q is significantly different to all other cases.

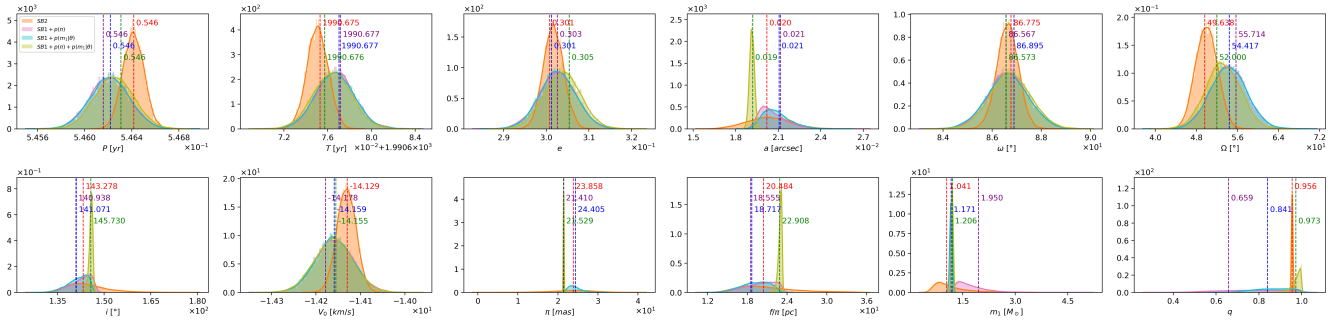

Figure 5.9: Marginal posterior distributions and MAP estimates of the orbital parameters of the YSC132AaAb binary system for the $SB1$, $SB1 + p(\pi)$, $SB1 + p(m_1|\theta)$ and $SB1 + p(\pi) + p(m_1|\theta)$ cases.

The projection of the posterior distributions on the observation space is presented in Figure 5.10. The trajectories of the MAP estimators in the orbit space presents evident differences between all the cases. The $SB1 + p(\pi)$ and $SB1 + p(m_1|\theta)$ cases presents a slight reduction of the projected uncertainty in the orbit space with respect to the $SB2$ case, while the mixed case $SB1 + p(\pi) + p(m_1|\theta)$ presents a significant uncertainty reduction. The orbital uncertainty projection of the mixed case also presents a different orbital shape compared to all other cases, with a worst fitting on the most precise observations (in rectangular coordinates $X = -15, Y = 15$). This reflects that the low standard deviations of the incorporated additional information play a major role in the inference procedure to the detriment of fitting the other positional observations. No significant differences are observed for the MAP and uncertainties projections in the radial velocities space between all four cases.

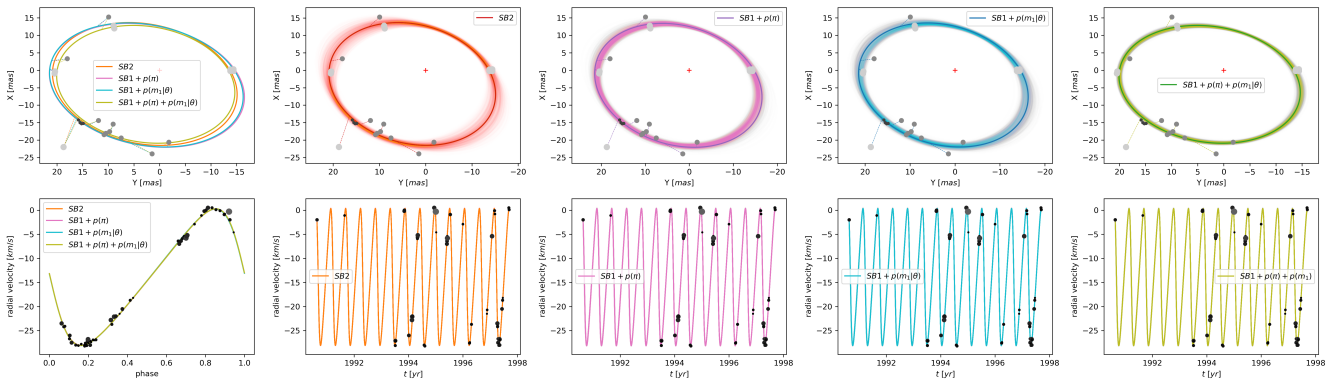

Figure 5.10: Estimated orbits and radial velocity curves of the YSC132AaAb binary system. First column: MAP point estimate projection of the posterior distribution for the $SB2$, $SB1 + p(\pi)$, $SB1 + p(m_1|\theta)$ and $SB1 + p(\pi) + p(m_1|\theta)$ cases. Second column: Projected posterior distribution of the $SB2$ case. Third column: Projected posterior distribution of $SB1 + p(\pi)$ case. Fourth column: Projected posterior distribution of the $SB1 + p(m_1|\theta)$ case. Fifth column: Projected posterior distribution of the $SB1 + p(\pi) + p(m_1|\theta)$ case.

5.2.2.3 HIP117186

Figure 5.11 shows that the posterior distributions of the $SB1 + p(\pi)$ and $SB1 + p(m_1|\theta)$ cases are almost equal except of the parameters π, m_1 and q, which are identified through the incorporation of the additional information $p(\pi)$ or $p(m_1|\theta)$. The posterior distribution of the other orbital parameters are equal to the posterior distributions of the $SB1$ case presented in the previous section. For the $SB1 + p(\pi)$ case, the posterior distribution of π is equal to the prior $p(\pi)$, while for the $SB1 + p(m_1|\theta)$ case, the uncertainty of the posterior distribution of m_1 is equal to the likelihood $p(m_1|\theta)$. All the studied cases with additional information presents a significant increase on the uncertainty of the posterior distribution of all the orbital parameters (with the exception of the angular parameters Ω and i) with respect to the full-information scenario $SB2$. The mixed case $SB1 + p(\pi) + p(m_1|\theta)$ presents the lowest uncertainty on π, m_1, q, closely followed by the $SB1 + p(m_1|\theta)$ and $SB1 + p(\pi)$ cases. The posterior distribution of the mixed case $SB1 + p(\pi) + p(m_1|\theta)$ presents a slight bias in all the orbital parameters except the angular parameters Ω and i, with respect to the $SB1$ case, and therefore, with respect to the $SB1 + p(\pi)$ and $SB1 + p(m_1|\theta)$ cases. It can be observed that, in the mixed case, the posterior distribution of π is in between the $SB1 + p(\pi)$ and $SB1 + p(m_1|\theta)$ but nearest to the first one. On the other hand, the posterior distribution of m_1 is in between the $SB1 + p(\pi)$ and $SB1 + p(m_1|\theta)$ distributions with a similar distance between them, and the posterior distribution of q is in between the $SB1 + p(\pi)$ and $SB1 + p(m_1|\theta)$ but nearest to the second one.

Finally, the projection of the estimated posterior distributions in the observation space is presented in Figure 5.12, where no significant differences are observed in the MAP and uncertainties projections between all the four cases.

5.2.3 Concluding Remarks

The developed experiments show an uncertainty reduction of the estimated posterior distributions concerning the $SB1$ case when additional information on the parallax or mass of the primary object is incorporated. In general, we observe that the more information is available, the less the uncertainty obtained in the estimates. Consequently, the case that incorporates both sources of additional information $SB1 + p(\pi) + p(m_1|\theta)$ presents the lower uncertainty. Due to the non-identifiability of the parameters π, m_1 and q in the single-lined binary system with a visual orbit model, the prior information $p(\pi)$ results

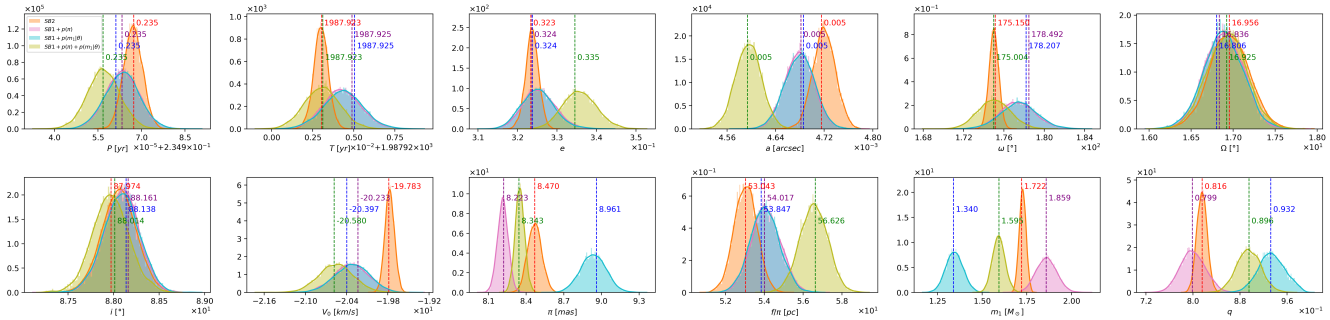

Figure 5.11: Marginal posterior distributions and MAP estimates of the orbital parameters of the HIP117186 binary system for the $SB1$, $SB1 + p(\pi)$, $SB1 + p(m_1|\theta)$ and $SB1 + p(\pi) + p(m_1|\theta)$ cases.

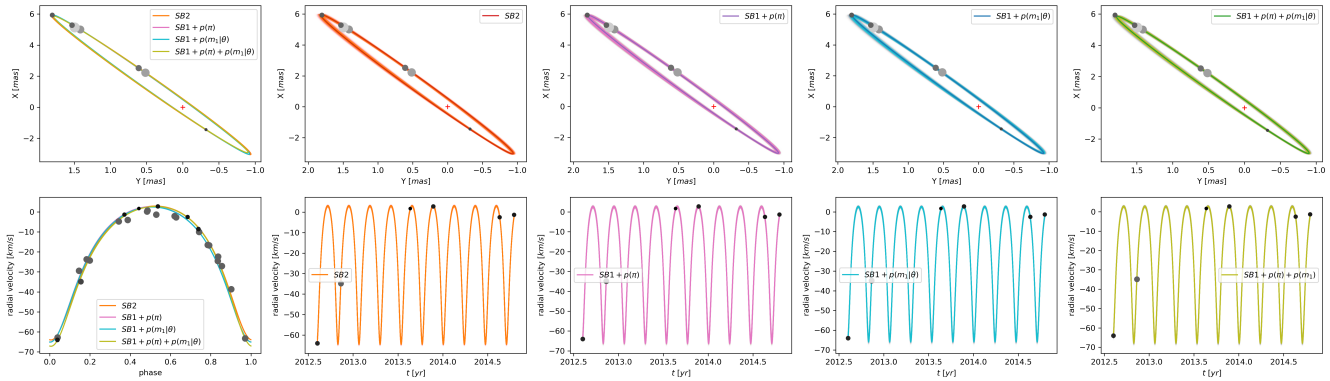

Figure 5.12: Estimated orbits and radial velocity curves of the HIP117186 binary system. First column: MAP point estimate projection of the posterior distribution for the $SB2$, $SB1 + p(\pi)$, $SB1+p(m_1|\theta)$ and $SB1+p(\pi)+p(m_1|\theta)$ cases. Second column: Projected posterior distribution of the $SB2$ case. Third column: Projected posterior distribution of $SB1+p(\pi)$ case. Fourth column: Projected posterior distribution of the $SB1 + p(m_1|\theta)$ case. Fifth column: Projected posterior distribution of the $SB1 + p(\pi) + p(m_1|\theta)$ case.

equal to the marginal posterior distribution of π and the likelihood term $p(m_1|\theta)$ results equal to the marginal posterior distribution of m_1. The major differences observed in the posterior distributions of the cases with additional information are in the trio of orbital parameters π, m_1, q. The marginal posterior distribution of those trio of parameters of the mixed case $SB1+p(\pi)+p(m_1|\theta)$ is in between the posterior distributions of the $SB1 + p(\pi)$ and $SB1 + p(m_1|\theta)$ cases. These distributions can be equidistant to the $SB1 + p(\pi)$ and $SB1 + p(m_1|\theta)$ cases, if both sources of information are equally likely according to the model, or can be near to one of them if one source of information is more likely than the other. The similarity of the posterior distribution observed with the $SB1 + p(\pi) + p(m_1|\theta)$ scenario to one of the simple cases, $SB1 + p(\pi)$ and $SB1 + p(m_1|\theta)$, allows determining the most reliable source of additional information according to the model and the observations. For example, if the posterior distributions (in the trio of parameters π, m_1, q) of the mixed case are nearest to the distribution of the $SB1+p(m_1|\theta)$ than the $SB1 + p(\pi)$, then the primary object mass additional observation is more reliable than the additional parallax observation since the last one results negligible for the inference process in the mixed scenario.

The joint estimation of the orbital and radial velocities observations and the incorporation of additional information have an impact on the estimated posterior distribution of some of the identifiable orbital parameters. This impact depends on the identifiable parameters depends on the uncertainty of the additional

Table 5.2: MAP estimates and 95% HDPIs of the marginal posterior distributions of the orbital parameters incorporating priors of m_1 and π of $SB2$ binary systems, considering the $SB2$, $SB1 + p(\pi)$, $SB1 + p(m_1|\theta)$, and $SB1 + p(\pi) + p(m_1|\theta)$ cases.

System	Case	P [yr]	T [yr]	e	a [']	Ω [°]	ω [°]	i [°]	V_0 [km/s]	π [mas]	f/π [pc]	m_1 [$M_\odot$]	q
CHR111	$SB2$	1.731	1,965.471	0.366	0.068	172.194	261.648	67.581	-9.590	36.665	9.572	1.368	0.541
		[1.73, 1.732]	[1965.452, 1965.488]	[0.348, 0.386]	[0.066, 0.069]	[168.423, 175.407]	[260.027, 263.951]	[62.243, 70.742]	[-9.742, -9.377]	[33.559, 38.62]	[9.214, 10.242]	[1.132, 1.739]	[0.493, 0.584]
	$SB1 + p(\pi)$	1.731	1,965.468	0.362	0.068	171.085	262.168	67.380	-9.531	39.113	9.600	1.077	0.601
		[1.73, 1.732]	[1965.454, 1965.492]	[0.346, 0.386]	[0.066, 0.069]	[168.731, 176.15]	[259.96, 263.976]	[61.975, 70.685]	[-9.713, -9.336]	[37.957, 40.408]	[9.207, 10.27]	[0.884, 1.211]	[0.556, 0.681]
	$SB1 + p(m_1\|\theta)$	1.731	1,965.471	0.364	0.068	172.181	261.729	66.418	-9.566	39.094	9.588	1.092	0.600
		[1.73, 1.732]	[1965.452, 1965.491]	[0.345, 0.386]	[0.066, 0.069]	[168.497, 175.902]	[260.019, 263.988]	[62.114, 70.835]	[-9.713, -9.349]	[37.47, 39.923]	[9.203, 10.243]	[1.046, 1.143]	[0.569, 0.641]
	$SB1 + p(\pi) + p(m_1\|\theta)$	1.731	1,965.475	0.364	0.068	172.577	261.601	67.535	-9.551	39.171	9.482	1.102	0.591
		[1.73, 1.732]	[1965.453, 1965.491]	[0.344, 0.383]	[0.066, 0.069]	[168.539, 175.718]	[259.927, 263.912]	[62.458, 70.829]	[-9.723, -9.345]	[38.068, 39.817]	[9.229, 10.106]	[1.046, 1.14]	[0.568, 0.637]
YSC132AaAb	$SB2$	0.546	1,990.675	0.301	0.020	86.775	49.638	143.278	-14.129	23.858	20.484	1.041	0.956
		[0.546, 0.547]	[1990.673, 1990.677]	[0.297, 0.307]	[0.017, 0.023]	[85.787, 87.484]	[45.879, 53.981]	[133.337, 157.681]	[-14.173, -14.09]	[13.194, 33.457]	[13.426, 34.668]	[0.495, 3.494]	[0.949, 0.963]
	$SB1 + p(\pi)$	0.546	1,990.677	0.303	0.021	86.567	55.714	140.938	-14.178	21.410	18.555	1.950	0.659
		[0.546, 0.547]	[1990.673, 1990.68]	[0.294, 0.311]	[0.019, 0.022]	[85.092, 88.254]	[47.554, 61.296]	[136.198, 147.662]	[-14.241, -14.075]	[21.271, 21.636]	[15.978, 23.353]	[1.182, 2.524]	[0.524, 1.0]
	$SB1 + p(m_1\|\theta)$	0.546	1,990.677	0.301	0.021	86.895	54.417	141.071	-14.159	24.405	18.717	1.171	0.841
		[0.546, 0.547]	[1990.673, 1990.68]	[0.294, 0.311]	[0.019, 0.023]	[84.988, 88.268]	[47.335, 61.069]	[135.668, 146.27]	[-14.24, -14.077]	[21.731, 27.054]	[15.412, 22.656]	[1.1, 1.236]	[0.727, 1.0]
	$SB1 + p(\pi) + p(m_1\|\theta)$	0.546	1,990.676	0.305	0.019	86.573	52.000	145.730	-14.155	21.529	22.908	1.206	0.973
		[0.546, 0.547]	[1990.673, 1990.68]	[0.296, 0.312]	[0.019, 0.019]	[85.116, 88.296]	[46.786, 59.727]	[144.563, 146.826]	[-14.243, -14.078]	[21.296, 21.66]	[22.432, 23.417]	[1.141, 1.26]	[0.935, 1.0]
HIP14157	$SB2$	0.119	1,993.795	0.761	0.006	175.114	19.437	91.532	30.682	19.436	24.367	1.003	0.900
		[0.119, 0.119]	[1993.795, 1993.796]	[0.758, 0.763]	[0.006, 0.006]	[174.829, 175.564]	[19.092, 19.924]	[90.104, 92.968]	[30.573, 30.78]	[19.279, 19.604]	[24.145, 24.616]	[0.982, 1.018]	[0.891, 0.91]
	$SB1 + p(\pi)$	0.119	1,993.795	0.762	0.006	174.912	19.532	91.432	30.692	19.524	24.333	0.985	0.905
		[0.119, 0.119]	[1993.795, 1993.796]	[0.758, 0.765]	[0.006, 0.006]	[174.469, 175.445]	[19.092, 19.937]	[89.917, 92.89]	[30.581, 30.844]	[19.425, 19.64]	[24.138, 24.645]	[0.947, 1.009]	[0.891, 0.932]
	$SB1 + p(m_1\|\theta)$	0.119	1,993.795	0.761	0.006	174.952	19.445	91.272	30.705	20.064	24.379	0.882	0.957
		[0.119, 0.119]	[1993.795, 1993.796]	[0.758, 0.765]	[0.006, 0.006]	[174.476, 175.429]	[19.093, 19.925]	[90.098, 92.977]	[30.573, 30.837]	[19.798, 20.543]	[24.146, 24.639]	[0.81, 0.924]	[0.937, 0.998]
	$SB1 + p(\pi) + p(m_1\|\theta)$	0.119	1,993.795	0.762	0.006	175.002	19.527	91.349	30.703	19.610	24.505	0.954	0.925
		[0.119, 0.119]	[1993.795, 1993.796]	[0.758, 0.765]	[0.006, 0.006]	[174.532, 175.496]	[19.08, 19.94]	[89.958, 92.911]	[30.579, 30.841]	[19.47, 19.677]	[24.233, 24.722]	[0.932, 0.99]	[0.9, 0.939]
HIP20601	$SB2$	0.428	1,971.983	0.852	0.011	-157.740	340.812	103.636	42.097	16.710	25.677	0.984	0.752
		[0.428, 0.428]	[1971.983, 1971.983]	[0.849, 0.854]	[0.011, 0.012]	[-158.465, -156.891]	[339.649, 341.814]	[101.114, 105.151]	[41.94, 42.251]	[16.429, 16.991]	[25.178, 26.099]	[0.947, 1.013]	[0.742, 0.759]
	$SB1 + p(\pi)$	0.428	1,971.983	0.852	0.011	-157.370	340.619	102.796	42.311	17.347	25.712	0.849	0.805
		[0.428, 0.428]	[1971.983, 1971.983]	[0.849, 0.855]	[0.011, 0.012]	[-158.309, -156.275]	[339.625, 341.903]	[100.728, 105.558]	[41.92, 42.486]	[17.072, 17.577]	[25.184, 26.188]	[0.797, 0.924]	[0.768, 0.836]
	$SB1 + p(m_1\|\theta)$	0.428	1,971.983	0.852	0.011	-157.447	340.930	103.282	42.262	17.044	25.667	0.916	0.778
		[0.428, 0.428]	[1971.983, 1971.983]	[0.849, 0.855]	[0.011, 0.012]	[-158.335, -156.458]	[339.542, 341.811]	[100.685, 105.232]	[41.913, 42.476]	[16.746, 17.346]	[25.171, 26.113]	[0.873, 0.954]	[0.755, 0.797]
	$SB1 + p(\pi) + p(m_1\|\theta)$	0.428	1,971.983	0.852	0.011	-156.987	340.443	102.590	42.255	17.186	25.450	0.905	0.777
		[0.428, 0.428]	[1971.983, 1971.983]	[0.849, 0.855]	[0.011, 0.012]	[-158.211, -156.252]	[339.572, 341.98]	[100.697, 105.291]	[41.93, 42.498]	[17.031, 17.421]	[25.069, 25.976]	[0.865, 0.934]	[0.761, 0.806]
HIP117186	$SB2$	0.235	1,987.923	0.323	0.005	175.150	16.956	87.974	-19.783	8.470	53.043	1.722	0.816
		[0.235, 0.235]	[1987.922, 1987.924]	[0.32, 0.327]	[0.005, 0.005]	[174.106, 175.971]	[16.5, 17.415]	[87.683, 88.42]	[-19.903, -19.632]	[8.366, 8.58]	[51.939, 54.212]	[1.69, 1.765]	[0.799, 0.833]
	$SB1 + p(\pi)$	0.235	1,987.925	0.324	0.005	178.492	16.836	88.161	-20.233	8.223	54.017	1.859	0.799
		[0.235, 0.235]	[1987.922, 1987.926]	[0.318, 0.334]	[0.005, 0.005]	[-178.896, 180.0]	[16.451, 17.332]	[87.722, 88.457]	[-20.812, -19.847]	[8.131, 8.299]	[52.632, 55.526]	[1.747, 1.968]	[0.759, 0.841]
	$SB1 + p(m_1\|\theta)$	0.235	1,987.925	0.324	0.005	178.207	16.806	88.138	-20.397	8.961	53.847	1.340	0.932
		[0.235, 0.235]	[1987.922, 1987.927]	[0.317, 0.334]	[0.005, 0.005]	[-179.186, 179.995]	[16.406, 17.32]	[87.693, 88.451]	[-20.797, -19.828]	[8.735, 9.127]	[52.571, 55.49]	[1.251, 1.439]	[0.892, 0.98]
	$SB1 + p(\pi) + p(m_1\|\theta)$	0.235	1,987.923	0.335	0.005	175.004	16.925	88.014	-20.580	8.343	56.626	1.595	0.896
		[0.235, 0.235]	[1987.921, 1987.925]	[0.328, 0.344]	[0.005, 0.005]	[172.09, 178.279]	[16.46, 17.383]	[87.595, 88.392]	[-21.03, -20.051]	[8.271, 8.421]	[55.163, 57.942]	[1.53, 1.669]	[0.854, 0.933]
HD6840	$SB2$	7.455	1,995.494	0.744	0.082	-144.994	152.001	50.669	-9.526	16.201	29.379	1.244	0.908
		[7.45, 7.458]	[1995.49, 1995.502]	[0.741, 0.746]	[0.082, 0.083]	[-145.468, -144.197]	[151.314, 152.497]	[49.926, 51.606]	[-9.596, -9.477]	[15.952, 16.572]	[28.751, 29.857]	[1.176, 1.294]	[0.894, 0.925]
	$SB1 + p(\pi)$	7.454	1,995.495	0.744	0.083	-145.217	152.208	50.819	-9.481	16.435	29.390	1.179	0.934
		[7.45, 7.458]	[1995.488, 1995.5]	[0.741, 0.746]	[0.082, 0.083]	[-145.946, -144.548]	[151.574, 152.808]	[49.962, 51.529]	[-9.55, -9.418]	[16.247, 16.987]	[28.909, 29.962]	[1.029, 1.223]	[0.915, 1.0]
	$SB1 + p(m_1\|\theta)$	7.455	1,995.494	0.744	0.083	-145.236	152.228	51.032	-9.474	16.910	29.275	1.057	0.980
		[7.45, 7.458]	[1995.488, 1995.501]	[0.741, 0.746]	[0.082, 0.083]	[-145.91, -144.531]	[151.622, 152.848]	[50.093, 51.629]	[-9.553, -9.418]	[16.636, 17.138]	[28.853, 29.855]	[1.015, 1.098]	[0.965, 1.0]
	$SB1 + p(\pi) + p(m_1\|\theta)$	7.454	1,995.493	0.744	0.082	-145.403	152.306	50.849	-9.490	16.806	29.429	1.074	0.979
		[7.45, 7.458]	[1995.489, 1995.501]	[0.741, 0.746]	[0.082, 0.083]	[-145.955, -144.571]	[151.589, 152.832]	[50.006, 51.444]	[-9.553, -9.419]	[16.619, 17.041]	[28.983, 29.899]	[1.027, 1.102]	[0.963, 1.0]
HIP108917	$SB2$	2.245	1,968.750	0.463	0.072	90.343	273.135	67.582	-10.796	37.243	7.239	1.043	0.369
		[2.245, 2.246]	[1968.744, 1968.756]	[0.452, 0.469]	[0.071, 0.072]	[89.919, 90.644]	[272.559, 273.573]	[67.044, 67.983]	[-11.394, -10.035]	[34.622, 40.09]	[6.073, 8.181]	[0.807, 1.266]	[0.3, 0.425]
	$SB1 + p(\pi)$	2.245	1,968.752	0.459	0.072	90.308	273.234	67.377	-11.090	31.858	7.116	1.749	0.293
		[2.245, 2.246]	[1968.745, 1968.758]	[0.452, 0.468]	[0.071, 0.072]	[89.925, 90.659]	[272.691, 273.706]	[67.005, 67.944]	[-11.747, -10.181]	[30.99, 33.415]	[6.027, 8.186]	[1.473, 1.93]	[0.241, 0.361]
	$SB1 + p(m_1\|\theta)$	2.245	1,968.753	0.459	0.072	90.297	273.152	67.455	-10.889	32.096	6.936	1.722	0.286
		[2.245, 2.246]	[1968.745, 1968.758]	[0.452, 0.468]	[0.071, 0.072]	[89.944, 90.665]	[272.676, 273.707]	[66.985, 67.941]	[-11.743, -10.158]	[30.969, 32.948]	[6.107, 8.256]	[1.593, 1.893]	[0.242, 0.348]
	$SB1 + p(\pi) + p(m_1\|\theta)$	2.245	1,968.751	0.460	0.072	90.354	272.995	67.447	-10.901	31.982	7.217	1.721	0.300
		[2.245, 2.246]	[1968.746, 1968.758]	[0.452, 0.469]	[0.071, 0.072]	[89.936, 90.669]	[272.694, 273.697]	[67.004, 67.955]	[-11.76, -10.195]	[31.284, 32.814]	[6.003, 8.099]	[1.601, 1.855]	[0.24, 0.346]
HIP677	$SB2$	0.265	1,988.847	0.535	0.024	78.339	104.952	106.172	-12.465	32.796	10.500	3.697	0.525
		[0.265, 0.265]	[1988.842, 1988.851]	[0.47, 0.589]	[0.022, 0.025]	[73.74, 81.112]	[101.653, 108.823]	[102.45, 109.408]	[-14.043, -9.567]	[29.447, 35.93]	[9.113, 12.139]	[2.983, 4.461]	[0.449, 0.6]
	$SB1 + p(\pi)$	0.265	1,988.847	0.520	0.023	77.123	105.402	106.471	-10.481	33.732	10.922	3.031	0.583
		[0.264, 0.265]	[1988.843, 1988.852]	[0.459, 0.594]	[0.022, 0.025]	[73.047, 81.221]	[100.932, 108.246]	[102.514, 109.817]	[-13.177, -7.438]	[32.956, 34.333]	[9.367, 12.515]	[2.408, 4.17]	[0.453, 0.718]
	$SB1 + p(m_1\|\theta)$	0.265	1,988.848	0.523	0.024	78.617	105.264	106.524	-10.154	34.209	10.857	3.063	0.591
		[0.264, 0.265]	[1988.843, 1988.852]	[0.462, 0.596]	[0.022, 0.026]	[72.973, 81.262]	[100.708, 108.133]	[102.46, 109.76]	[-13.267, -7.524]	[30.839, 38.681]	[9.255, 12.434]	[2.032, 3.879]	[0.477, 0.734]
	$SB1 + p(\pi) + p(m_1\|\theta)$	0.265	1,988.848	0.530	0.024	77.903	103.764	105.992	-9.832	33.724	10.897	3.223	0.581
		[0.264, 0.265]	[1988.843, 1988.851]	[0.461, 0.582]	[0.022, 0.025]	[73.038, 80.879]	[100.862, 108.126]	[102.974, 109.822]	[-13.25, -7.476]	[32.968, 34.315]	[9.693, 12.377]	[2.548, 3.737]	[0.478, 0.704]

(a) Results reported by other authors. (b) Results obtained in this work.

information sources compared with the uncertainty quality of the other observations. For example, the binary system YSC132AaAb shows the impact of the additional information on the posterior distribution of the identifiable orbital parameters a and i. In contrast, the binary system HIP108917 has a null impact on the marginal distributions of the identifiable orbital parameters when adding additional information.

The projections of the estimated posterior distribution and the MAP estimates on the observation space presents no appreciable differences between all the studied cases ($SB2$, $SB1 + p(\pi)$, $SB1 + p(m_1|\theta)$ and $SB1 + p(\pi) + p(m_1|\theta)$). The only significant difference is observed in the orbit of the YSC132AaAb system where the mixed case $SB1 + p(\pi) + p(m_1|\theta)$ presents the lower uncertainty, even lower than the full-information case $SB2$, but at the cost of a slight worst fitting of some of the orbital observations.

Focusing on the mass ratio q estimation, the posterior distribution of the $SB1 + p(\pi)$, $SB1 + p(m_1|\theta)$ and $SB1 + p(\pi) + p(m_1|\theta)$ cases that offers the highest similarity to the full information scenario $SB2$, according to the Kullback-Leibler divergence measure, is the $SB1 + p(\pi)$ case, followed by the mixed case $SB1 + p(\pi) + p(m_1|\theta)$ and the $SB1 + p(m_1|\theta)$ case. However, the lowest mean absolute error between the MAP estimates is reached by the $SB1 + p(\pi) + p(m_1|\theta)$ case (4.92%), followed by the $SB1 + p(\pi)$ (7.41%) and $SB1 + p(m_1|\theta)$ (7.44%) cases, as well as the high posterior density interval range as discussed

Table 5.3: Comparison of the marginal posterior distribution of the mass ratio q between the $SB1$ with additional information cases and the full-information $SB2$ case.

	MAP estimate absolute error [%]			HPDI length			KL-Divergence		
System	$SB1+p(\pi)$	$SB1+p(m_1\|\theta)$	$SB1+p(\pi)+p(m_1\|\theta)$	$SB1+p(\pi)$	$SB1+p(m_1\|\theta)$	$SB1+p(\pi)+p(m_1\|\theta)$	$SB1+p(\pi)$	$SB1+p(m_1\|\theta)$	$SB1+p(\pi)+p(m_1\|\theta)$
CHR111	6.05	5.88	5.02	0.125	0.072	0.069	3.804	5.778	5.647
YSC132AaAb	29.67	11.49	1.73	0.476	0.273	0.065	3.533	3.014	2.146
HIP14157	0.54	5.78	2.55	0.041	0.061	0.039	0.826	7.038	2.346
HIP20601	5.36	2.62	2.58	0.068	0.042	0.045	4.723	3.471	4.76
HIP117186	1.66	11.67	7.98	0.082	0.088	0.079	0.772	9.201	7.285
HD6840	2.6	7.21	7.03	0.085	0.035	0.037	2.507	8.844	8.821
HIP108917	7.6	8.28	6.91	0.12	0.106	0.106	2.064	2.856	2.932
HIP677	5.81	6.57	5.58	0.265	0.257	0.226	0.485	0.834	0.728
mean	7.41	7.44	4.92	0.158	0.117	0.083	2.339	5.129	4.333
std	8.72	2.83	2.22	0.136	0.088	0.058	1.481	2.857	2.609

previously. The obtained results indicate that the closest estimation to the full-information scenario is obtained by incorporating prior information on the parallax π. However, the most robust point estimates are obtained by incorporating both sources of additional information, correcting the estimation through $p(m_1|\theta)$ when $p(\pi)$ is incorrect. This is illustrated in the YSC132AaAb system, at the cost of a slight increase in the average estimation error. According to all the comparison metrics, the lowest performance is achieved by the $SB1 + p(m_1|\theta)$ case. This is attributed to the fact that the additional information on m_1 comes from a rough and older empirical rule (that relates the mass with the object's spectral type). In contrast, the additional information on π was recently updated with a high precision instrument.

5.3 Application to Unresolved Single-lined Spectroscopic Binaries with a Visual Orbit

The inference of twelve single-lined spectroscopic binaries with a visual orbit using the approaches $SB1$, $SB1+p(\pi)$, $SB1+p(m_1|\theta)$ and $SB1+p(\pi)+p(m_1|\theta)$ is performed. The estimates and their uncertainties are compared in the parameters space through visualization of the posterior marginal distributions, as well as in the observations space, by projecting 1000 randomly selected samples of the posterior distribution on the observation space, drawing trajectories from the time of the first observation time t_0 to the first completion of the orbit $t_0 + P$. The MAP estimators and their respective 95% high posterior density interval estimates are summarized in Table 5.4.

To avoid redundancy, the analysis is focused only on three of the twelve systems resolved: HIP3504, HIP99675, and HIP109951.

5.3.1 HIP3504

The available data consists of few and imprecise astrometric observations, which only cover three distant points of the orbit. Very few and imprecise radial velocity observations of the primary object are also available. The observations and it uncertainties are visualized in Figure 5.14.

Figure 5.13 shows the posterior distributions of the three cases with additional information as well as the reference case without additional information. These distributions are considerably different in their MAP values and dispersion compared to the $SB1$ case. The most significant difference is presented in the period P, where the posterior distribution of the $SB1$ case presents an extremely high uncertainty with a MAP of 31.15 $[yr]$, while the posterior distribution of all the cases with additional information presents a much less uncertainty with a MAP of 2.48 $[yr]$. The uncertainty of the orbital parameters T, e, ω, i is higher in the cases with additional information than the $SB1$ case, while the uncertainty of the orbital parameters $P, a, \Omega, V_0, f/\pi$ is lower than the $SB1$ case. The marginal posterior distribution of

the mass ratio q shows that the uncertainty of the $SB1 + p(\pi)$ case is the highest, while the uncertainty of $SB1 + p(m_1|\theta)$ and $SB1 + p(\pi) + p(m_1|\theta)$ are almost identical.

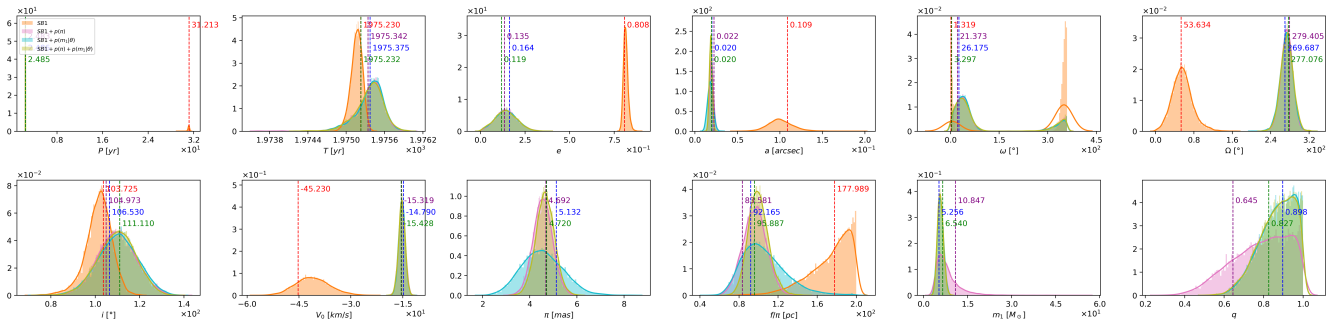

Figure 5.13: Marginal posterior distributions and MAP estimates of the orbital parameters of the HIP3504 binary system for the $SB1$, $SB1 + p(\pi)$, $SB1 + p(m_1|\theta)$ and $SB1 + p(\pi) + p(m_1|\theta)$ cases.

The projection of the posterior distributions on the observation space is presented in Figure 5.14. The MAP estimates in the orbit space of the cases with additional information are almost identical but very different than the $SB1$ case. This is consistent with the important difference of the posterior distributions of the period P. The orbital uncertainty of the $SB1$ case is very high in this analysis, projecting a dense cloud of possible orbits. In contrast, the orbital uncertainty with additional information $SB1 + p(\pi)$, $SB1 + p(m_1|\theta)$ and $SB1 + p(\pi) + p(m_1|\theta)$ are confined to a delimited ring of possible orbits. All those solutions are less uncertain than the $SB1$. The projected uncertainty in the orbit space of the $SB1 + p(\pi)$ and $SB1 + p(m_1|\theta)$ cases are very similar, while the mixed case $SB1 + p(\pi) + p(m_1|\theta)$ presents a slight lower uncertainty. Similarly, the uncertainty in the radial velocity projection of the $SB1$ case is reduced in the times with observations but extremely higher in the times far from the observations. In contrast, the radial velocity uncertainty of the cases with additional information presents no variation between the times with and without the observations, preserving the uncertainty than the $SB1$ case in the times with observations. No significant differences in the radial velocity curve's uncertainty are observed between the cases with additional information. Another relevant difference between the cases with and without additional information is that the period of the radial velocity curve in the $SB1 + p(\pi)$, $SB1 + p(m_1|\theta)$, $SB1 + p(\pi) + p(m_1|\theta)$ cases presents a visible lower period than the $SB1$ case. The period of the first ones is visible in the figure's time window, while the period of the last case is much higher and is not visible in the figure's time window. This behavior coincides with the dramatic differences between the marginal posterior distributions of the period in the cases with and without additional information presented in Figure 5.13.

The tremendous differences observed between the cases with and without additional information show the crucial relevance of the prior information in determining some orbital parameters, the orbit, and the radial velocity curves of binary systems that have not enough observations available.

5.3.2 HIP99675

This object's data consists of few and highly imprecise astrometric observations in two extreme zones of the orbit and abundant and highly precise radial velocity observations of the primary object. The observations and it uncertainties are visualized in Figure 5.16.

Figure 5.15 shows that the posterior distributions of all the cases are identical for almost all the orbital parameters, with the exception of $a, i, \pi, f/\pi, m_1, q$, where the most significant differences are observed on the parameters π, m_1, q. The marginal posterior distribution π of the mixed case $SB1 + p(\pi) + p(m_1|\theta)$

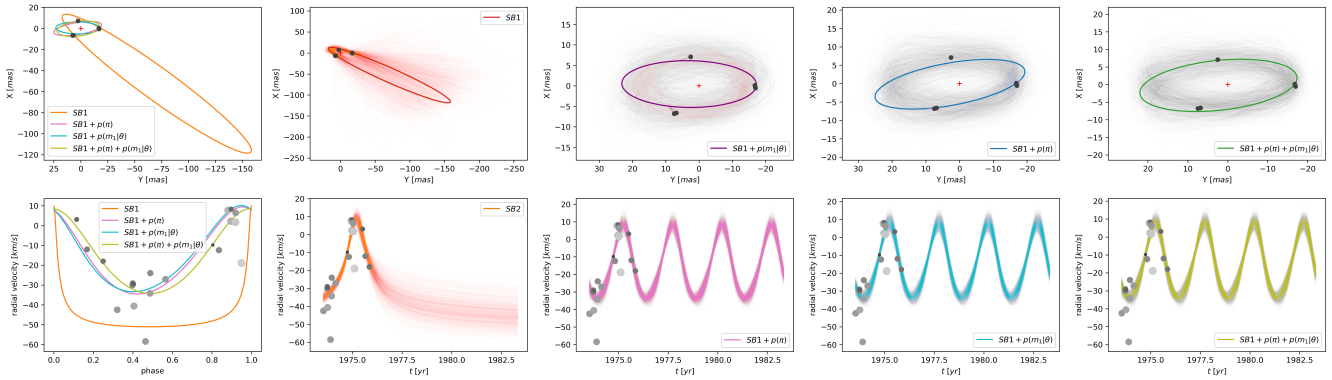

Figure 5.14: Estimated orbits and radial velocity curves of the HIP3504 binary system. First column: MAP point estimate projection of the posterior distribution for the $SB1$, $SB1 + p(\pi)$, $SB1 + p(m_1|\theta)$ and $SB1 + p(\pi) + p(m_1|\theta)$ cases. Second column: Projected posterior distribution of the $SB1$ case. Third column: Projected posterior distribution of the $SB1 + p(\pi)$ case. Fourth column: Projected posterior distribution of the $SB1 + p(m_1|\theta)$ case. Fifth column: Projected posterior distribution of the $SB1 + p(\pi) + p(m_1|\theta)$ case.

is in between the $SB1 + p(\pi)$ and $SB1 + p(m_1|\theta)$ cases presenting the lowest uncertainty, where its MAP estimation is almost equidistant to the other cases. The marginal posterior distribution m_1 of the mixed case is almost equal to the obtained with the $SB1 + p(m_1|\theta)$ case, but the MAP estimate is very different than the one obtained with the $SB1 + p(\pi)$ case, which has the lowest uncertainty. The posterior distributions of the mass ratio q of the $SB1 + p(\pi) + p(m_1|\theta)$ and $SB1 + p(m_1|\theta)$ are very similar, but quite different to the $SB1 + p(\pi)$ case. The $SB1 + p(\pi)$ offers the posterior distribution with the less uncertainty, followed by the $SB1 + p(\pi) + p(m_1|\theta)$ and $SB1 + p(m_1|\theta)$ cases.

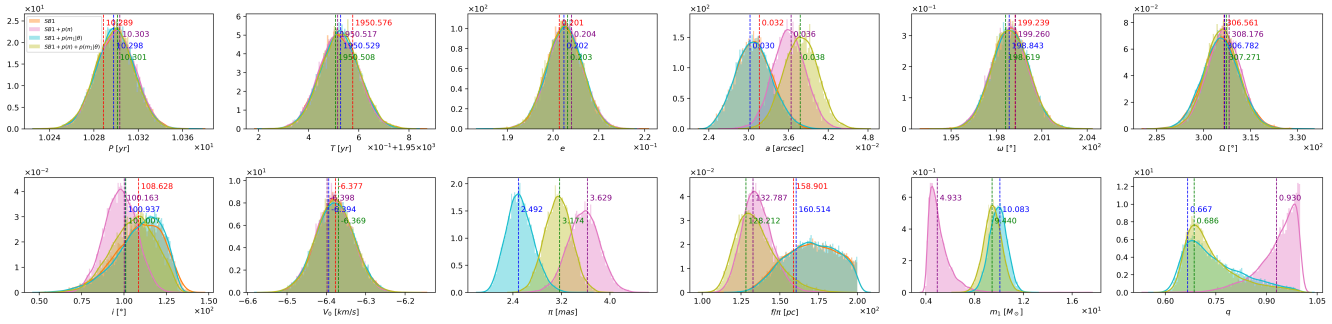

Figure 5.15: Marginal posterior distributions and MAP estimates of the orbital parameters of the HIP99675 binary system for the $SB1$, $SB1 + p(\pi)$, $SB1 + p(m_1|\theta)$ and $SB1 + p(\pi) + p(m_1|\theta)$ cases.

The obtained posterior distributions in the observation space are presented in Figure 5.16. The MAP estimates in the orbit space of all the cases are significantly different, and none of them fit the positional observations adequately. The orbit uncertainty obtained from those distributions is extremely high in all the cases, expressed as a dense cloud of possible orbits. The $SB1$ case presents the higher uncertainty in the orbit space, followed by the $SB1 + p(m_1|\theta)$ and $SB1 + p(\pi) + p(m_1|\theta)$ cases, being the $SB1 + p(\pi)$ the less uncertain case. The MAP estimates in the radial velocity space of all the cases are identical, with almost no projected uncertainty.

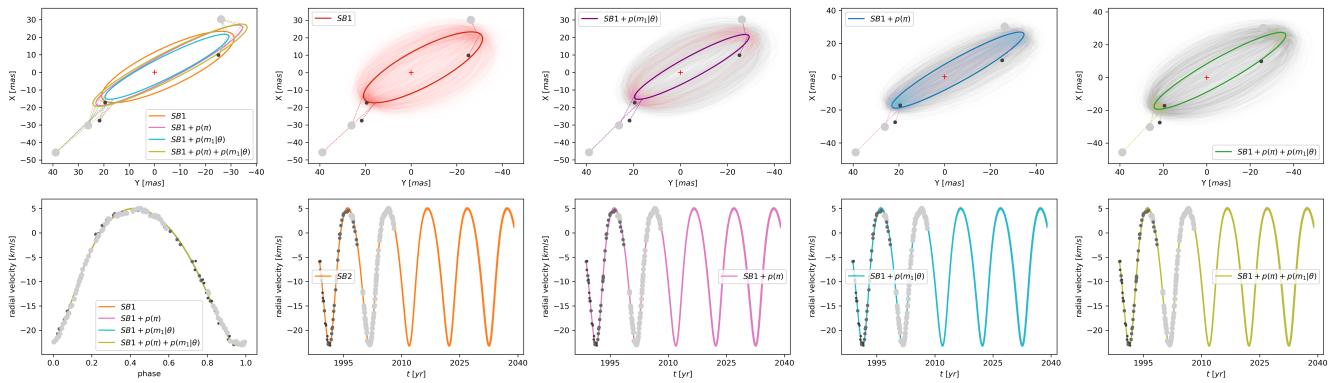

Figure 5.16: Estimated orbits and radial velocity curves of the HIP99675 binary system. First column: MAP point estimate projection of the posterior distribution for the $SB1$, $SB1 + p(\pi)$, $SB1 + p(m_1|\theta)$ and $SB1 + p(\pi) + p(m_1|\theta)$ cases. Second column: Projected posterior distribution of the $SB1$ case. Third column: Projected posterior distribution of the $SB1 + p(\pi)$ case. Fourth column: Projected posterior distribution of the $SB1 + p(m_1|\theta)$ case. Fifth column: Projected posterior distribution of the $SB1 + p(\pi) + p(m_1|\theta)$ case.

5.3.3 HIP109951

The available data consists of abundant and precise astrometric observations covering less than a half of the orbit with only one less precise observation in the other half. There are abundant and highly imprecise radial velocity observations of the primary object, mostly concentrated in a small segment of the phase near the periastron. The observations and it uncertainties are visualized in Figure 5.18.

Figure 5.17 shows the posterior distributions of all the cases are similar in almost all the orbital parameters, with the exception of the parameters $i, \pi, f/\pi, m_1, q$. The most significant differences are observed in the parameters π, m_1, q. The marginal posterior distribution π of the mixed case $SB1 + p(\pi) + p(m_1|\theta)$ is in between the $SB1 + p(\pi)$ and $SB1 + p(m_1|\theta)$ cases presenting the lowest uncertainty, where its MAP estimation is almost equidistant to the other cases. The marginal posterior distribution m_1 of the mixed case is almost equal to the $SB1+p(m_1|\theta)$ case, but very different in its MAP estimates to the $SB1 + p(\pi)$ case, which has the lowest uncertainty. The posterior distribution of the mass ratio q of the $SB1+p(\pi)$ and $SB1+p(m_1|\theta)$ cases are very similar, but they are quite different to the $SB1+p(\pi)+p(m_1|\theta)$ case, where the mixed case $SB1 + p(\pi) + p(m_1|\theta)$ presents the lowest uncertainty, distantly followed by the $SB1 + p(m_1|\theta)$ and $SB1 + p(\pi)$ cases. In all the cases, the uncertainty of the period P is very high. This is reflected in a wide dispersion of the corresponding marginal posterior distributions.

The obtained posterior distributions in the observation space are presented in Figure 5.18. The MAP estimates in the orbit and radial velocity spaces show slight but almost negligible differences between all the cases. The orbit uncertainty (obtained from the posterior distribution) is high in the orbit segment with no observations but almost null in the orbit segment with observations. The $SB1$ case presents the highest uncertainty in the orbit space, followed by the $SB1+p(m_1|\theta)$ and $SB1+p(\pi)$ cases, and concluding with the $SB1 + p(\pi) + p(m_1|\theta)$ case. This last case has considerably lower uncertainty when compared to all the other cases. The uncertainty in the radial velocity is low in the zones with observations but high in the zones with no observations. Unlike all the other studied systems, the uncertainty in the radial velocity increases with time. This behavior is attributed to the high uncertainty of the orbital period P, as all the positional and radial velocity observations are constrained to the same period. This fact does not allow determining an accurate estimation of the orbital period. The lack of a good determination of the period causes that the possible radial velocity trajectories to get out of phase, increasing the uncertainty

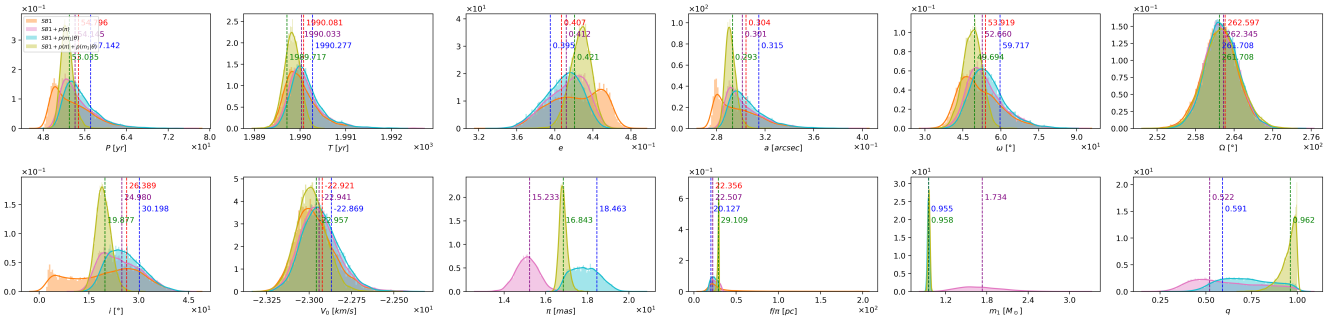

Figure 5.17: Marginal posterior distributions and MAP estimates of the orbital parameters of the HIP109951 binary system for the $SB1$, $SB1 + p(\pi)$, $SB1 + p(m_1|\theta)$ and $SB1 + p(\pi) + p(m_1|\theta)$ cases.

along the time. Similarly to the results obtained in the orbit space, the $SB1$ case presents the higher posterior uncertainty in the radial velocity space, followed by the results obtained in the $SB1 + p(m_1|\theta)$ and $SB1 + p(\pi)$ cases. In this scenario, the mixed case $SB1 + p(\pi) + p(m_1|\theta)$ has the lowest uncertainty, being considerably lower compared to the uncertainties in the radial velocity space of all the other cases.

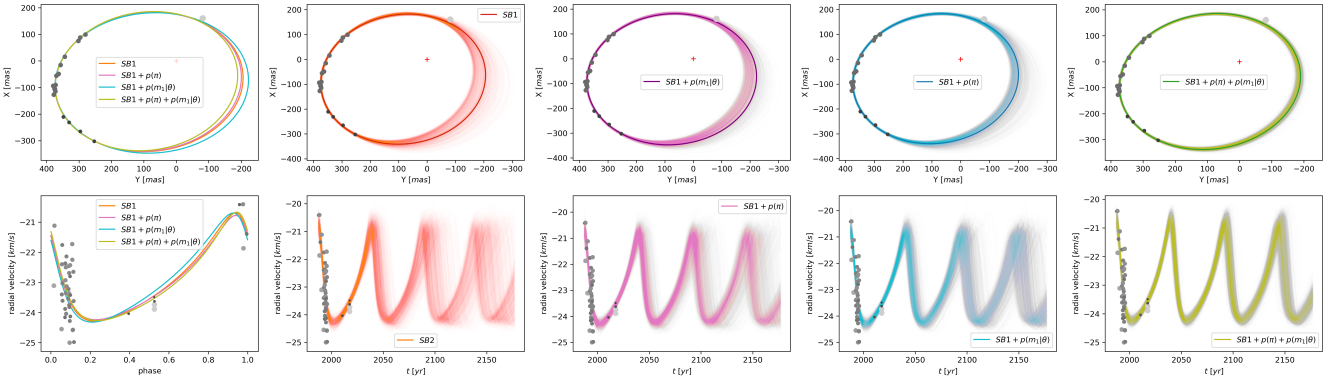

Figure 5.18: Estimated orbits and radial velocity curves of the HIP109951 binary system. First column: MAP point estimate projection of the posterior distribution for the $SB1$, $SB1 + p(\pi)$, $SB1 + p(m_1|\theta)$ and $SB1 + p(\pi) + p(m_1|\theta)$ cases. Second column: Projected posterior distribution of the $SB1$ case. Third column: Projected posterior distribution of the $SB1 + p(\pi)$ case. Fourth column: Projected posterior distribution of the $SB1 + p(m_1|\theta)$ case. Fifth column: Projected posterior distribution of the $SB1 + p(\pi) + p(m_1|\theta)$ case.

5.4 Concluding Remarks

The experimental results show the efficacy of the proposed Bayesian inference methodology to characterize and visualize the uncertainty of single-lined visual-spectroscopic binary systems.

Interesting results are obtained in some of the binary systems evaluated. For example, the HIP3504 shows an extremely high uncertainty on its orbit and its radial velocity curve due to the few observations available. However, incorporating the priors radically reduces the uncertainty of the estimates, reaching completely different solutions to the ones obtained without the prior incorporation. This is reflected in smaller orbits and a radial velocity curve with a lower period. The results obtained in the system HIP99675 show that the prior incorporation's effect is mostly concentrated in the trio of soft-identifiable parameters

π, m_1 and q, slightly affecting the other orbital parameters. This behavior expresses the robustness of the estimation when abundant and precise observations are available. On the soft-identifiable parameters, we observe that the posterior distributions of the $SB1 + p(m_1|\theta)$ and the mixed case $SB1 + p(\pi) + p(m_1|\theta)$ are similar, but very different to the case $SB1 + p(\pi)$. This behavior can be interpreted as the prior information of the primary object mass $p(m_1|\theta)$ is more reliable to the prior on the parallax $p(\pi)$ for this system. Finally, the results obtained in the system HIP109951 reflect the value of providing a good characterization of the uncertainty. In this case, we observe that the uncertainty in the half of the orbit with no observations is considerably higher than the other half, but most importantly, the uncertainty of the radial velocity curve visibly increases with time. This behavior is attributed to the high uncertainty of the system's period, making the trajectories of the radial velocity get out of phase with time. On the prior information incorporated, we observe that the posterior distribution of the cases $SB1 + p(\pi)$ and $p(m_1|\theta)$ are very similar to the one obtained in the $SB1$ case. However, the mixed case $SB1+p(\pi)+p(m_1|\theta)$ shows a visible uncertainty reduction in all the orbital parameters with a very similar MAP value. Regarding the posterior distribution of the mass ratio q, we observe that the cases $SB1 + p(\pi)$ and $p(m_1|\theta)$ provide a poor estimation, reflected in a very high uncertainty. However, the mixed case $SB1 + p(\pi) + p(m_1|\theta)$ provides a much more constrained posterior distribution than the other cases, considerably differing in their MAP estimates. It is important to note that, as the posterior distribution on the mass ratio q of the $SB1 + p(\pi)$ and $p(m_1|\theta)$ cases presents a very high uncertainty (with the shape of uniform distributions), it corresponding MAP estimate are not very reliable, but the mixed case $SB1 + p(\pi) + p(m_1|\theta)$ it does. This last result shows the relevance of incorporating both sources of prior information to obtain more accurate estimates.

Table 5.4: MAP estimates and 95% HDPIs of the marginal posterior distributions of the orbital parameters incorporating priors of m_1 and π of unresolved $SB1$ binary systems, considering the $SB1$, $SB1 + p(\pi)$, $SB1 + p(m_1|\theta)$, and $SB1 + p(\pi) + p(m_1|\theta)$ cases.

System	Case	P [yr]	T [yr]	e	a ["]	Ω [°]	ω [°]	i [°]	V_0 [km/s]	π [mas]	f/π [pc]	m_1 [$M_\odot$]	q
DSG7AaAb	$SB1$	0.619	1,987.813	0.308	0.016	20.562	164.346	121.321	-51.301	34.312	-	-	-
		[0.618, 0.619]	[1987.806, 1987.82]	[0.294, 0.328]	[0.014, 0.018]	[16.955, 24.04]	[154.169, 170.929]	[112.756, 129.763]	[-51.626, -51.017]	[28.861, 42.885]	-	-	-
	$SB1 + p(\pi)$	0.619	1,987.813	0.307	0.017	20.168	163.589	119.719	-51.282	14.711	32.186	2.013	0.899
		[0.618, 0.619]	[1987.807, 1987.82]	[0.292, 0.326]	[0.015, 0.019]	[17.231, 24.235]	[153.245, 169.21]	[111.265, 125.124]	[-51.613, -51.014]	[11.902, 17.697]	[27.04, 37.262]	[1.301, 3.435]	[0.694, 1.0]
	$SB1 + p(m_1\|\theta)$	0.619	1,987.815	0.309	0.017	21.384	160.222	117.032	-51.347	15.443	31.238	1.715	0.932
		[0.618, 0.619]	[1987.807, 1987.82]	[0.293, 0.328]	[0.014, 0.018]	[17.416, 24.365]	[152.776, 168.884]	[111.111, 123.22]	[-51.626, -51.02]	[13.11, 17.639]	[27.642, 37.226]	[1.339, 2.034]	[0.884, 1.0]
	$SB1 + p(\pi) + p(m_1\|\theta)$	0.619	1,987.816	0.312	0.017	21.639	159.469	114.537	-51.325	16.572	29.431	1.532	0.952
		[0.618, 0.619]	[1987.808, 1987.821]	[0.292, 0.326]	[0.015, 0.019]	[17.61, 24.454]	[152.635, 168.411]	[110.206, 121.517]	[-51.627, -51.024]	[14.033, 18.374]	[26.549, 34.804]	[1.303, 1.898]	[0.893, 1.0]
HIP171	$SB1$	26.607	1,882.987	0.357	0.817	-80.847	109.296	49.751	-36.288	5.654	-	-	-
		[26.589, 26.628]	[1882.89, 1883.063]	[0.354, 0.36]	[0.814, 0.82]	[-81.216, -80.591]	[108.951, 109.669]	[49.48, 50.062]	[-36.325, -36.249]	[5.564, 5.757]	-	-	-
	$SB1 + p(\pi)$	26.602	1,883.001	0.357	0.818	-80.922	109.322	49.835	-36.291	79.079	5.652	0.864	0.808
		[26.589, 26.628]	[1882.893, 1883.064]	[0.355, 0.36]	[0.814, 0.82]	[-81.213, -80.599]	[108.975, 109.685]	[49.482, 50.061]	[-36.325, -36.249]	[78.013, 80.163]	[5.569, 5.76]	[0.811, 0.91]	[0.779, 0.843]
	$SB1 + p(m_1\|\theta)$	26.610	1,882.968	0.357	0.817	-80.930	109.336	49.752	-36.293	77.207	5.650	0.944	0.774
		[26.589, 26.627]	[1882.891, 1883.062]	[0.354, 0.36]	[0.814, 0.82]	[-81.207, -80.585]	[108.956, 109.668]	[49.491, 50.064]	[-36.325, -36.251]	[76.208, 77.7]	[5.569, 5.761]	[0.926, 0.984]	[0.753, 0.793]
	$SB1 + p(\pi) + p(m_1\|\theta)$	26.601	1,883.006	0.358	0.819	-80.928	109.294	49.947	-36.297	77.790	5.610	0.929	0.774
		[26.587, 26.626]	[1882.897, 1883.068]	[0.355, 0.36]	[0.816, 0.822]	[-81.259, -80.633]	[108.981, 109.707]	[49.593, 50.152]	[-36.333, -36.257]	[77.008, 78.305]	[5.528, 5.713]	[0.907, 0.96]	[0.755, 0.797]
HIP3504	$SB1$	31.199	1,975.236	0.809	0.105	1.898	57.719	102.241	-44.998	180.352	-	-	-
		[30.878, 31.406]	[1974.988, 1975.333]	[0.79, 0.842]	[0.073, 0.133]	[-24.568, 10.92]	[14.701, 98.393]	[89.906, 113.391]	[-50.265, -31.248]	[128.634, 199.98]	-	-	-
	$SB1 + p(\pi)$	2.485	1,975.308	0.153	0.020	15.070	273.213	104.978	-15.102	4.603	93.073	7.434	0.750
		[2.471, 2.498]	[1974.915, 1975.75]	[0.029, 0.261]	[0.016, 0.025]	[-44.073, 83.271]	[248.11, 299.545]	[91.98, 125.174]	[-16.976, -13.357]	[3.807, 5.436]	[70.569, 117.562]	[2.857, 16.251]	[0.52, 1.0]
	$SB1 + p(m_1\|\theta)$	2.484	1,975.434	0.147	0.021	34.461	271.723	108.288	-14.685	5.349	83.650	5.689	0.810
		[2.471, 2.498]	[1974.903, 1975.78]	[0.025, 0.261]	[0.013, 0.026]	[-48.982, 84.83]	[247.578, 299.215]	[92.736, 125.822]	[-17.063, -13.458]	[3.232, 6.598]	[63.383, 136.914]	[3.788, 7.761]	[0.698, 1.0]
	$SB1 + p(\pi) + p(m_1\|\theta)$	2.486	1,975.371	0.133	0.019	25.676	268.836	111.311	-15.069	4.649	96.647	5.912	0.816
		[2.471, 2.497]	[1974.903, 1975.794]	[0.012, 0.254]	[0.016, 0.022]	[-44.991, 89.803]	[248.867, 299.28]	[92.33, 124.997]	[-17.054, -13.455]	[3.985, 5.492]	[79.424, 118.919]	[3.831, 7.687]	[0.696, 1.0]
HIP6564	$SB1$	16.082	1,924.566	0.934	0.197	169.555	208.995	120.905	99.037	28.652	-	-	-
		[16.07, 16.104]	[1924.499, 1924.602]	[0.926, 0.94]	[0.195, 0.198]	[167.549, 172.737]	[207.729, 210.742]	[118.652, 122.969]	[92.996, 105.86]	[24.411, 34.134]	-	-	-
	$SB1 + p(\pi)$	16.084	1,924.558	0.941	0.196	168.334	208.062	122.587	91.730	20.839	23.186	1.663	0.935
		[16.065, 16.096]	[1924.523, 1924.616]	[0.936, 0.945]	[0.195, 0.198]	[165.541, 170.235]	[206.383, 209.228]	[120.515, 124.465]	[89.08, 94.431]	[19.848, 22.102]	[21.443, 24.79]	[1.364, 1.939]	[0.85, 1.0]
	$SB1 + p(m_1\|\theta)$	16.077	1,924.578	0.939	0.196	168.461	208.256	121.999	93.101	19.835	24.283	1.945	0.929
		[16.065, 16.098]	[1924.516, 1924.612]	[0.933, 0.944]	[0.195, 0.198]	[166.001, 171.036]	[206.683, 209.712]	[119.964, 124.115]	[89.723, 98.066]	[17.447, 21.821]	[21.766, 27.746]	[1.379, 2.792]	[0.837, 1.0]
	$SB1 + p(\pi) + p(m_1\|\theta)$	16.079	1,924.574	0.941	0.196	167.476	207.638	122.240	91.327	21.333	22.826	1.548	0.949
		[16.065, 16.095]	[1924.524, 1924.616]	[0.936, 0.945]	[0.195, 0.198]	[165.257, 170.025]	[206.272, 209.15]	[120.555, 124.59]	[88.976, 94.209]	[19.961, 22.187]	[21.435, 24.661]	[1.366, 1.917]	[0.854, 1.0]
HIP7918	$SB1$	19.731	1,996.964	0.439	0.607	-158.428	208.520	97.638	3.319	2.932	-	-	-
		[19.701, 19.76]	[1996.911, 1997.005]	[0.434, 0.444]	[0.603, 0.613]	[-159.474, -157.442]	[208.046, 208.815]	[97.04, 98.016]	[3.306, 3.332]	[2.893, 2.964]	-	-	-
	$SB1 + p(\pi)$	19.736	1,996.966	0.439	0.608	-158.395	208.375	97.473	3.320	76.457	2.940	0.998	0.290
		[19.702, 19.761]	[1996.91, 1997.005]	[0.434, 0.444]	[0.603, 0.613]	[-159.487, -157.437]	[208.04, 208.806]	[97.044, 98.013]	[3.305, 3.331]	[76.104, 76.95]	[2.894, 2.965]	[0.967, 1.031]	[0.284, 0.294]
	$SB1 + p(m_1\|\theta)$	19.733	1,996.954	0.439	0.608	-158.593	208.326	97.485	3.321	76.113	2.931	1.016	0.287
		[19.7, 19.76]	[1996.914, 1997.008]	[0.434, 0.444]	[0.603, 0.613]	[-159.498, -157.464]	[208.028, 208.81]	[97.031, 98.013]	[3.305, 3.331]	[74.781, 77.181]	[2.895, 2.967]	[0.978, 1.069]	[0.281, 0.292]
	$SB1 + p(\pi) + p(m_1\|\theta)$	19.736	1,996.974	0.439	0.610	-158.210	208.420	97.530	3.318	76.439	2.917	1.013	0.287
		[19.703, 19.761]	[1996.914, 1997.009]	[0.434, 0.443]	[0.604, 0.613]	[-159.462, -157.405]	[208.03, 208.797]	[97.012, 97.982]	[3.306, 3.331]	[76.055, 76.864]	[2.891, 2.956]	[0.98, 1.034]	[0.283, 0.292]
HIP65982	$SB1$	3.232	1,984.708	0.519	0.099	6.536	302.129	24.603	-51.178	5.901	-	-	-
		[3.229, 3.237]	[1984.666, 1984.734]	[0.503, 0.537]	[0.096, 0.101]	[-0.428, 12.052]	[297.343, 308.595]	[16.08, 29.728]	[-51.298, -51.028]	[4.506, 8.741]	-	-	-
	$SB1 + p(\pi)$	3.232	1,984.704	0.517	0.099	6.389	302.747	25.116	-51.155	37.519	5.566	1.397	0.264
		[3.229, 3.237]	[1984.667, 1984.733]	[0.503, 0.536]	[0.096, 0.101]	[-0.369, 12.046]	[297.143, 308.285]	[16.281, 29.594]	[-51.296, -51.021]	[36.603, 39.054]	[4.54, 8.769]	[1.027, 1.562]	[0.198, 0.486]
	$SB1 + p(m_1\|\theta)$	3.232	1,984.699	0.519	0.099	5.493	303.220	24.894	-51.183	42.272	5.747	0.921	0.321
		[3.229, 3.236]	[1984.665, 1984.732]	[0.503, 0.537]	[0.096, 0.101]	[-0.763, 11.87]	[297.351, 308.72]	[15.84, 29.851]	[-51.3, -51.025]	[39.346, 44.395]	[4.365, 8.733]	[0.873, 0.95]	[0.239, 0.527]
	$SB1 + p(\pi) + p(m_1\|\theta)$	3.233	1,984.700	0.535	0.096	6.149	302.863	14.636	-51.166	38.329	10.059	0.924	0.627
		[3.229, 3.237]	[1984.672, 1984.736]	[0.523, 0.547]	[0.095, 0.097]	[0.401, 12.521]	[296.444, 308.151]	[11.343, 17.272]	[-51.301, -51.028]	[36.84, 39.36]	[8.526, 12.468]	[0.875, 0.955]	[0.498, 0.851]
HIP69962	$SB1$	18.818	1,984.953	0.305	0.350	72.448	285.221	107.667	11.571	7.199	-	-	-
		[18.7, 18.924]	[1984.786, 1985.072]	[0.29, 0.313]	[0.345, 0.352]	[70.498, 73.54]	[284.721, 285.867]	[107.304, 108.436]	[11.384, 11.737]	[6.617, 7.678]	-	-	-
	$SB1 + p(\pi)$	18.767	1,984.971	0.303	0.348	71.781	285.351	107.942	11.579	44.139	7.158	0.949	0.462
		[18.705, 18.927]	[1984.789, 1985.07]	[0.29, 0.313]	[0.345, 0.352]	[70.466, 73.535]	[284.734, 285.886]	[107.311, 108.442]	[11.378, 11.737]	[42.918, 45.902]	[6.591, 7.683]	[0.819, 1.056]	[0.409, 0.521]
	$SB1 + p(m_1\|\theta)$	18.797	1,984.931	0.302	0.349	71.887	285.489	107.837	11.571	49.275	7.110	0.652	0.539
		[18.705, 18.926]	[1984.786, 1985.063]	[0.29, 0.313]	[0.345, 0.352]	[70.53, 73.537]	[284.718, 285.883]	[107.331, 108.45]	[11.38, 11.738]	[47.659, 51.386]	[6.597, 7.671]	[0.566, 0.724]	[0.486, 0.608]
	$SB1 + p(\pi) + p(m_1\|\theta)$	18.813	1,984.902	0.299	0.348	71.599	285.228	108.026	11.556	46.853	7.273	0.761	0.517
		[18.716, 18.936]	[1984.759, 1985.033]	[0.288, 0.31]	[0.344, 0.351]	[70.174, 73.252]	[284.75, 285.911]	[107.439, 108.564]	[11.392, 11.755]	[45.673, 47.729]	[6.952, 8.024]	[0.697, 0.813]	[0.481, 0.596]
HIP78401	$SB1$	10.808	1,979.079	0.939	0.099	-179.128	353.834	31.992	-6.466	56.608	-	-	-
		[10.807, 10.809]	[1979.075, 1979.082]	[0.938, 0.94]	[0.099, 0.099]	[-179.999, 179.87]	[353.096, 355.043]	[31.271, 33.056]	[-7.175, -5.705]	[54.489, 58.405]	-	-	-
	$SB1 + p(\pi)$	10.809	1,979.072	0.935	0.099	173.325	0.037	35.083	-5.945	6.485	53.135	20.132	0.526
		[10.807, 10.81]	[1979.069, 1979.076]	[0.934, 0.936]	[0.099, 0.099]	[173.24, 173.42]	[0.0, 0.074]	[34.439, 35.999]	[-6.609, -5.106]	[5.233, 8.643]	[51.425, 54.781]	[5.715, 37.566]	[0.361, 0.823]
	$SB1 + p(m_1\|\theta)$	10.808	1,979.079	0.939	0.099	-179.362	354.018	32.093	-6.580	6.243	56.182	22.199	0.540
		[10.807, 10.809]	[1979.075, 1979.082]	[0.938, 0.94]	[0.099, 0.099]	[-179.999, 179.876]	[353.112, 355.062]	[31.262, 33.094]	[-7.133, -5.65]	[5.597, 8.479]	[54.357, 58.246]	[5.915, 29.157]	[0.442, 0.894]
	$SB1 + p(\pi) + p(m_1\|\theta)$	10.808	1,979.079	0.939	0.099	-179.323	353.986	31.943	-6.607	6.446	56.402	19.762	0.571
		[10.807, 10.809]	[1979.075, 1979.082]	[0.938, 0.94]	[0.099, 0.099]	[-180.0, 179.865]	[353.071, 355.035]	[31.285, 33.07]	[-7.182, -5.7]	[5.668, 7.741]	[54.315, 58.225]	[8.639, 28.756]	[0.458, 0.767]
HIP79101	$SB1$	1.545	1,966.733	0.465	0.033	-7.311	186.279	24.931	-16.719	7.867	-	-	-
		[1.544, 1.547]	[1966.718, 1966.748]	[0.447, 0.487]	[0.031, 0.035]	[-11.795, -2.96]	[180.631, 192.371]	[0.929, 35.108]	[-16.768, -16.647]	[4.325, 141.128]	-	-	-
	$SB1 + p(\pi)$	1.545	1,966.731	0.464	0.034	-8.159	187.152	28.483	-16.692	15.118	6.814	4.347	0.115
		[1.544, 1.547]	[1966.719, 1966.748]	[0.445, 0.487]	[0.032, 0.035]	[-11.56, -2.804]	[180.686, 192.096]	[5.485, 35.469]	[-16.767, -16.648]	[14.153, 15.698]	[4.868, 30.797]	[2.177, 5.203]	[0.069, 0.841]
	$SB1 + p(m_1\|\theta)$	1.545	1,966.735	0.471	0.034	-7.039	185.636	21.705	-16.713	17.597	8.931	2.449	0.186
		[1.544, 1.546]	[1966.718, 1966.748]	[0.446, 0.488]	[0.032, 0.035]	[-11.521, -2.708]	[180.37, 192.054]	[5.324, 35.809]	[-16.764, -16.647]	[14.182, 19.957]	[4.717, 30.313]	[1.733, 3.286]	[0.087, 0.812]
	$SB1 + p(\pi) + p(m_1\|\theta)$	1.545	1,966.737	0.469	0.033	-6.511	186.223	8.417	-16.709	15.189	23.156	2.736	0.543
		[1.544, 1.546]	[1966.718, 1966.748]	[0.449, 0.489]	[0.031, 0.034]	[-11.883, -2.89]	[180.923, 192.778]	[5.389, 13.403]	[-16.761, -16.642]	[14.267, 15.779]	[14.335, 34.332]	[2.106, 3.332]	[0.293, 0.999]
HIP99675	$SB1$	10.300	1,950.523	0.203	0.032	-160.963	306.392	109.643	-6.364	155.926	-	-	-
		[10.263, 10.331]	[1950.376, 1950.68]	[0.195, 0.21]	[0.025, 0.036]	[-163.437, -158.66]	[294.404, 317.684]	[80.399, 133.173]	[-6.47, -6.285]	[142.134, 199.951]	-	-	-
	$SB1 + p(\pi)$	10.309	1,950.479	0.204	0.037	-161.315	305.268	100.073	-6.375	3.727	131.027	4.686	0.954
		[10.266, 10.334]	[1950.367, 1950.676]	[0.195, 0.21]	[0.032, 0.041]	[-163.532, -158.776]	[296.412, 317.427]	[79.396, 115.39]	[-6.473, -6.285]	[3.093, 4.084]	[117.144, 152.989]	[4.024, 6.894]	[0.814, 1.0]
	$SB1 + p(m_1\|\theta)$	10.297	1,950.535	0.201	0.029	-160.958	304.857	116.110	-6.360	2.353	184.152	9.871	0.765
		[10.266, 10.332]	[1950.378, 1950.677]	[0.195, 0.21]	[0.025, 0.036]	[-163.391, -158.676]	[294.217, 317.669]	[82.309, 133.486]	[-6.472, -6.288]	[1.977, 2.982]	[135.573, 226.945]	[8.562, 11.368]	[0.629, 0.973]
	$SB1 + p(\pi) + p(m_1\|\theta)$	10.301	1,950.514	0.203	0.040	-161.350	306.934	102.004	-6.383	3.353	121.954	9.437	0.692
		[10.266, 10.333]	[1950.375, 1950.678]	[0.195, 0.21]	[0.033, 0.043]	[-163.343, -158.71]	[295.453, 316.188]	[82.462, 128.976]	[-6.469, -6.285]	[2.758, 3.595]	[111.512, 158.606]	[8.033, 10.914]	[0.63, 0.895]
HIP109951	$SB1$	58.309	1,990.498	0.391	0.321	62.426	262.000	32.115	-22.847	19.402	-	-	-
		[48.269, 61.633]	[1989.425, 1990.908]	[0.372, 0.465]	[0.275, 0.335]	[38.049, 69.562]	[256.403, 267.143]	[2.498, 35.157]	[-23.179, -22.73]	[13.987, 140.695]	-	-	-
	$SB1 + p(\pi)$	56.369	1,990.203	0.400	0.311	58.320	261.506	28.946	-22.871	14.899	20.748	1.986	0.447
		[50.086, 62.173]	[1989.552, 1990.979]	[0.37, 0.446]	[0.283, 0.337]	[42.264, 70.681]	[257.095, 267.473]	[15.397, 36.152]	[-23.149, -22.715]	[14.115, 16.223]	[16.554, 32.343]	[1.179, 2.36]	[0.321, 0.934]
	$SB1 + p(m_1\|\theta)$	56.941	1,990.286	0.392	0.314	57.656	263.599	30.022	-22.919	18.445	20.313	0.952	0.599
		[50.63, 63.167]	[1989.571, 1991.056]	[0.366, 0.441]	[0.286, 0.342]	[43.646, 72.714]	[257.021, 267.45]	[17.793, 37.547]	[-23.137, -22.7]	[16.816, 19.264]	[16.618, 29.321]	[0.926, 0.983]	[0.46, 0.968]
	$SB1 + p(\pi) + p(m_1\|\theta)$	52.515	1,989.840	0.430	0.291	49.172	261.551	19.005	-23.031	16.905	28.513	0.957	0.931
		[50.437, 55.072]	[1989.473, 1990.164]	[0.405, 0.449]	[0.284, 0.301]	[41.478, 56.262]	[256.624, 267.574]	[15.233, 23.905]	[-23.153, -22.809]	[16.507, 17.27]	[26.722, 30.285]	[0.935, 0.992]	[0.854, 1.0]
HIP115126	$SB1$	6.325	1,980.741	0.160	0.191	-146.961	341.392	50.015	-1.608	8.650	-	-	-
		[6.322, 6.328]	[1980.724, 1980.755]	[0.158, 0.163]	[0.189, 0.193]	[-147.419, -146.405]	[340.75, 341.798]	[48.775, 50.709]	[-1.623, -1.595]	[8.507, 8.862]	-	-	-
	$SB1 + p(\pi)$	6.326	1,980.738	0.161	0.191	-146.777	341.275	49.632	-1.608	45.134	8.702	1.144	0.647
		[6.322, 6.328]	[1980.724, 1980.754]	[0.158, 0.163]	[0.189, 0.192]	[-147.437, -146.421]	[340.773, 341.808]	[48.849, 50.751]	[-1.623, -1.595]	[43.862, 46.023]	[8.506, 8.847]	[1.06, 1.281]	[0.608, 0.673]
	$SB1 + p(m_1\|\theta)$	6.325	1,980.741	0.160	0.191	-146.900	341.256	50.010	-1.611	44.152	8.653	1.246	0.618
		[6.322, 6.328]	[1980.723, 1980.754]	[0.158, 0.163]	[0.189, 0.193]	[-147.452, -146.431]	[340.753, 341.795]	[48.801, 50.734]	[-1.623, -1.595]	[42.976, 46.613]	[8.505, 8.855]	[1.027, 1.364]	[0.595, 0.679]
	$SB1 + p(\pi) + p(m_1\|\theta)$	6.326	1,980.737	0.160	0.191	-146.774	341.189	49.866	-1.609	44.806	8.663	1.183	0.634
		[6.322, 6.328]	[1980.723, 1980.754]	[0.158, 0.163]	[0.189, 0.193]	[-147.424, -146.402]	[340.802, 341.814]	[48.83, 50.773]	[-1.623, -1.596]	[43.965, 45.8]	[8.501, 8.848]	[1.086, 1.273]	[0.608, 0.665]

(a) Results reported by other authors. (b) Results obtained in this work.

Chapter 6

Bayesian Inference in Hierarchical Stellar Systems

The vast advances in observational astronomy and the numerous resources destined for the construction of highly precise telescopes allow detecting subtle details of the dynamics and discovering new components of stellar systems once undetectable. The new resources increase the interest for the analysis of multiple stellar systems of various architectures.

This chapter introduces a Bayesian inference methodology for orbital parameters estimation in multiple stellar systems using the hierarchical approximation. The developed methodology provides a joint estimation of the posterior distribution of the orbital parameters of hierarchical stellar systems with any architecture and multiplicity. Section 6.1 introduces a Bayesian model for hierarchical stellar systems and discusses the possibilities and limitations for the estimation of the orbital parameters. Section 6.2 presents the evaluation of the proposed methodology on several hierarchical systems with different architectures, multiplicities, and sources of observations, analyzing and discussing the obtained results.

6.1 Bayesian Model

The Bayesian statistical model used to perform the inference in hierarchical stellar systems is the Keplerian orbital model introduced in Section 2.1.3. This model is adapted to determine the position observations and the radial velocities observations simultaneously applied to each of the binary sub-systems that compose it. Following the Bayesian modeling presented in Section 5.1 for binary stellar systems, our model assumes that the observations are independent and follow a Gaussian distribution with uniform priors distributions on the parameters space (considering it physical valid ranges described in Section 2.1).

The expressions for the determination of the position and radial velocities for a 3-hierarchical stellar system presented in Section 2.2.1 directly extends to n-hierarchical scenario, where we identify the same three important cases: positional observation formulae when the inner binary system is resolved; positional observation formulae when the inner binary system is unresolved and radial velocities observation formulae.

Let $\{S_i\}_{i=1}^{n}$ be the n-stars or components of an hierarchical stellar system, $C_k = \bigcup_{i=1}^{k}\{S_i\}$ be the subgroups or clusters of stars that satisfies $C_1 \subset C_2 \subset ... \subset C_n$, $\vec{r}_{S_k}$ and $\vec{r}_{C_k}$ the position vector of the star S_k and be the center of mass of the cluster C_k, respectively. The position vector of a component S_n

relative to principal component of the system S_1 can be expressed as follows:

$$\vec{r}_{S_1S_n} = \sum_{i=1}^{n-2} \vec{r}_{C_iC_{i+1}} + \vec{r}_{C_{n-1}S_n} = \sum_{i=1}^{n-2} \vec{r}_{C_iS_{i+1}} \left(\frac{q_{C_iS_{i+1}}}{1+q_{C_iS_{i+1}}} \right) + \vec{r}_{C_{n-1}S_n}. \tag{6.1}$$

The hierarchical structure of the orbital model imposes that the position vectors $\left\{\vec{r}_{C_iS_{i+1}}\right\}_{i=1}^{n-1}$ describes independent binary systems, consequently, the position vector $\vec{r}_{S_1S_n}$ is characterized by the set orbital parameters $\bigcup_{i=1}^{n-1}\{P_{C_iS_{i+1}}, T_{C_iS_{i+1}}, e_{C_iS_{i+1}}, a_{C_iS_{i+1}}, \omega_{C_iS_{i+1}}, \Omega_{C_iS_{i+1}}, i_{C_iS_{i+1}}, q_{C_iS_{i+1}}\} \setminus \{q_{C_{n-1}S_n}\}$.

If the inner binary systems are unresolved up to a binary system $\vec{r}_{C_{m-1},S_m}$ and denoting $\vec{r}_{C_m^*}$ the position vector of the photocenter of the cluster C_m, the relative position vector of a component S_n relative to the photocenter of C_m, can be expressed as follows:

$$\vec{r}_{C_m^*S_n} = \vec{r}_{C_m^*C_m} + \sum_{i=m}^{n-2} \vec{r}_{C_iC_{i+1}} + \vec{r}_{C_{n-1}S_n} = \vec{r}_{C_m^*C_m} + \sum_{i=m}^{n-2} \vec{r}_{C_iS_{i+1}} \left(\frac{q_{C_iS_{i+1}}}{1+q_{C_iS_{i+1}}} \right) + \vec{r}_{C_{n-1}S_n}, \tag{6.2}$$

with $\vec{r}_{C_m^*C_m} = \vec{r}_{C_m} - \vec{r}_{C_m^*}$ a combinatorial term that involves the orbital parameters $\{m_1, ...m_m\} \cup \{L_1, ..., L_m\}$.

Considering V_{S_j} and V_{C_j} the radial velocities of the component S_j and center of mass of the cluster C_j, respectively, they can be described as:

$$\begin{aligned} V_{C_j}(t) =& V_0 + \sum_{i=1}^{n-1} K_{S_{i+1}}[e_{C_iS_{i+1}} \cos\omega_{C_iS_{i+1}} + \cos(\nu_{C_iS_{i+1}}(t) + \omega_{C_iS_{i+1}})],\ \forall j \in \{1, ..., n-1\}, \\ V_{S_j}(t) =& V_{C_{j-1}}(t) - K_{S_j}[e_{C_{j-1}S_j} \cos\omega_{C_{j-1}S_j} + \cos(\nu_{C_{j-1}S_j}(t) + \omega_{C_{j-1}S_j})],\ \forall j \in \{2, ..., n\}. \end{aligned} \tag{6.3}$$

Finally, we have to impose an additional restriction on the mass of the binary systems, such as they respect the hierarchical approximation. Thereby, as $\vec{r}_{C_iS_{i+1}}, \forall i \in \{1, ..., n-1\}$ describes a binary systems with a primary object C_i and a companion object S_{i+1}, respectively, it total mass must be equal to the primary object mass of the subsequent binary systems $\vec{r}_{C_{i+1}S_{i+2}}$, i.e., $m_{C_i} + m_{S_{i+1}} \stackrel{!}{=} m_{C_{i+1}}$. According to (2.18) and assuming that $\pi_{C_iS_{i+1}} = \pi_{C_{i+1}S_{i+2}}, \forall i \in \{1, ..., n-1\}$ (the parallax is the same for the whole hierarchical system), the additional restriction on the mass can be reduced to the following expression:

$$a_{C_iS_{i+1}} = a_{C_{i+1}S_{i+2}} \cdot \sqrt[3]{\left(\frac{P_{C_iS_{i+1}}}{P_{C_{i+1}S_{i+2}}} \right)^2 \cdot \frac{1}{1+q_{C_{i+1}S_{i+2}}}}. \tag{6.4}$$

The mass restriction (6.4) implies that the semi-major axis parameter of all the binary systems $a_{C_iS_{i+1}}, \forall i \in \{1, ..., n-2\}$ with the exception of the most outer system $a_{C_{n-1}S_n}$ are deterministically determined by the orbital model, substantially reducing the parameters space of the inference. Note that (6.4) still satisfying the physical valid range of the semi-major axis $a \in (0, +\infty)$, since all the parameters involved are greater than zero. Analogously, the mass restriction in (6.4) can be restated in terms of the period P of the binary systems instead of the semi-major axis a, still satisfying it physical valid range $P \in (0, +\infty)$. However, (6.4) cannot be restated in terms of the mass ratio q, since the obtained deterministic expression does not satisfy the physical valid range of that parameter $q \in (0, 1]$.

The main advantage of the joint estimation of the set of orbital parameters (of all the binaries that compose the hierarchical system) is the information shared by the different observation sources. This strategy infers the orbital parameters that are non-identifiable under the separate binary modeling setting. For example, the mass ratio of a visual triple stellar system can be estimated only through the positional

observations of the inner and the outer binaries, not requiring observations of the radial velocities or additional information about the system.

Another great advantage of the Bayesian methodology in hierarchical stellar systems is the flexibility for incorporating prior information. Depending on the multiplicity, architecture, sources, and precision of the observations of the stellar system, some of the orbital parameters can be identifiable, but others do not. In many cases, the coverage of the observations of many outer systems is poor due to the high periods of revolution involved. In addition, many inner systems present no positional observations due to their high tightness, where the precision of the observational instruments is insufficient to identify both components of the inner binary. When the absence or the low quality of the observational sources indicates that some parameters are not identifiable, suitable informative priors can be used to constrain the orbital solution of the whole system, allowing to determine some orbital parameters that cannot be determined with the available observations. In this context, we say that a parameter is weak-identifiable if it becomes identifiable by incorporating an informative prior.

6.2 Experiments

This section performs the inference of the six hierarchical stellar systems through the proposed Bayesian inference methodology, considering systems of different multiplicities, architectures, and available observational sources. The estimates and the uncertainty in the parameters space are analyzed through visualization of the posterior marginal distributions. The posterior distribution is also computed in the observations space by projecting 1000 randomly selected samples of the posterior distribution, drawing trajectories from the first observation time t_0 to the first completion of the orbit $t_0 + P$. The maximum a posteriori (MAP) estimate and the 95% high posterior density interval estimates are summarized in Table 6.1.

The inference process is performed through the simulation of 10000 samples of the respective posterior distributions following the sampling method described in Section 5.1.

For the sake of space, the analysis is presented only on four of the eight systems: LHS1070, HIP110960, KUI99M and HIP51966.

6.2.1 LHS1070

The LHS1070 (GJ2005) is a triple hierarchical system composed of a tighter inner pair B_a, B_b that orbits a brighter primary component A. This system's data consists of numerous and precise positional observations of the inner pair B_aB_b that cover the entire orbit and precise observations of the primary component of the inner pair B_a relative to the brighter component A. No radial velocity observations are available for any of the system components. The observations and it uncertainties are visualized in Figure 6.2. The LHS1070 hierarchical system was previously studied by [Köhler et al., 2012, Tokovinin, 2018, Villegas et al., 2021].

The available sources of information make the set of astrometric orbital parameters of the inner B_a, B_b and the outer A, B systems, as well as the mass ratio of the inner system $q_{B_aB_b}$ identifiable. Consequently, no informative priors on the parameters of the system are used for the inference process.

The posterior distributions presented in Figure 6.1 shows that the orbital parameters of the inner pair B_a, B_b are well determined (being Gaussian distributions with low dispersion), which is attributed to the abundance, high precision, and high orbit coverage of the positional observations. However, the marginal posterior distribution of the orbital parameters of the outer pair A, B presents a significant dispersion and

skewness on $P_{AB}, T_{AB}, e_{AB}, a_{AB}, \omega_{AB}$, which is attributed to the reduced orbit coverage of the respective positional observations. In particular, the high uncertainty of the period P_{AB} is an expected result since the corresponding positional observations do not cover the same zones of the orbit in different periods. This is a consequence of the prohibitive time scale of observation (hundred of years). It is important to notice that, even with the low orbit coverage of the positional observation of the outer pair A, B, the inclination i_{AB} is determined with good precision due to the distribution of the observations in the orbital space and the information-sharing between the well determined inner orbit B_a, B_b and the outer orbit A, B. Finally, the marginal posterior distribution of the mass ratio $q_{B_a B_b}$ is well determined, presenting a reduced uncertainty but a high skewness to its maximum possible value ($q = 1$).

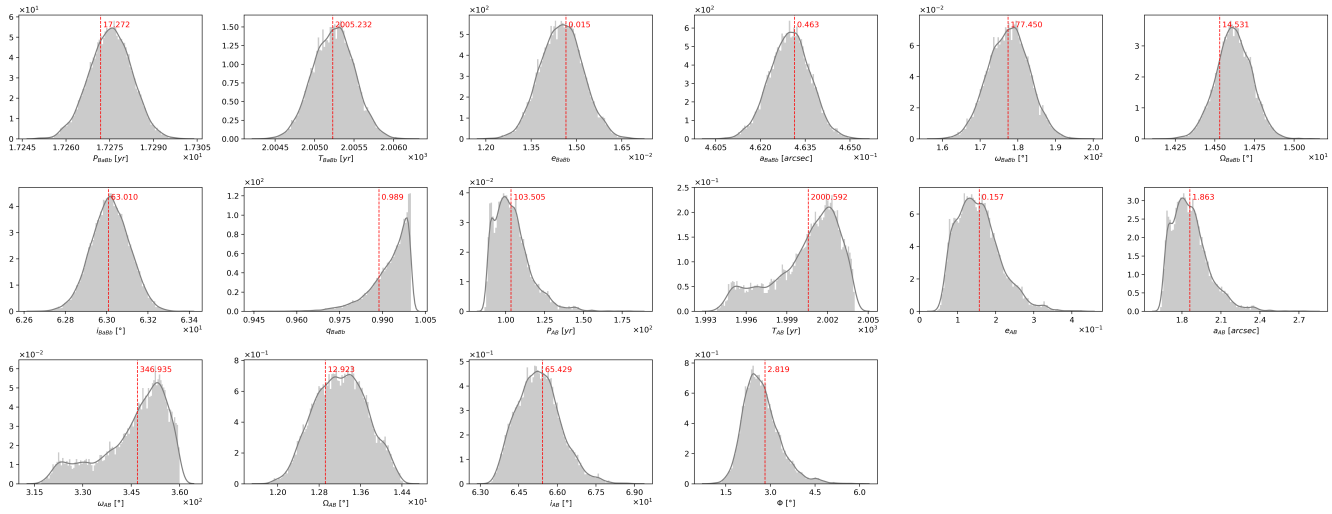

Figure 6.1: Marginal posterior distributions and MAP estimates of the orbital parameters of the hierarchical system LHS1070.

The posterior distributions in the observation space are presented in Figure 6.2. No uncertainty is observed in the posterior distribution of the inner orbit B_a, B_b, which perfectly fits the corresponding observations. Similar behavior is observed in the segments of the outer orbit populated by the positional observations. The upper orbital lobe, which has observations along all the curve, presents no observable uncertainty. In contrast, the subsequent lobe, populated with precise observation only on the initial small portion of the curve, presents a slight but visible uncertainty increase. All the other segments of the orbit with no observations present a considerable increase of the uncertainty, presenting three or four additional possible lobes. The observed indeterminacy of the outer orbit shape of the LHS1070 system is a clear example of the advantages of the Bayesian inference approach over the deterministic estimation approach (commonly used in the literature), offering more complete inference perspective of this stellar system.

6.2.2 HIP110960

The HIP110960 is a triple hierarchical system composed of a tighter inner pair A_a, A_b orbited by a fainter companion object component B. The available data consists of five positional observations of the inner pair $A_a A_b$ (distributed along all the orbit), numerous and imprecise positional observations of the outer component B relative to the center of light of the inner system A^* (that covers less than a half of the orbit), and few but precise recent positional observation of the outer component B relative to the primary component of the inner system A_a. No radial velocity observations are available for any of the system components. The observations and its uncertainties are visualized in Figure 6.4. The HIP110960 hierarchical system was previously studied by [Heintz, 1984, Tokovinin, 2016].

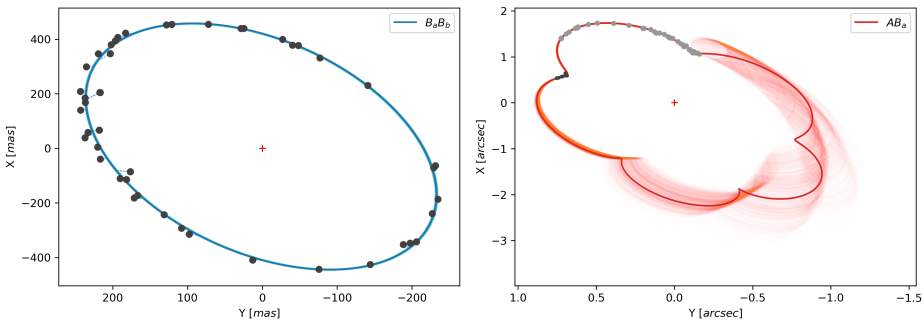

Figure 6.2: Estimated orbits of the hierarchical system LHS1070.

The available sources of information make the set of astrometric orbital parameters of the inner A_a, A_b and the outer A_a, B systems, as well as the mass ratio of the inner system $q_{A_aA_b}$ and the luminosity ratio $L_{A_aA_b}$ identifiable. Therefore, no informative priors on the parameters are used for the inference process.

The posterior distributions presented in Figure 6.3 shows that the orbital parameters $P_{A_aA_b}$ and $T_{A_aA_b}$ of the inner pair are well determined, (being Gaussian distributions with low dispersion), while the posterior distribution of the orbital parameters $e_{A_aA_b}$ and $a_{A_aA_b}$ presents a slight increase of skewness and dispersion. The angular parameters $\omega_{A_aA_b}, \Omega_{A_aA_b}, i_{A_aA_b}$ are highly uncertain presenting a considerably higher dispersion on the respective posterior distributions. This distribution looks like uniform distributions with a pronounced mode on one bound of the respective parameter space. The irregular shape of the angular parameters posterior distribution of the inner pair is attributed to the lack of orbit coverage of the positional observations, which are only distributed on a quarter of the whole orbit. The orbital parameters posterior distributions of the outer orbit A, B presents all a very high dispersion and with a high skewness on the parameters P_{AB}, T_{AB} and e_{AB}. The high uncertainty of the outer orbit orbital parameters is also attributed to the low orbit coverage of the respective positional observations. As there are positional observations from A^*B and A_aB simultaneously, the mass and light ratios $q_{A_aA_b}$ and $L_{A_aA_b}$ can be determinate without the incorporation of additional information in the form of informative priors. The mass ratio $q_{A_aA_b}$ posterior distribution has a Gaussian shape centered in the center value of the valid physical range ($q \in (0, 1]$) with a moderate dispersion. On the other hand, the light ratio $L_{A_aA_b}$ posterior distribution has a moderate skewness to its range lower bound with low uncertainty.

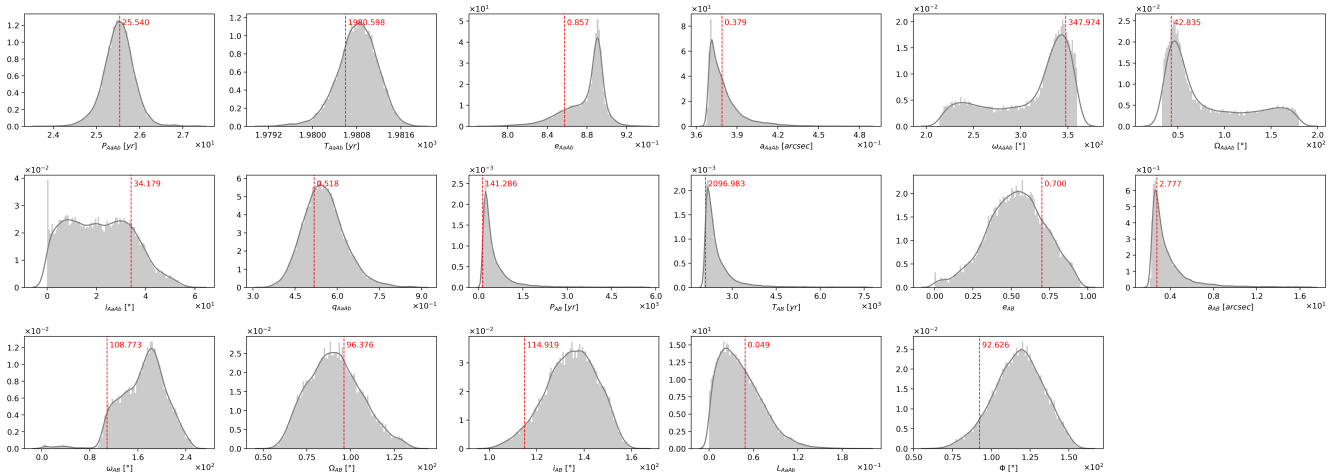

Figure 6.3: Marginal posterior distributions and MAP estimates of the orbital parameters of the hierarchical system HIP110960.

The posterior distributions in the observation space are presented in Figure 6.4. A low uncertainty is observed in the inner orbit B_a, B_b in the segment populated with observations, presenting an almost perfect fitting. The uncertainty slightly increases in the zones of the orbit without observations. Similar behavior is observed in the sections of the outer orbit populated with the positional observations. However, the projected uncertainty of the orbit increases dramatically in the areas far from the observations (which extends further from the plot limits). A significant uncertainty reduction is observed at the end of the periods P_{AB} of the orbits, which is restricted to a well-determined narrow range.

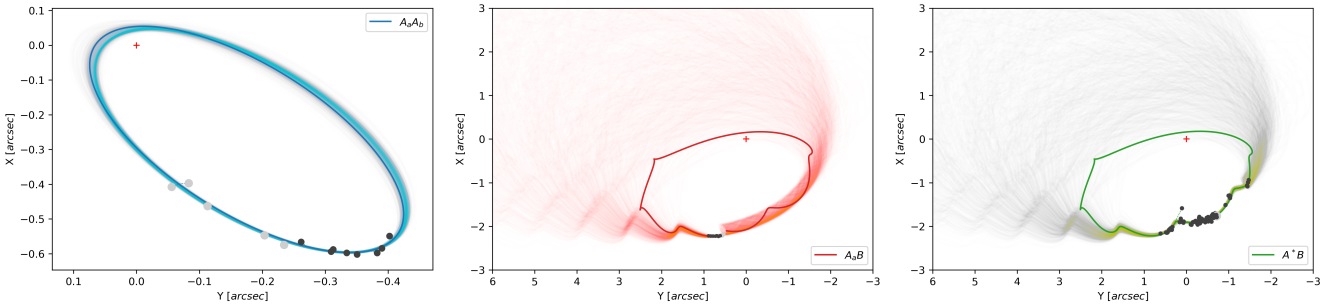

Figure 6.4: Estimated orbits of the hierarchical system HIP110960.

6.2.3 KUI99M

The KUI99M (HIP111805) is a triple hierarchical system composed of a tighter inner pair A_a, A_b orbited by a fainter companion object component B. The available data consists of numerous and precise positional observations of the inner A_aA_b and the outer pair A_a, B distributed along both orbits. In addition, there are precise and abundant radial velocity observations of the primary object A_a relative to the center of mass of the whole system and scarce and imprecise radial velocity observations of the companion object A_b relative also to the center of mass of the whole system. No radial velocity observations of the outer component B are available. The observations and their uncertainties are visualized in Figure 6.6. The KUI99M hierarchical system was previously studied by [Tokovinin and Latham, 2017, Villegas et al., 2021].

The available sources of observations make the set of astrometric and spectroscopic orbital parameters of the inner A_a, A_b and the outer A_a, B systems, as well as the mass ratio of the inner system $q_{A_aA_b}$ identifiable. Consequently, no informative priors on the parameters are used for the inference process.

The posterior distributions presented in Figure 6.5 show that the astrometric and spectroscopic orbital parameters of the inner A_a, A_b, and outer $A.B$ pairs are well determined, presenting a Gaussian shape with low dispersion. It is important to note that the mass ratio parameter $q_{A_aA_b}$ is determined through the positional and radial velocities of the inner system A_a, A_b, as well as the gravitational interaction between the inner A_a, A_b and outer A, B orbits. Therefore, there is information sharing between the two sources of information that allows a low uncertainty estimate of $q_{A_aA_b}$. As the same parameter of parallax π is assumed for the inner and outer systems, the positional observations of the inner and outer orbit and the radial velocities observations of both bodies of the inner orbit allows to estimate the mass ratio of the outer system q_{AB} with low uncertainty. This is attributed to the information-sharing between all the sources of information, even in absence of explicit radial velocity observations of the outer companion object V_B.

The posterior distributions in the observation space are presented in Figure 6.6. No uncertainty is observed in the posterior distribution of the inner A_a, A_b and outer orbits A, B, perfectly fitting the corresponding observations. The same behavior is observed in the radial velocity curve of the primary component of the inner orbit V_{A_a}. However, the radial velocity curve of the companion object V_{A_b} does

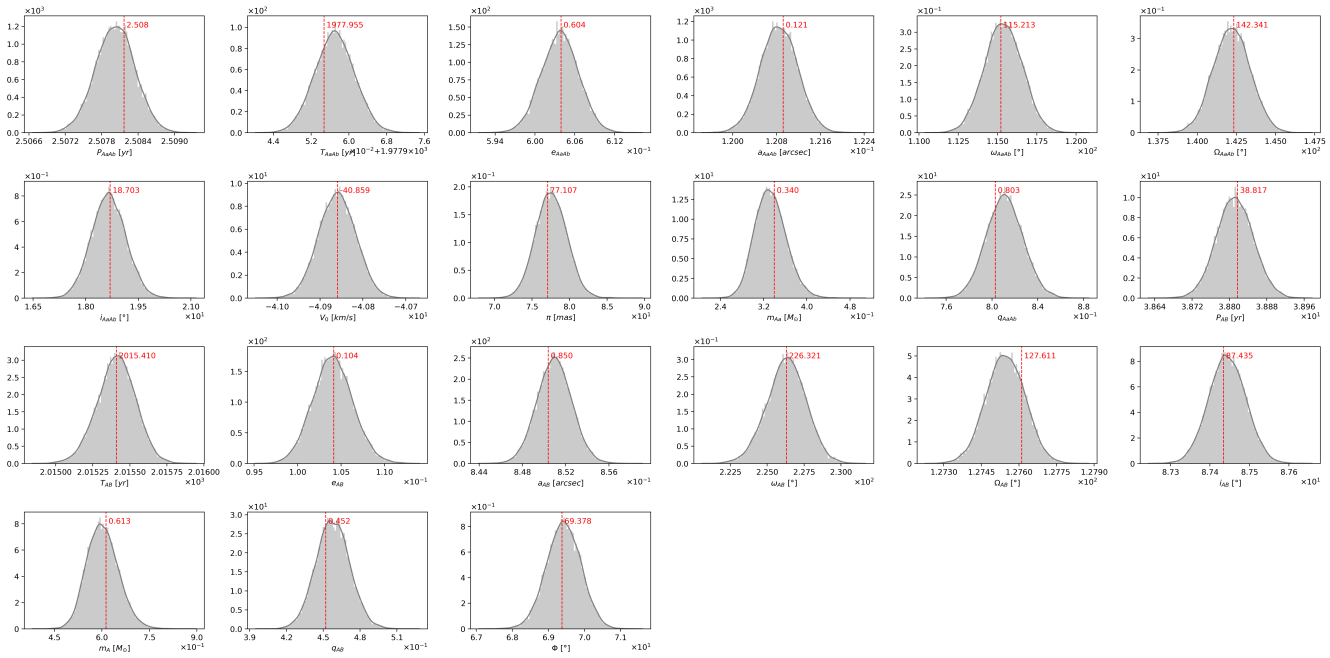

Figure 6.5: Marginal posterior distributions and MAP estimates of the orbital parameters of the hierarchical system KUI99M.

not reach the observations, failing to fit them adequately. The fitting of the V_{A_b} observations is displaced by the information provided by the other sources of information as the positional observations and the gravitational interaction between the inner and outer orbits, providing an estimate with almost no uncertainty even if it does not reach the corresponding observations. The radial velocity observations V_{A_a} and V_{A_b} allows to provide a confident estimate of the radial velocity curve of the outer object V_A and, due to the information sharing between all the sources of information, the radial velocity curve of the companion object of the outer orbit V_B is also determined, but with a slight but visible increase of uncertainty in the peak and trough of the curve.

6.2.4 HIP51966

The HIP51966 is a quadruple hierarchical system composed of a tighter inner pair A_{a1}, A_{a2} that is orbited by a fainter object A_b, forming an inner triple system that is in turn orbited by an outer object B. The available data consists of numerous and noisy positional observations of the component A_b relative to the light center of the inner system A^* (distributed along all the orbit) and few but precise observations of the outer component B (relative to the light center of the inner system A_a^*). No positional observations of the inner pair are available, i.e., the inner system is unresolved. However, numerous and highly precise radial velocity observations of the primary component of the inner pair A_a (relative to the center of mass of the whole system) are available. No radial velocity observations for the other system components are available. The observations and their uncertainties are visualized in Figure 6.8. The HIP51966 hierarchical system was previously studied by [Tokovinin et al., 2015].

As only radial velocity observations of the inner orbit A_{a1}, A_{a2} are available, only its spectroscopic orbital parameters are identifiable. To identify the astrometric orbital parameters of the inner orbit as well as the mass ratio $q_{A_{a1}A_{a2}}$, an informative Gaussian prior is imposed on the system parallax $\pi \sim \mathcal{N}\left(26.3 \cdot 10^{-3}, (0.401 \cdot 10^{-3})^2\right)$ as reported in [Brown et al., 2018]. The weak-identifiability of the astrometric orbital parameter of the inner systems makes the astrometric orbital parameters of the middle

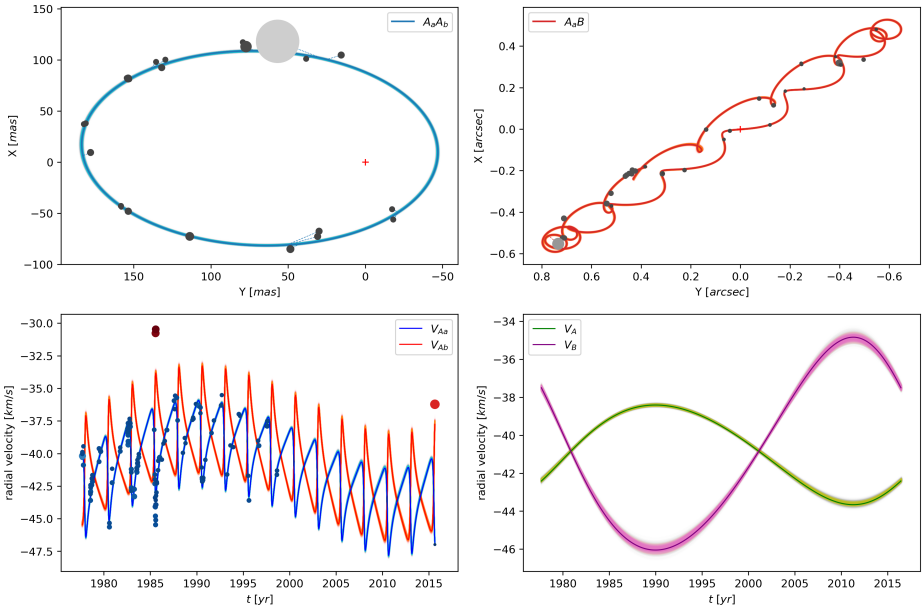

Figure 6.6: Estimated orbits and radial velocity curves of the hierarchical system KUI99M.

A_aA_b and the outer AB systems also identifiable. The positional observations $A_a^*A_b$ are not sufficiently precise to capture the subtle wobble movement of the companion object A_b due to gravitational interaction with the inner no-observable component A_{a1}. For address this issue, [Tokovinin et al., 2015] proposes fixing the values of the inclination $i_{A_{a1}A_{a2}}$ and the longitude of the ascending node $\Omega_{A_{a1}A_{a2}}$. In contrast, we avoid using informative priors on those parameters to study the full-expressiveness of the available positional data of the pair A_a, A_b on the determination of the inner orbit $A_{a1}A_{a2}$.

The obtained posterior distributions presented in Figure 6.7 show that the spectroscopic parameters of the inner pair are determined with low uncertainty. This is attributed to the available precise radial velocity observations of the primary component A_{a1}. It is observed that the posterior distribution of the weak-identifiable parameters presents evident differences compared to the prior distributions used. The posterior distribution of the parallax π is slightly shifted to higher values compared to the assumed prior, but with a similar dispersion. The posterior distribution of the light ratio $L_{A_{a1}A_{a2}}$ presents a higher skewness to the zero value compared to the assumed prior and lower dispersion. Surprisingly, the angular astrometric parameters $\Omega_{A_{a1}A_{a2}}, i_{A_{a1}A_{a2}}$, are well determined, even considering that no positional data of the inner system A_{a1}, A_{a2} is available. This is attributed to the information provided by the informative priors on π assumed and the information sharing between the inner orbit (with no positional observations) and the middle orbit (with positional observations). It is important to remark that the mass ratio $q_{A_{a1}A_{a2}}$ and the semimajor axis $a_{A_{a1}A_{a2}}$ of the inner orbit are both identifiable: $a_{A_{a1}A_{a2}}$ is deterministically determined (eq. (6.4)) as a function of $P_{A_{a1}A_{a2}}$ (determined through the radial velocities observations of A_{a1}), $P_{A_aA_b}$ (determined through the available positional observations of A_a^*, A_b) and $q_{A_aA_b}$ (determined through the positional observations of A_a^*, B), allowing $q_{A_{a1}A_{a2}}$ to be identified through the interaction between the inner and the middle orbit. Compared to the orbital estimation in [Tokovinin et al., 2015], it is important to remark that [Tokovinin et al., 2015] completely ignores the angular parameter $\Omega_{A_{a1}A_{a2}}$ in his estimation due to the absence of positional observation of the inner orbit. In contrast, this work incorporates that parameter in the inference, showing a surprisingly not very high uncertainty, with a 95% HPDI of [53.17, 90.01]. Additionally, the estimated inclination of the inner orbit $i_{A_{a1}A_{a2}}$ ($\sim$127°) obtained in this work strongly differs from the fixed value assumed by [Tokovinin et al., 2015] (57°), questioning his assumption of co-planarity between the inner and middle orbits. Well-determined estimations are obtained in the

majority of the inner orbit orbital parameters. However, the posterior distribution of the mass ratio $q_{A_{a1}A_{a2}}$ presents a somewhat high uncertainty, where the corresponding posterior distribution covers all the valid range [0,1] but is concentrated in it first half [0,0.5]. This uncertainty is attributed to the absence of positional observations of the inner orbit. Finally, the orbital parameters of the middle system A_a, A_b are all well-determined, presenting a lower uncertainty, including the mass ratio $q_{A_aA_b}$. However, all the orbital parameters of the outer system A, B present an extremely high uncertainty. The respective posterior distributions have wide dispersion and irregular shapes, attributed to the low orbital coverage of the available positional observations.

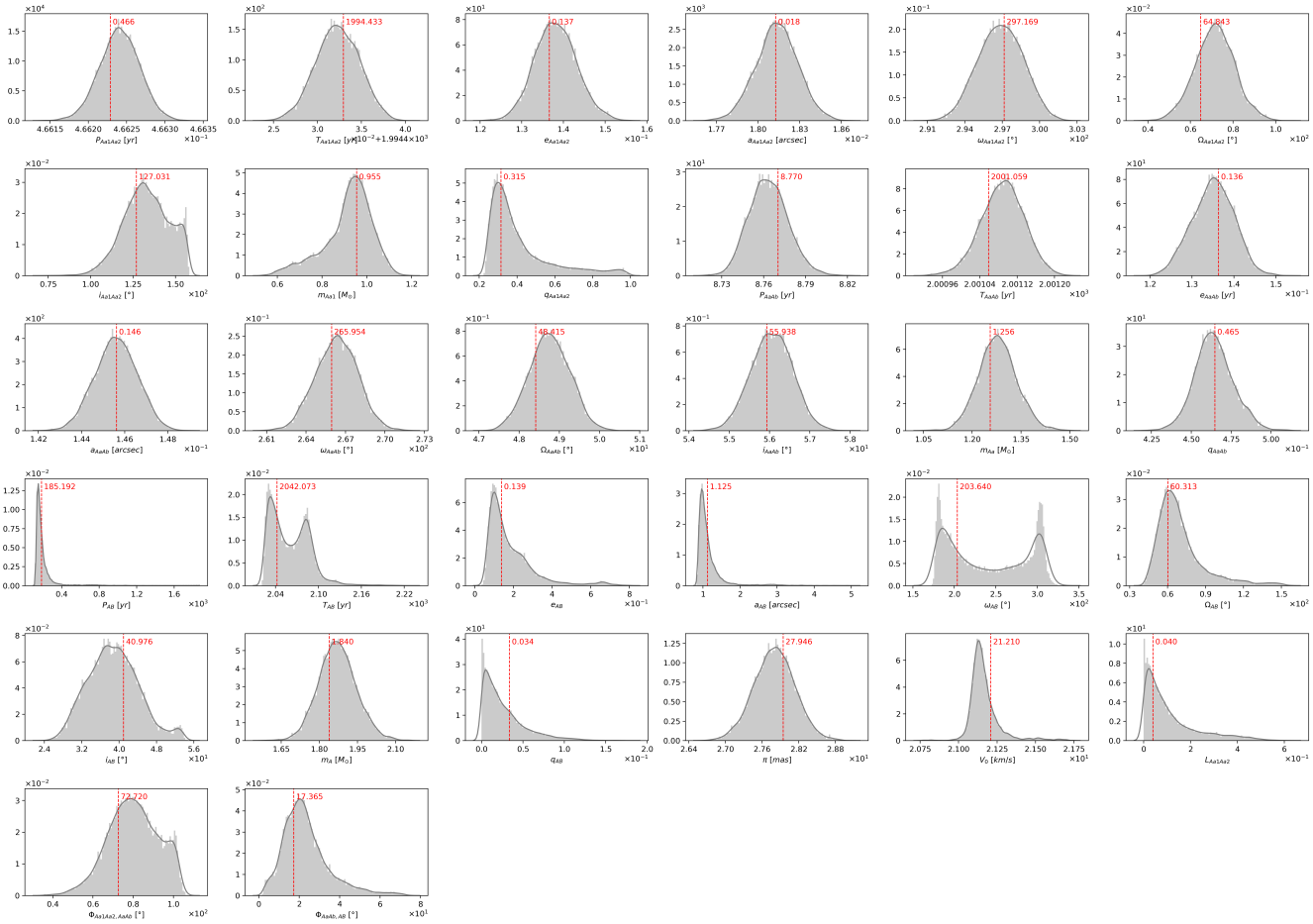

Figure 6.7: Marginal posterior distributions and MAP estimates of the orbital parameters of the hierarchical system HIP51966.

The obtained posterior distributions in the observation space are presented in Figure 6.8. A high uncertainty is observed in the inner orbit A_{a1}, A_{a2}, showing a diffuse cloud of possible orbits attributed to the somewhat high uncertainty obtained in the angular parameters $\Omega_{A_{a1}A_{a2}}$ and $i_{A_{a1}A_{a2}}$ (due to the absence of positional observations in A_{a1}, A_{a2}), but the size of the orbit is invariant in all the projected orbits even in the absence of positional observations (due to the precise estimations of $e_{A_{a1}A_{a2}}$ and $a_{A_{a1}A_{a2}}$). This is attributed to the prior information on the parallax π and information provided by the radial velocities observations and the positional observations of the middle orbit. The projection of the radial velocity curves $V_{A_{a1}}$ shows no uncertainty, which is expected considering the quantity and quality of the corresponding observations. Low uncertainty is observed in the middle orbit A_a^*, A_b with a good fitting of the available positional observations. High uncertainty is observed in the zones with no observation of the outer orbit A_a^*, B. However, almost no uncertainty is observed in the orbit zone populated with positional

observations, with perfect observations fitting.

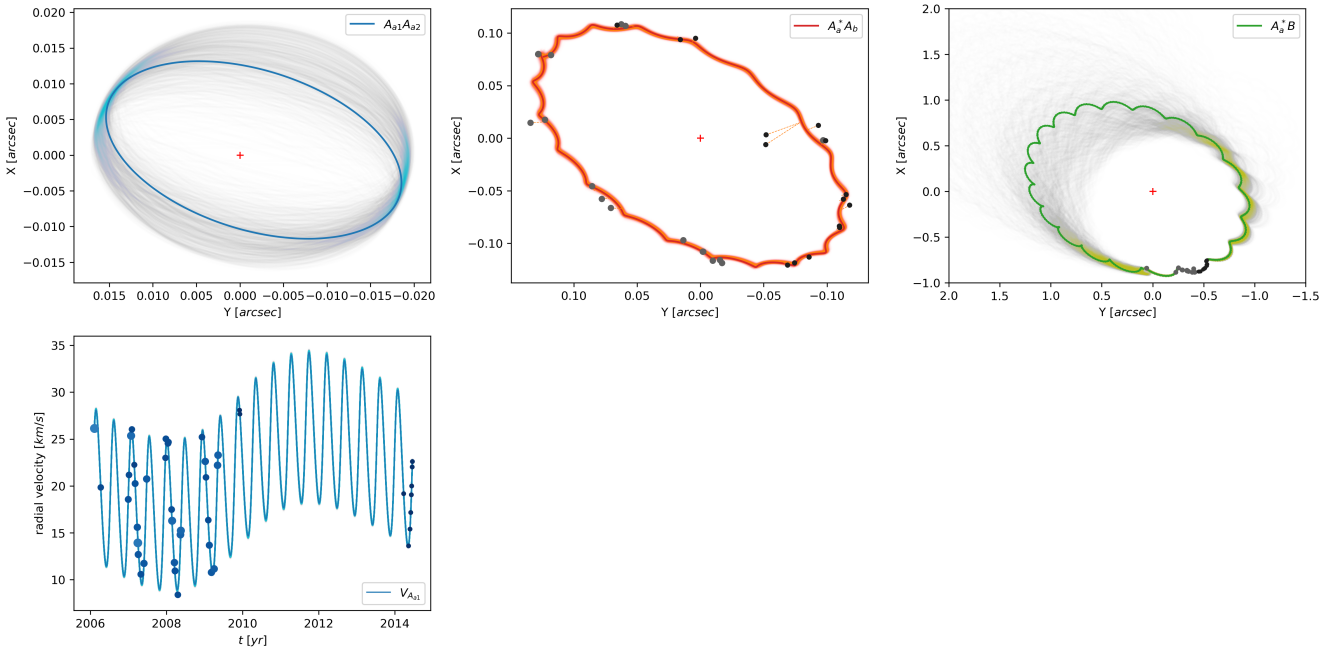

Figure 6.8: Estimated orbits and radial velocity curves of the hierarchical system HIP51966.

6.3 Concluding Remarks

The experimental results show the efficacy and precision of the proposed Bayesian inference methodology on several hierarchical stellar systems of different multiplicities, architectures, and sources of observations available.

Interesting results are shown in some of the hierarchical systems evaluated. For example, in the systems LHS1070 and HIP110960, the posterior distributions in the orbits space show an extremely high uncertainty in the zones of the outer orbits far from the observations. These high uncertainties translate into multiple probable orbits of very different shapes and periods, being an interesting insight to the analysis of these objects, since the majority of the related works in the hierarchical systems field base its analysis in a single deterministic orbit, completely ignoring the diversity of other possible orbits. In the case of the KUI99M system, the results show a posterior distribution with low dispersion in all the parameters, which is translated into orbits and radial velocities curves with almost no uncertainty that adequately fits the available data in both the inner and the outer binary system. This result shows the efficacy of the Bayesian methodology in performing a precise inference when multiple sources of data are available. The most interesting result is presented in the case of the quadruple system HIP51966, which has multiple missing sources of observations (positional and radial velocities observations of the binaries that compose the hierarchical system). However, the information-sharing between the different sources of observations and the prior information incorporated allowed a good estimation of the orbital parameters. The obtained results were similar to those reported by other authors but without fixing some of the orbital parameters involved. Specifically, we impose an informative prior only on the parallax of the system, while the other work needs to fix three of the orbital parameters of the inner orbit to perform a good estimation. This capability of the proposed method allows a better expressiveness of the data in the results, avoiding inducing strong biases to the estimation procedure (due to the parameter fixing).

Table 6.1: MAP estimates and 95% HDPIs of the marginal posterior distributions of the orbital parameters of several hierarchical stellar systems.

System	Subsystem	Author	P $[yr]$	T $[yr]$	e	a ‖	ω $[^\circ]$	Ω $[^\circ]$	i $[^\circ]$	K_1 $[km/s]$	K_2 $[km/s]$	V_0 $[km/s]$	π $[mas]$	q	L	Φ $[^\circ]$
LHS1070	inner	(a)	17.247	2,006.440	0.017	0.460	202.530	14.820	62.040	-	-	-	-	0.957	-	0.935
		(b)	17.272	2,005.232	0.015	0.463	177.450	14.531	63.010	-	-	-	-	0.989	-	2.819
			[17.261, 17.289]	[2004.756, 2005.758]	[0.013, 0.016]	[0.462, 0.464]	[167.449, 188.341]	[14.396, 14.837]	[62.839, 63.201]	-	-	-	-	[0.979, 1.0]	-	[1.6, 3.965]
	outer	(a)	77.620	2,049.670	0.039	1.528	210.700	13.900	62.500	-	-	-	-	-	-	-
		(b)	103.505	2,000.592	0.157	1.863	346.935	12.923	65.429	-	-	-	-	-	-	-
			[87.322, 126.559]	[1995.358, 2003.946]	[0.065, 0.27]	[1.65, 2.142]	[323.963, 359.872]	[12.314, 14.177]	[63.768, 66.919]	-	-	-	-	-	-	-
HIP111805	inner	(a)	1.501	1,986.093	0.022	0.038	232.900	334.500	87.400	13.130	19.210	-	-	0.493	-	175.673
		(b)	1.501	1,900.962	0.012	0.034	337.938	336.213	83.368	12.909	15.829	-	-	0.816	-	171.294
			[1.501, 1.502]	[1900.93, 1901.045]	[0.0, 0.03]	[0.033, 0.034]	[328.072, 359.929]	[333.41, 339.84]	[79.489, 85.321]	[12.634, 13.132]	[14.74, 16.518]	-	-	[0.772, 0.879]	-	[167.43, 173.123]
	outer	(a)	30.127	2,010.179	0.324	0.336	84.920	154.250	88.280	6.057	8.600	-22.580	-	0.705	-	-
		(b)	29.979	1,920.049	0.324	0.332	84.436	154.164	88.170	6.239	8.903	-22.547	23.011	0.701	-	-
			[29.926, 30.045]	[1919.857, 1920.217]	[0.319, 0.333]	[0.331, 0.334]	[84.039, 84.987]	[154.019, 154.361]	[88.051, 88.353]	[5.947, 6.468]	[8.424, 9.227]	[-22.735, -22.409]	[22.499, 23.906]	[0.658, 0.75]	-	-
KUI99M	inner	(a)	2.510	2,000.518	0.617	0.124	109.700	147.100	24.100	3.270	6.930	-	-	0.840	-	64.791
		(b)	2.508	1,977.955	0.605	0.121	114.896	142.642	19.255	3.336	4.100	-	-	0.814	-	68.909
			[2.507, 2.509]	[1977.95, 1977.966]	[0.6, 0.611]	[0.12, 0.121]	[113.399, 117.952]	[139.507, 144.003]	[18.419, 20.296]	[3.251, 3.412]	[3.912, 4.284]	-	-	[0.782, 0.845]	-	[67.839, 69.667]
	outer	(a)	38.679	2,016.110	0.118	0.855	233.400	127.600	87.400	2.660	-	-41.110	-	0.445	-	-
		(b)	38.786	2,015.527	0.106	0.852	227.540	127.520	87.462	2.625	5.608	-40.845	79.871	0.468	-	-
			[38.718, 38.866]	[2015.193, 2015.685]	[0.1, 0.109]	[0.849, 0.855]	[224.161, 229.193]	[127.394, 127.697]	[87.365, 87.548]	[2.498, 2.72]	[5.255, 5.903]	[-40.909, -40.744]	[76.356, 84.316]	[0.441, 0.493]	-	-
HIP110960	inner	(a)	25.950	2,006.520	0.872	0.110	100.900	293.700	11.800	-	-	-	-	-1.400	-	153.045
		(b)	25.540	1,980.598	0.857	0.379	347.974	42.835	34.179	-	-	-	-	0.518	0.049	92.626
			[24.817, 26.223]	[1980.115, 1981.512]	[0.834, 0.911]	[0.368, 0.407]	[228.529, 359.948]	[33.731, 167.391]	[0.009, 43.094]	-	-	-	-	[0.415, 0.704]	[0.0, 0.097]	[85.144, 147.367]
	outer	(a)	540	1,981.500	0.419	3.496	269.300	131.300	142.000	-	-	-	-	-	-	-
		(b)	141.286	2,096.983	0.700	2.777	108.773	96.376	114.919	-	-	-	-	-	-	-
			[95.625, 1989.406]	[2037.897, 3962.022]	[0.191, 0.916]	[2.152, 8.665]	[101.679, 232.78]	[63.201, 120.781]	[113.521, 154.486]	-	-	-	-	-	-	-
HIP113726	inner	(a)	-	-	-	-	-	-	-	-	-	-	-	-	-	-
		(b)	5.623	1,975.722	0.245	0.056	217.376	95.406	161.201	-	-	-	-	0.532	0.539	71.296
			[5.592, 5.709]	[1975.631, 1975.754]	[0.155, 0.289]	[0.05, 0.062]	[127.395, 295.975]	[12.716, 179.923]	[147.549, 179.995]	-	-	-	-	[0.111, 1.0]	[0.111, 1.0]	[51.413, 96.908]
	outer	(a)	-	-	-	-	-	-	-	-	-	-	-	-	-	-
		(b)	154.907	2,006.496	0.429	0.337	121.398	4.347	110.161	-	-	-	-	-	-	-
			[115.884, 190.946]	[2004.709, 2009.122]	[0.328, 0.502]	[0.299, 0.374]	[109.922, 136.957]	[3.262, 5.566]	[108.095, 111.684]	-	-	-	-	-	-	-
HIP51966	inner	(a)	0.466	2,003.292	0.135	0.018	297.600	-	57	8.080	-	-	-	-	-	-
		(b)	0.466	1,994.433	0.137	0.018	297.169	64.843	127.031	8.005	25.393	-	-	0.315	0.040	72.720
			[0.466, 0.466]	[1994.428, 1994.437]	[0.128, 0.148]	[0.018, 0.018]	[293.368, 300.198]	[53.17, 90.01]	[110.428, 157.258]	[7.93, 8.098]	[8.222, 31.467]	-	-	[0.232, 0.813]	[0.0, 0.376]	[58.533, 103.593]
	middle	(a)	8.846	2,009.816	0.125	0.150	263.900	50.400	56.600	4.710	-	-	-	-	-	10.797
		(b)	8.770	2,001.059	0.136	0.146	265.954	48.415	55.938	4.697	10.102	-	-	0.465	-	17.365
			[8.738, 8.792]	[2000.998, 2001.178]	[0.125, 0.144]	[0.144, 0.148]	[263.281, 269.416]	[47.779, 49.721]	[55.089, 57.073]	[4.618, 4.833]	[9.713, 10.595]	-	-	[0.441, 0.489]	-	[2.725, 47.036]
	outer	(a)	205	2,048.600	0.301	1.334	219.700	63.200	54.200	0.210	-	21.160	-	-	-	-
		(b)	185.192	2,042.073	0.139	1.125	203.640	60.313	40.976	0.140	4.147	21.210	27.946	0.034	-	-
			[118.646, 350.193]	[2023.037, 2102.572]	[0.047, 0.435]	[0.836, 1.716]	[177.163, 308.672]	[43.035, 110.905]	[28.276, 49.543]	[0.0, 0.329]	[3.619, 4.403]	[21.013, 21.381]	[27.14, 28.461]	[0.0, 0.08]	-	-

(a) Results reported by other authors. (b) Results obtained in this work.

All the mentioned capabilities make our Bayesian inference method an interesting analysis tool that could enhance the study of multiple stellar systems and simplify the estimation process. The mentioned simplification refers to the capacity of the method to be applied on hierarchical systems of any multiplicity and architecture, also allowing to incorporate any prior information to constrain the orbital estimation straightforwardly.

Chapter 7

An Application: Optimal Measurement Time

The technological advances in observational astronomy have allowed the discovery of an enormous number of new celestial objects of great interest to the astronomical community, motivating numerous observation campaigns to collect data for its study. Unfortunately, the number of celestial objects of study exceeds the observational capacity due to the limited high precision observational instruments. This limitation causes fierce competition among astronomers for precious observational time, limiting the number of observations that can be made. In that context, it is crucial to determine the adequate observational times to define observational planning that adapts to the time constraints and, at the same time, that allows characterizing the object of study as best as possible.

The current chapter introduces a methodology for the determination of optimal times of measurement in hierarchical stellar systems based on the Maximum Entropy Sampling [Shewry and Wynn, 1987] criterion. The proposed method uses the estimated posterior distribution of hierarchical stellar systems to generate a probability distribution of the time of measurement that provides the highest information about the system parameters. Section 7.1 introduces the maximum entropy sampling criterion and discuss its connections with the classical Bayesian optimal design problem; Section 7.2 presents the methodology of application of the maximum entropy sampling criterion on the hierarchical stellar system model, and Section 7.3 evaluates the optimal sampling methodology on three stellar systems with exhaustive analysis and discussion of the obtained results.

7.1 Maximum Entropy Sampling

The Maximum Entropy Sampling [Shewry and Wynn, 1987] is an information theoretic-based criterion that proposes to select experiments or samples that maximizes the gain in information for prediction at unsampled sites.

The concept of information is defined in terms of the entropy [Cover, 1999], which is a measure of uncertainty or average level of information of a random variable. Let X be a discrete random variable with alphabet $\mathcal{X}$ and probability mass function $p(x) = \mathbb{P}(X = x), x \in \mathcal{X}$, the entropy of X is formally defined as follows:

$$H(X) = \sum_{x \in \mathcal{X}} -p(x) \log p(x). \tag{7.1}$$

The analogous concept of entropy for continuous random variables is the differential entropy. Let X a continuous random variable with support on $\mathbb{X}$ and probability density function $f(x)$, the differential entropy of X is defined as follows:

$$h(X) = \mathbb{E}\{-\log f(x)\} = -\int_{\mathbb{X}} f(x)\log f(x)dx. \tag{7.2}$$

In the information-theoretic language, the Maximum Entropy Sampling criterion attempts to choose an experiment or sample that minimizes the uncertainty or entropy of a system conditioned to it. Let $S = [1, ..., N]$ be a finite system described by a random vector Γ_i, with $i \in S$ the observation site. Let $s, \bar{s}$ two partitions of S, i.e, $S = s\dot{\cup}\bar{s}$, where $\Gamma_i, i \in s$ are the observations or samples of the process. The objective of the problem is to interpolate the unobserved sites of the system $\Gamma_i, i \in \bar{s}$ from the observed portion of the system $\Gamma_i, i \in s$. By the chain rule property of the entropy on Γ_S, the following expression is obtained:

$$H(\Gamma_S) = H(\Gamma_s) + \mathbb{E}_{\Gamma_S}\{H(\Gamma_{\bar{s}}|\Gamma_s)\}. \tag{7.3}$$

The maximisation of the information on the unsampled sites conditioned to the samples is achieved by minimizing the term $\mathbb{E}_{\Gamma_S}\{H(\Gamma_s|\Gamma_{\bar{s}})\}$ in (7.3), but as $H(\Gamma_S)$ is fixed and assuming it is finite, the problem reduces to the maximization of $H(\Gamma_s)$, i.e., to select the samples $s^* = \arg\max_{s\subseteq S} H(\Gamma_s)$ with the highest observed uncertainty.

The maximum entropy sampling criterion is philosophically different from other classical optimal design approaches. The proposed criterion wishes to absorb the maximum amount of variability into the sample so that conditionally on the sample, the unsampled population has minimum variability. In particular, the maximum entropy sampling criterion is different from the classical Bayesian optimal experimental design, where the figure of merit is the minimization of the posterior entropy. However, [Sebastiani and Wynn, 2000] showed that both problems are dual if an additional condition is satisfied. Following the Bayesian experimental framework presented in Section 3.1, let $y \in Y$ the observations of a system, $\theta \in \Theta$ its parameters and $\xi \in \Xi$ an experiment. By choosing $\Gamma_s \equiv Y|\xi$ and $\Gamma_{\bar{s}} \equiv \Theta$ in (7.3), the following expression is obtained:

$$H(Y, \Theta|\xi) = H(Y|\xi) + \mathbb{E}_Y\{H(\Theta|Y, \xi)\}. \tag{7.4}$$

If $H(Y, \Theta|\xi)$ is fixed in the sense that is independent of any experiment $\xi \in \Xi$ and all the terms in (7.4) are finite, the maximization of $\mathbb{E}_Y\{H(\theta|Y, \xi)\}$ is achieved by minimizing $H(Y|\xi)$. In other words, the maximum entropy sampling criteria is equivalent to the Bayesian optimal design criterion if the joint distribution of (Y, Θ) is not functionally dependent on the experiment ξ. By interchanging the role of Y and Θ in (7.4), it can be showed that if $\mathbb{E}_\Theta\{H(Y|\theta, \xi)\}$ does not depend on the design of t, then the term $H(Y, \Theta|\xi)$ neither. This reformulation is an easy way to check the assumption that allows the duality between the maximum entropy sampling criterion and the Bayesian optimal experiment design.

It is relevant to mention that the maximum entropy sampling criterion is widely used by the machine learning community for supervised learning tasks, under the name of Uncertainty Sampling [Lewis and Catlett, 1994, Zhu et al., 2008, Yang et al., 2015, Sharma and Bilgic, 2017, Nguyen et al., 2019].

7.2 Application to Hierarchical Stellar Systems

The present section introduces the methodology for determining a probability distribution of the optimal measurement time of a stellar system based on the maximum entropy sampling criterion.

Following the classical Bayesian experimental framework: Let $y \in Y$ and $\theta \in \Theta$ be the observations and parameters of a system, respectively. The Bayesian optimal design problem is to find the experiment $\xi \in \Xi$ that minimizes the average entropy of the posterior distribution $\mathbb{E}_Y\{H(\Theta|Y,\xi)\}$. The same framework can be used for determining the optimal measurement time in a hierarchical stellar system. Considering Y is the observation space, Θ is the orbital parameters space (that characterize the stellar system), and ξ is the time of measurement of the system (denoted as t hereinafter). Thereby, the optimal Bayesian design problem is to find the time of measurement t from a finite set S that maximizes the information of the system orbital parameters, or equivalently, that minimizes the posterior entropy $\mathbb{E}_Y\{H(\Theta|Y,t)\}$.

According to the Bayesian modeling of binary stellar systems introduced in Section 5, it is assumed that each of the observations distributes as a Gaussian with its mean determined by the Keplerian orbital model at a determined time t and its variance determined by the observation error. Therefore, let $Y = (Y_1, ..., Y_n)$ be a finite set of new observations of the stellar system that satisfies that:

$$Y_i|(\theta, t_i) = f(\theta, t_i) + \varepsilon_i, \tag{7.5}$$

where $\varepsilon_i \sim \mathcal{N}(0, \sigma_i^2)$, σ_i is the observation error and $f(\cdot)$ is the function that maps the times t_i into the observation space Y_i given the stellar system orbital parameters θ (according to the ephemerides formuale of the orbital Keplerian model).

It is additionally assumed that the random variables ε are independent of $f(\theta, t_i)$ and $\sigma_i = \sigma, \forall i \in [1, ..., n]$. Then, it is satisfied that $H(Y|\theta, t) = H(\varepsilon)$, i.e., $H(Y|\theta, t)$ is functionally independent from the choice of the measurement time t. The fulfillment of the additional hypothesis is a sufficient condition to prove the duality between the optimal Bayesian design problem and the maximum entropy sampling criterion [Sebastiani and Wynn, 2000] as discussed in Section 7.1. Therefore, using the equation (7.4) and noting that the assumption makes the term $H(Y, \Theta|t)$ fixed for any t, the time of measurement that maximizes the information of the orbital parameters of the hierarchical stellar system is the time $t^* \in S$ that maximizes the entropy in the observation space:

$$t^* = \arg\max_{t \in S} H(Y|t). \tag{7.6}$$

As each orbital parameter configuration $\theta \in \Theta$ induces orbits in the observational space with different periods P, a suitable selection of the set of measurement times S must be defined to ensure a full exploration of each system θ. In that manner, it is proposed to determine the optimal normalized time of measurement or phase $\tau = (t - t_0)/P$ using the maximum entropy sampling criteria on the first period of each orbital parameter configuration θ, instead of applying it directly on the original time of measurement t. t_0 is a fixed initial time of evaluation. Therefore, the problem reduces to:

$$\tau^* = \arg\min_{\tau \in [0,1]} \mathbb{E}_Y\{H(\Theta|Y,\tau)\} = \arg\max_{\tau \in [0,1]} H(Y|\tau), \tag{7.7}$$

which induces a probability distribution on the time of optimal measurement t^*:

$$t^* = \tau^* P + t_0. \tag{7.8}$$

The proposed time of measurement methodology based on the maximum entropy sampling criterion (instead of using the classical Bayesian optimal design approach) has three main advantages in the context of hierarchical stellar systems:

- The maximum entropy sampling criterion reduces the optimal measurement time estimation problem to compute the entropy in the observation space Y. This computation is at most 4-dimensional

(considering the 2-dimensional positional observations and the radial velocities observation of both the main and the companion object of the binary system) instead of computing the entropy in the orbital parameters space Θ that has tens of dimensions. This last space increases with the system multiplicity in hierarchical stellar systems. This simplification reduces the computational cost of the estimation of the entropy drastically.

- The reduction of the computational cost of the entropy estimation also allows to perform the algorithm in a dense array of normalized times $\tau \in [0, 1]$, allowing to determine the normalized time of maximum entropy in the observational space τ^* with high precision.
- The classical Bayesian optimal design approach requires estimating the posterior distribution incorporating virtual samples at each of the evaluated times $\tau_i, i \in [0, 1]$. This approach limits the set of times to evaluate to a few due to the high computational costs involved in the Bayesian inference procedure. In contrast, the maximum entropy sampling approach requires only one computation of the posterior distribution to compute the entropy of the projected posterior distribution into the observation space Y given the available observations $\mathcal{D}$ of the stellar system.

7.3 Experiments

The proposed methodology for determining the optimal measurement time based on the maximum entropy sampling criterion is performed on three binary systems studied in the previous chapters: the double-lined visual spectroscopic binary system YSC132AaAb, the single-lined visual spectroscopic binary system HIP99675, and the visual triple system LHS1070.

The entropy of the projection in the positional observation space (the orbit space X, Y) of 1000 randomly selected samples θ of the posterior distribution $p(\Theta|\mathcal{D})$ given the set of available observations of the stellar system $\mathcal{D}$, is computed along the valid range of the normalized time $\tau \in [0, 1]$, using the k-nearest neighbour entropy estimation method [Kozachenko and Leonenko, 1987]. Five representative normalized times $\tau_i, i \in [1, 2, 3, 4, 5]$ of the curve of estimated entropies of the observational space are selected, comparing the entropy values $H(X, Y|\mathcal{D}, \tau_i)$ with the observed projection of the posterior distribution at those times. To verify the duality between the maximum entropy sampling criterion and the Bayesian optimal design problem, new Bayesian inferences are performed to estimate the posterior distributions $p(\Theta|\mathcal{D} \cup \{d_{\tau_i}\}), i \in [1, 2, 3, 4, 5]$, with $d_{\tau_i} \sim \mathcal{N}(f(\theta_{MAP}, \tau_i), \sigma^2)$ a virtual observation determined by the Keplerian model on the times τ_i with fixed variance σ^2 for each of the times, considering the maximum a posteriori estimate of the posterior distribution θ_{MAP}. The virtual posterior distributions and its projection on the observation space are compared, and the estimated entropies of the virtual posterior distributions $H(\Theta|\mathcal{D} \cup \{d_{\tau_i}\})$ are compared with the respective entropies of the observation space $H(X, Y|\mathcal{D}, \tau_i)$. Finally, the normalized time of highest observation entropy is selected to generate the posterior distribution of the time of optimal measurement t^*.

7.3.1 YSC132AaAb

The computed posterior distribution in the observation space of the system YSC132AaAb and the corresponding estimated entropy curve $H(X, Y|\mathcal{D}, \tau)$ along the normalized times $\tau \in [0, 1]$, with t_0 equals to latest available observation in $\mathcal{D}$, are presented in Figure 7.1.

The estimated entropy curve shows two local maximum values with an accentuated local minimum in between and two minimum values at the beginning $\tau = 0$ and the end $\tau = 1$ of the curve. The behavior of the estimated entropy curve coincides with the observed dispersion of the posterior distribution projected in the observation space, where the times of minimum entropy are the corresponding zones in the orbit

populated with high precision observations ($\tau = 0$), presenting a narrow band of the projected orbits. The entropy increase with τ as well as the dispersion of the projected orbits as no precise observations are presented in that zone of the orbital space (τ_1). This behavior persists until reaching the first local maximum of the entropy, where the dispersion of the projected orbits is highest (τ_2) due to the scarce and low precise observations in that zone of the orbit. The entropy starts to decrease as well as the dispersion of the projected orbits until reaching a local minimum value (τ_3). It is interesting to note that the projected orbits become narrower (with lower uncertainty) even in the absence of positional observations near that zone, attributed to underlying uncertainty reduction of some of the orbital parameters through the observations in other zones of the orbit. The posterior entropy in that zone (τ_3) is higher than the beginning of the curve ($\tau = 0$). However, the projected orbit is narrower. This last phenomenon is attributed to the fact that the uncertainty of the projected orbit (τ_3) is not reflected in the width of the orbit but along the orbit, due to the uncertainty on the period P of the system. The estimated entropy curves increase again until reaching a second local maximum (τ_4), which is slightly higher than the previous local maximum (τ_2). The projected orbits reach the higher dispersion due to the total absence of observations in that orbital zone. Finally, the entropy curve starts to decrease at the end of the observational period ($\tau = 1$) reaching a similar minimum as the beginning of the curve ($\tau = 0$) due to the periodicity of the orbit.

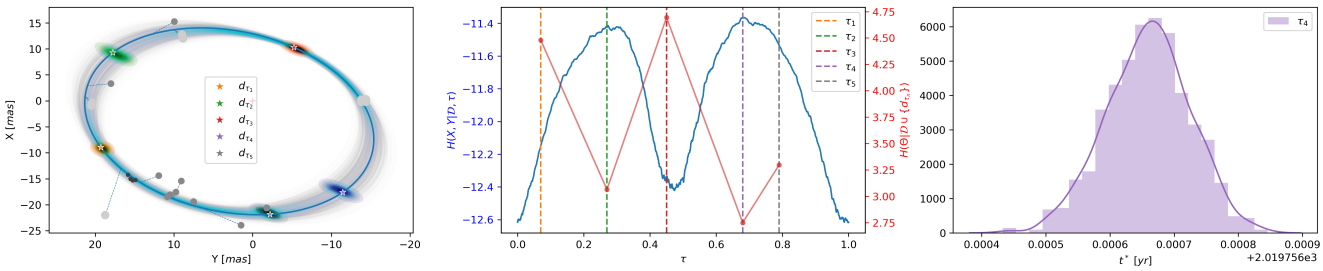

Figure 7.1: Optimal measurement time analysis on the stellar system YSC132AaAb. Left: Estimated entropy curve of the observation space $H(X, Y|\mathcal{D}, \tau)$ and estimated posterior distribution entropies $H(\Theta|\mathcal{D}\cup\{d_{\tau_i}\})$ incorporating different virtual observations d_{τ_i} at five representative times $\tau_i, i \in [1, 2, 3, 4, 5]$. Middle: Projection of the posterior distribution $p(X, Y|\mathcal{D})$ along $\tau \in [0, 1]$, highlighting the projection at the representative times $\tau_i, i \in [1, 2, 3, 4, 5]$. Right: Optimal time measurement distribution induced by the normalized time of highest entropy in the observation space $\tau^* = \arg\max_{\tau\in[0,1]} H(X, Y|\mathcal{D}, \tau)$.

The posterior distribution is updated through the incorporation of virtual observations d_{τ_i} to the set of available observations $\mathcal{D}$. Each virtual observation is generated by means of the projection of the maximum a posterior estimate of the posterior distribution θ_{MAP} in the observation space at the normalized times τ_i, with a fixed identically distributed stochastic uncertainty $\varepsilon_i \sim \mathcal{N}(0, \sigma^2)$, i.e., considering that the virtual observations distributes as $d_{\tau_i} \sim \mathcal{N}(f(\theta_{MAP}, \tau_i), \sigma^2)$, with $f(\cdot)$ a deterministic mapping determined by the Keplerian orbital model. The marginal updated posterior distribution by the incorporation of each sample d_{τ_i} are presented in Figure 7.2 and the corresponding projections on the observation space are presented in Figure 7.3. The comparison of the marginal distributions shows different magnitudes of the orbital parameters uncertainty, where the virtual observations d_{τ_i} that reaches the highest uncertainty reduction in orbital parameters space are the observation in which the corresponding entropy $H(X, Y|\mathcal{D}, \tau_i)$ is higher and viceversa. This (opposite) behavior is clearly reflected in the comparison between the estimated values of $H(X, Y|\mathcal{D}, \tau_i)$ and $H(\Theta|\mathcal{D}\cup\{d_{\tau_i}\})$ presented in the middle panel of the Figure 7.1, confirming the duality between the maximum entropy sampling criterion and the Bayesian optimal design problem. The projected updated posterior distributions show a considerable uncertainty reduction when the virtual measurement is taken in the highest uncertainty zones. On the other hand, the uncertainty reduction is negligible

when the virtual measurement was taken in the zones of lower uncertainty, satisfying the objective of the maximum entropy sampling principle.

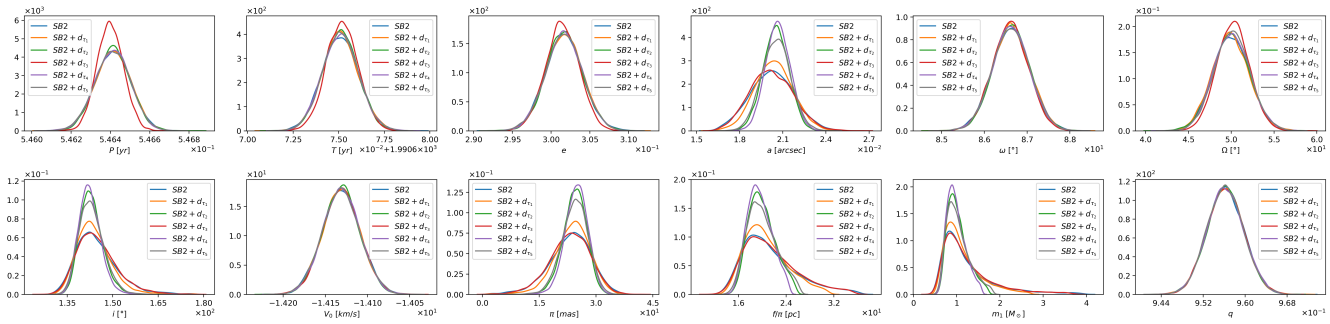

Figure 7.2: Marginal posterior distributions of the stellar system YSC132AaAb incorporating the virtual observations $d_{\tau_i} \sim \mathcal{N}(f(\theta_{MAP}, \tau_i), \sigma^2)$, with σ^2 fixed to the minimum variance of the system available observations $\mathcal{D}$.

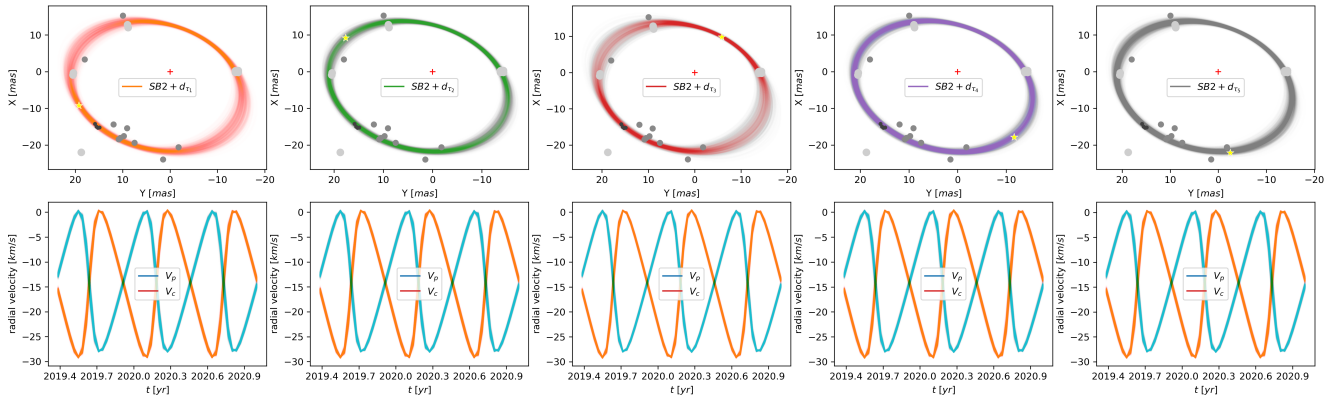

Figure 7.3: Projection of the posterior distribution of the stellar system YSC132AaAb in the observation space.

Finally, the distribution of the optimal times of measurement t^* at the normalized time of maximum uncertainty of the observation space (τ_4) is presented in the right panel of Figure 7.1. The estimated distribution has a clear defined Gaussian shape with extremely low dispersion. This result reflects that the period P of the stellar system is well determined due to the presence of abundant and precise radial velocities observations.

7.3.2 HIP99675

The computed posterior distribution in the observation space of the system HIP99675 and the corresponding estimated entropy curve $H(X, Y|\mathcal{D}, \tau)$ along the normalized times $\tau \in [0, 1]$, with t_0 equals to latest available observation in $\mathcal{D}$, are presented in Figure 7.4.

The estimated entropy curve has a symmetrical behavior around the point of minimum entropy (τ_3), where the entropy increases as far from that point ($\tau < \tau_3$ and $\tau > \tau_3$) in a wavy manner, presenting multiple local minimums and maximums but permanently preserving it average increasing behavior. The behavior of the estimated entropy curve apparently does not relate to the observed posterior orbits that form a diffuse cloud of lines with no clearly defined regions due to the scarce and low precise available positional observations. However, the projection of the posterior distribution at the five normalized times

selected from the estimated entropy curve $\tau_i, i \in [1,2,3,4,5]$ shows that the uncertainties are perpendicularly distributed with respect to the well-defined orbits, and the overlap of these orbits forms the previously not interpretable diffuse cloud. The zone in the orbit space with the lowest entropy (τ_3) coincides with most of the available positional observations being near that zone. In contrast, the opposite zone in the orbit presents the highest entropy (τ_1, τ_5). The wavy behavior of the entropy curve is reflected in the shape of the projection of the posterior distribution in the orbit space, where local minimums correspond to long but narrower projections (τ_4) and local maximums corresponds to slightly shorter but wider projections (τ_2).

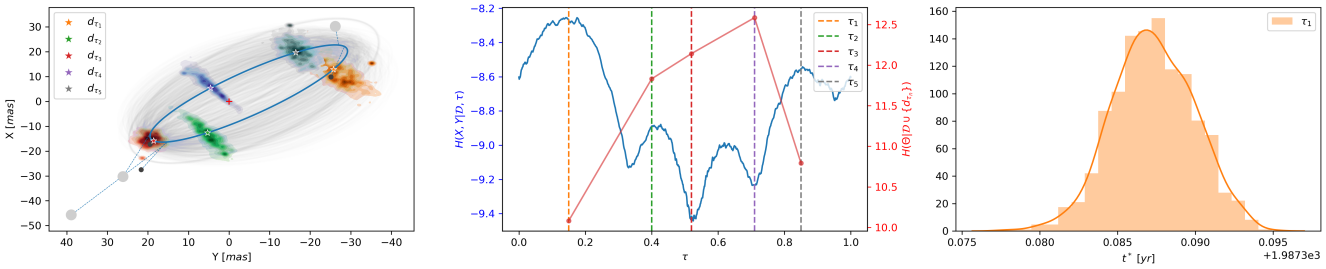

Figure 7.4: Optimal measurement time analysis on the stellar system HIP99675. Left: Estimated entropy curve of the observation space $H(X, Y|\mathcal{D}, \tau)$ and estimated posterior distribution entropies $H(\Theta|\mathcal{D} \cup \{d_{\tau_i}\})$ incorporating different virtual observations d_{τ_i} at five representative times $\tau_i, i \in [1,2,3,4,5]$. Middle: Projection of the posterior distribution $p(X, Y|\mathcal{D})$ along $\tau \in [0,1]$, highlighting the projection at the representative times $\tau_i, i \in [1,2,3,4,5]$. Right: Optimal time measurement distribution induced by the normalized time of highest entropy in the observation space $\tau^* = \arg\max_{\tau \in [0,1]} H(X, Y|\mathcal{D}, \tau)$.

Then, the posterior distribution is updated through the incorporation of virtual observations d_{τ_i} to the set of available observations $\mathcal{D}$. Each virtual observation is generated by means of the projection of the maximum a posterior estimate of the posterior distribution θ_{MAP} in the observation space at the normalized times τ_i, with a fixed identically distributed stochastic uncertainty $\varepsilon_i \sim \mathcal{N}(0, \sigma^2)$, i.e., considering that the virtual observations distributes as $d_{\tau_i} \sim \mathcal{N}(f(\theta_{MAP}, \tau_i), \sigma^2)$, with $f(\cdot)$ a deterministic mapping determined by the Keplerian orbital model. The marginal (updated) posterior distributions obtained by incorporating each sample d_{τ_i} are presented in Figure 7.5 an the corresponding projections on the observation space are presented in Figure 7.6. The posterior distribution updated with the virtual observation generated through the normalized time of lowest entropy (d_{τ_3}) presents an uncertainty reduction on some of the marginal distributions of the parameters (a, Ω), but a considerable increase of the uncertainty in other orbital parameters (i). The same trade-off between the uncertainty reduction and increase between some of the orbital parameters is observed in each case, inducing different projected orbits of the updated posterior distributions. We observe that the virtual observations incorporated reduce the projected uncertainty in zones of the orbit space near the measurements. For example, the d_{τ_3} observations highly reduce the uncertainty on the length of the orbit, coinciding with a reduction of the uncertainty of the updated marginal posterior distributions of a and Ω. However, this virtual observation does not provide any information about the inclination of the orbit, expressed by a high increase on the updated marginal posterior distributions of i. The computed entropy of the updated posterior distribution $H(\Theta|\mathcal{D} \cup \{d_{\tau_i}\})$ takes into account the mentioned trade-off between the increase and reduction of uncertainty of the posterior distribution along its different dimensions. Indeed, the comparison between the entropies $H(X, Y|\mathcal{D}, \tau_i)$ and $H(\Theta|\mathcal{D} \cup \{d_{\tau_i}\})$ presented in the middle panel of the Figure 7.4 shows that the duality between the maximum entropy sampling criterion and the Bayesian optimal design problem is preserved, as expected. However, an unexpected result that disagrees with the dual principle is observed in the pair (τ_3, τ_4), where $H(\Theta|\mathcal{D} \cup \{d_{\tau_3}\})$ should been greater than $H(\Theta|\mathcal{D} \cup \{d_{\tau_4}\})$, since $H(X, Y|\mathcal{D}, \tau_3)$

is lower than $H(X, Y|\mathcal{D}, \tau_4)$. This result is attributed to the estimation error of $H(\Theta|\mathcal{D} \cup \{d_{\tau_i}\})$ due to the high dimension of the parameter space Θ. Nevertheless, the observed discrepancy is minimal in magnitude, and all the other normalized observation times chosen satisfy the aforementioned duality principle.

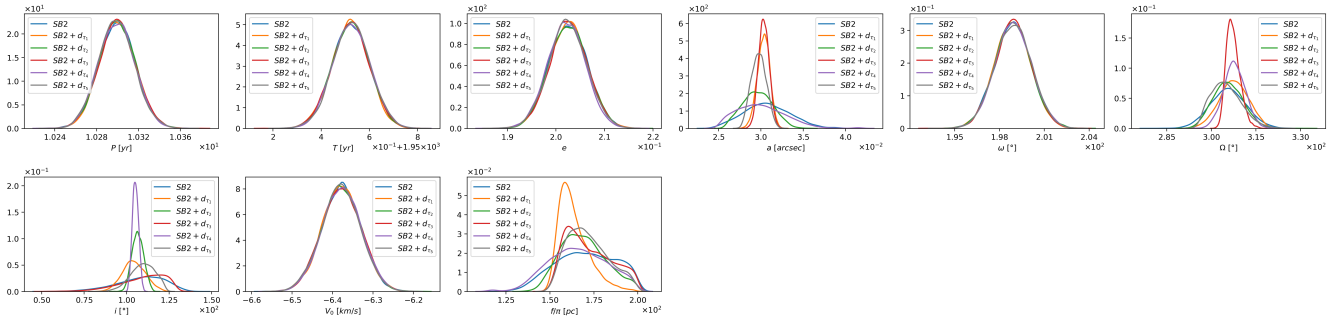

Figure 7.5: Marginal posterior distributions of the stellar system HIP99675 incorporating the virtual observations $d_{\tau_i} \sim \mathcal{N}(f(\theta_{MAP}, \tau_i), \sigma^2)$, with σ^2 fixed to 0.1 times the minimum variance of the system available observations $\mathcal{D}$.

The distribution of the optimal times of measurement t^* at the normalized time of maximum uncertainty of the observation space (τ_1) is presented in the right panel of Figure 7.4. The estimated distribution has a clearly defined Gaussian shape with an extremely low dispersion, which reflects that the period P of the stellar system is well determined due to the presence of abundant and precise radial velocities observations as in the system studied previously.

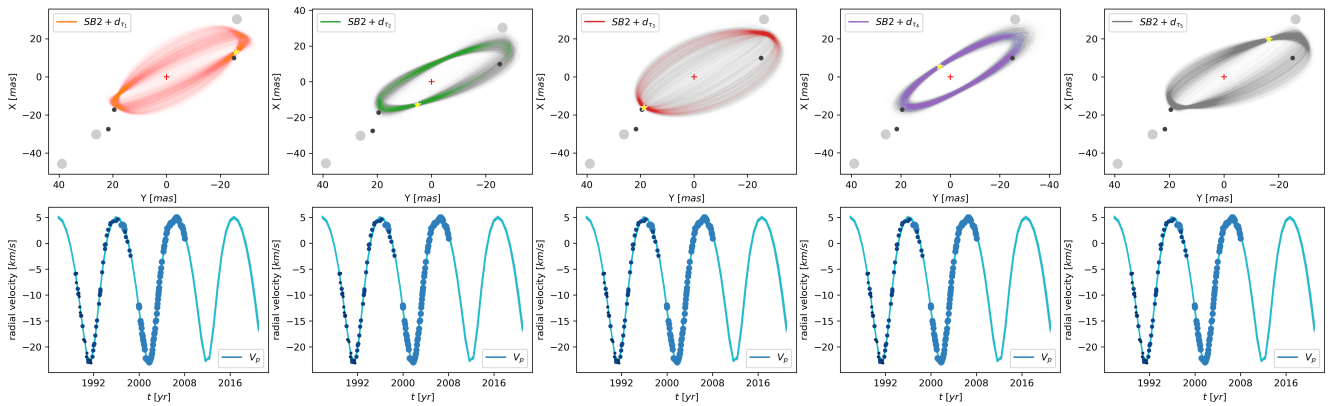

Figure 7.6: Projection of the posterior distribution of the stellar system HIP99675 in the observation space.

7.3.3 LHS1070

The computed posterior distribution in the observation space of the system HIP99675 and the corresponding estimated entropy curve $H(X, Y|\mathcal{D}, \tau)$ along the normalized times $\tau \in [0, 1]$, with t_0 equals to latest available observation in $\mathcal{D}$, are presented in Figure 7.7. Five representative normalized times $\tau_i, i \in [1, 2, 3, 4, 5]$ of the estimated entropy curve were selected for the analysis.

The estimated entropy curve presents a global minimum at the beginning $\tau = 0$, which coincides with the zone in the orbit populated with abundant and precise positional observations. The entropy curve increases steeply from the positional observations in the orbit space until reaching a maximum value at τ_3. The rest of the curve presents a constant value with a slight decrease in τ_4. The entropy curve presents

a clearly defined wavy behavior along all the normalized times τ evaluated. This behavior coincides with the lobes of the projected orbits in the observation space. The projection of the posterior distribution in the orbit space at the normalized times $\tau_i, i \in [1, 2, 3, 4, 5]$ shows that the posterior distributions increase its dispersion in zones far from the available observations, forming wavy patterns. This pattern can be attributed to the shift increase between the projected orbits due to the high uncertainty on the period P orbital parameter of the purely astrometric hierarchical system (without radial velocities observations).

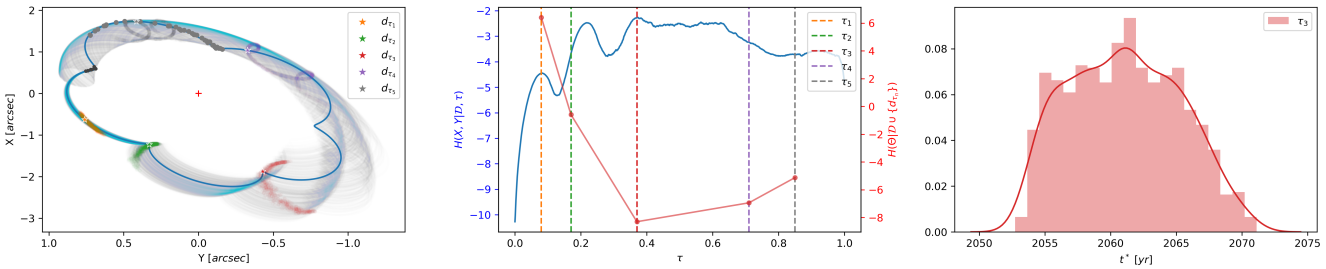

Figure 7.7: Optimal measurement time analysis on the stellar system LHS1070. Left: Estimated entropy curve of the observation space $H(X, Y|\mathcal{D}, \tau)$ and estimated posterior distribution entropies $H(\Theta|\mathcal{D} \cup \{d_{\tau_i}\})$ incorporating different virtual observations d_{τ_i} at five representative times $\tau_i, i \in [1, 2, 3, 4, 5]$. Middle: Projection of the posterior distribution $p(X, Y|\mathcal{D})$ along $\tau \in [0, 1]$, highlighting the projection at the representative times $\tau_i, i \in [1, 2, 3, 4, 5]$. Right: Optimal time measurement distribution induced by the normalized time of highest entropy in the observation space $\tau^* = \arg\max_{\tau \in [0,1]} H(X, Y|\mathcal{D}, \tau)$.

The posterior distribution is updated through the incorporation of virtual observations d_{τ_i} to the set of original observations $\mathcal{D}$. Each virtual observation is generated utilizing the maximum a posteriori estimate of the posterior distribution, i.e., θ_{MAP}, in the observation space at the normalized times τ_i, with a fixed identically distributed stochastic uncertainty $\varepsilon_i \sim \mathcal{N}(0, \sigma^2)$. More precisely, we consider that the virtual observations distributes as $d_{\tau_i} \sim \mathcal{N}(f(\theta_{MAP}, \tau_i), \sigma^2)$, with $f(\cdot)$ a deterministic mapping determined by the Keplerian orbital model. The updated posterior distribution (by incorporating the samples d_{τ_i}) is presented in Figure 7.8. The corresponding projections on the observation space are presented in Figure 7.9. The comparison of these marginal distributions shows different magnitudes of uncertainty of the orbital parameters of the outer orbit AB, where the virtual observations d_{τ_i} that reaches the highest uncertainty reduction in the orbital parameters space are the observations in which the corresponding entropy $H(X, Y|\mathcal{D}, \tau_i)$ is the highest and viceversa. This opposite behavior is clearly reflected in the comparison between the estimated values of $H(X, Y|\mathcal{D}, \tau_i)$ and $H(\Theta|\mathcal{D} \cup \{d_{\tau_i}\})$ presented in the middle panel of the Figure 7.1, which verifies the duality between the maximum entropy sampling criterion and the Bayesian optimal design problem. The updated posterior distributions show a considerable uncertainty reduction in the zones where the virtual measurement was taken in the zones of highest uncertainty. On the other hand, the uncertainty reduction is negligible when the virtual measurement was taken in the zones of lower uncertainty, satisfying the objective of the maximum entropy sampling principle.

The distribution of the optimal times of measurement t^* at the normalized time of maximum uncertainty of the observation space (τ_3) is presented in the right panel of the Figure 7.7. Unlike the previously studied cases, the estimated distribution presents a high dispersion, which incorporates the high uncertainty on the period P of the outer system AB due to the low coverage of positional observation in the orbit space and the absence of radial velocities observations.

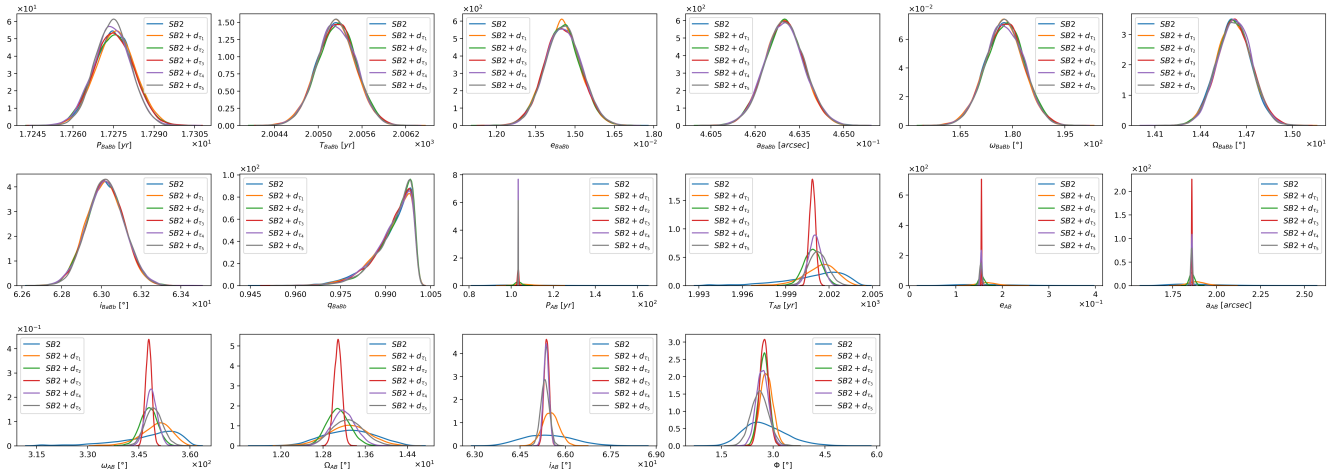

Figure 7.8: Marginal posterior distributions of the stellar system LHS1070 incorporating the virtual observations $d_{\tau_i} \sim \mathcal{N}(f(\theta_{MAP}, \tau_i), \sigma^2)$, with σ^2 fixed to the minimum variance of the system available observations $\mathcal{D}$.

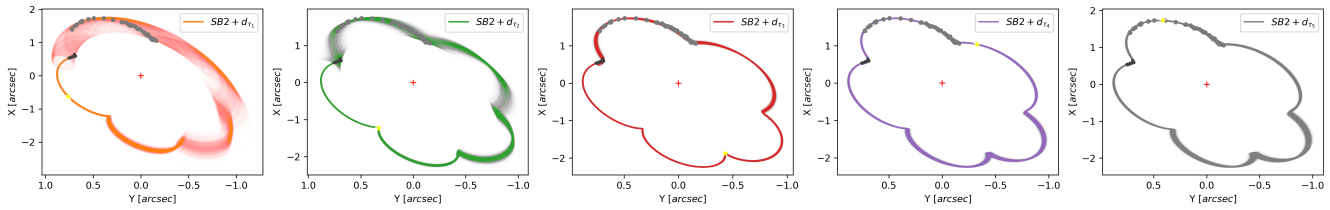

Figure 7.9: Projection of the posterior distribution of the stellar system LHS1070 in the observation space.

Chapter 8

Conclusions and Future Work

The proposed Bayesian methodology for the inference of the orbital parameters in binary and hierarchical stellar systems allows providing a computationally efficient, robust, and precise estimation of the corresponding joint posterior distribution. This inference is implemented by the No-U-Turn sampler Markov chain Monte Carlo algorithm, which allows the incorporation of prior information of the stellar system to constrain the inference in scenarios with inaccurate or missing data. The flexibility of this sampling scheme to add prior information was beneficial for the estimation of the individual masses in single-lined visual-spectroscopic binaries and the estimation of hierarchical stellar systems of any multiplicity and architecture.

An exhaustive experimental analysis was considered for the validation of the proposed methodology in single-lined visual-spectroscopic binaries. We study the quality of the obtained inference by comparing the estimated posterior distribution of well-determined double-lined visual-spectroscopic binaries with its single-lined visual spectroscopic counterpart by omitting the radial velocities observations of the companion object. Our results show a negligible difference between the estimated posterior distributions of the orbital parameters (and their uncertainties) compared to the benchmark case in which the radial velocities of both components are considered. The incorporation of prior information of the system (the parallax and the primary object mass) to estimate its individual masses was also studied. The empirical results show that the posterior distribution of the mass ratio of the system, and hence its individual masses, can be estimated with good precision. The incorporation of prior distributions makes those parameters identifiable, where the obtained estimations —position, dispersion and shape of the marginal posterior distribution— strongly depend on the prior chosen. This prior knowledge also has an influence on the estimated posterior distribution of the other orbital parameters. The impact of incorporating priors on the inference of previous-identifiable orbital parameters (that are already identifiable without the incorporation of the priors) directly depends on the abundance, precision, and orbital coverage of the available observations. In particular, it is observed that if the system is precisely determined, the impact of the prior on the estimation of those set of parameters is negligible.

On the specifics, the numerical results show that the lowest estimation error (from the average MAP of all the systems analyzed) on the system's mass ratio, compared to the full-information scenario (with both radial velocities observations), was achieved by the mixed prior case that incorporates a parallax and a primary object mass priors simultaneously (4.92%) and the highest error was obtained by the incorporation of a prior on only the primary object mass (7.44%), achieving a percentage error lower than 8%. However, it is shown that the greatest similitude of the marginal posterior distribution to the full-information scenario —in the KL-divergence sense— was achieved by incorporating a prior on the system

parallax alone. However, the lowest similitude was obtained by incorporating a prior on the primary object mass alone, which is attributed to the fact that the parallax prior information is more recent and precise than the primary object mass prior. The incorporation of both priors was even more beneficial to the accuracy of the MAP estimates where, as more information is provided, a better estimation can be obtained. The differences between the estimated posterior distributions in all the studied cases were also analyzed in the corresponding observations space, providing a better understanding of the effect of the different sources of information on the shape and uncertainty of the orbit and radial velocity curves of the stellar systems.

An analogous experimental setting was developed for the evaluation of the Bayesian inference in hierarchical stellar systems. Here, the flexibility for incorporating prior information of the system was crucial for the identifiability of the orbital parameters. The inherent complexity, diversity, and lack of information sources make this estimation task very challenging. As the phenomenological interaction between the orbital parameters of the hierarchical stellar system is much more complex than the binary case, the obtained posterior distributions present complex and diverse shapes. This is also observed on the observations space, where a wide variety of orbits of diverse shapes, sizes, and periods are probable, showing the relevance of the proposed Bayesian tool for a deeper analysis of those types of stellar systems, in contrast to the deterministic estimation of the orbital parameters widely used by the astronomical community.

A direct application of the estimated posterior distribution of the orbital parameters of the stellar systems was also explored: the determination of optimal time of observation, i.e., the time that mostly reduces the uncertainty of the orbital parameters that characterize the stellar system. The proposed Bayesian methodology allows to provide a temporal characterization of the uncertainty of the system. Then, the determination of the time ranges that provide the highest and lowest information about the system through the incorporation of a new observation can be addressed and provide a probability distribution of the optimal time of measurement. The theoretical foundations and practical advantages of the proposed methodology were extensively discussed in this thesis, showing that the method, based on the maximum entropy sampling principle, effectively solves the Bayesian optimal design problem, as well as showing its computational efficiency compared to a naive approach to solve the problem, allowing to compute an information gain curve in a continuous range of time that effectively characterizes the variability of the system. The proposed framework was validated in three different stellar systems, showing the suitability and capabilities of the method for the determination of the optimal time of measurement. We also show that the generated information gain curve of the stellar systems allows performing a deeper analysis on the uncertainty of the orbits and radial velocity curves of binary systems.

The present work addresses the orbital parameters estimation of binary and hierarchical stellar systems through a Bayesian perspective, emphasizing the importance of providing a complete posterior distribution characterization of the orbital parameters. This approach not only allows to provide an uncertainty quantification (as many classical optimization-based methods roughly provide), but it also allows to visualize the uncertainty of the system in the observation space. This last information is a relevant input for other statistical methods, such as the proposed optimal measurement time methodology. However, many other statistical-based applications of interest for the astronomical community can also be developed. In this context, a possible future work could be the formulation of a statistical hypothesis test for determining which object of a binary system is effectively the main object and which is the companion object, which is roughly determined by the astronomer without further foundation, as well as to explore other novel statistical applications that makes use of the posterior distribution of the orbital parameters of binary and hierarchical stellar systems. Another interesting future work could be the extension of the Bayesian methodology developed for hierarchical stellar systems to triple stellar systems without imposing any simplification of its dynamics, working directly on the solution of the 3-body problem. This approach could be helpful in the study of the multiple system's stability under the Bayesian approach.

Bibliography

[Abadi et al., 2016] Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., et al. (2016). Tensorflow: A system for large-scale machine learning. In *12th {USENIX} symposium on operating systems design and implementation ({OSDI} 16)*, pages 265–283.

[Aller et al., 1996] Aller, L., Appenzeller, I., Baschek, B., Butler, K., De Loore, C., Duerbeck, H., El Eid, M., Fink, H., Herczeg, T., Richtler, T., et al. (1996). Landolt-börnstein: Numerical data and functional relationships in science and technology-new series" gruppe/group 6 astronomy and astrophysics" volume 3 voigt: Astronomy and astrophysics. extension and supplement to volume 2" stars and star clusters. *lbor*, 3.

[Bingham et al., 2019] Bingham, E., Chen, J. P., Jankowiak, M., Obermeyer, F., Pradhan, N., Karaletsos, T., Singh, R., Szerlip, P., Horsfall, P., and Goodman, N. D. (2019). Pyro: Deep universal probabilistic programming. *The Journal of Machine Learning Research*, 20(1):973–978.

[Bouffanais and Porter, 2019] Bouffanais, Y. and Porter, E. K. (2019). Bayesian inference for binary neutron star inspirals using a hamiltonian monte carlo algorithm. *Physical Review D*, 100(10):104023.

[Brown et al., 2018] Brown, A., Vallenari, A., Prusti, T., De Bruijne, J., Babusiaux, C., Bailer-Jones, C., Biermann, M., Evans, D. W., Eyer, L., Jansen, F., et al. (2018). Gaia data release 2-summary of the contents and survey properties. *Astronomy & astrophysics*, 616:A1.

[Burgasser et al., 2012] Burgasser, A. J., Luk, C., Dhital, S., Gagliuffi, D. B., Nicholls, C. P., Prato, L., West, A. A., and Lépine, S. (2012). Discovery of a very low mass triple with late-m and t dwarf components: Lp 704-48/sdss j0006–0852ab. *The Astrophysical Journal*, 757(2):110.

[Carpenter et al., 2017] Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P., and Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of statistical software*, 76(1):1–32.

[Cid Palacios, 1958] Cid Palacios, R. (1958). On the necessary and sufficient observations for determination of elliptic orbits in double stars. *AJ*, 63:395.

[Cover, 1999] Cover, T. M. (1999). *Elements of information theory*. John Wiley & Sons.

[Czekala et al., 2017] Czekala, I., Andrews, S. M., Torres, G., Rodriguez, J. E., Jensen, E. L., Stassun, K. G., Latham, D. W., Wilner, D. J., Gully-Santiago, M. A., Grankin, K. N., et al. (2017). The architecture of the gw ori young triple-star system and its disk: Dynamical masses, mutual inclinations, and recurrent eclipses. *The Astrophysical Journal*, 851(2):132.

[Docobo, 1985] Docobo, J. (1985). On the analytic calculation of visual double star orbits. *Celestial Mechanics*, 36(2):143–153.

[Docobo et al., 1992] Docobo, J., Ling, J., and Prieto, C. (1992). Adaptation of docobo's method for the calculation of orbits of spectroscopic-interferometric binaries with mixed data. In *International Astronomical Union Colloquium*, volume 135, pages 220–222. Cambridge University Press.

[Docobo et al., 2008] Docobo, J., Tamazian, V., Balega, Y., Andrade, M., Schertl, D., Weigelt, G., Campo, P., and Palacios, M. (2008). A methodology for studying physical and dynamical properties of multiple stars. application to the system of red dwarfs gl 22. *Astronomy & Astrophysics*, 478(1):187–191.

[Docobo et al., 2018] Docobo, J. A., Tamazian, V. S., Campo, P. P., and Piccotti, L. (2018). Visual orbit and individual masses of the single-lined spectroscopic binary 94 aqr a (hd 219834a; mca 74). *The Astronomical Journal*, 156(3):85.

[Ford, 2005] Ford, E. B. (2005). Quantifying the uncertainty in the orbits of extrasolar planets. *The Astronomical Journal*, 129(3):1706.

[Geman and Geman, 1984] Geman, S. and Geman, D. (1984). Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE Transactions on pattern analysis and machine intelligence*, (6):721–741.

[Gregory, 2005] Gregory, P. (2005). A bayesian analysis of extrasolar planet data for hd 73526. *The Astrophysical Journal*, 631(2):1198.

[Gregory, 2011] Gregory, P. C. (2011). Bayesian exoplanet tests of a new method for mcmc sampling in highly correlated model parameter spaces. *Monthly Notices of the Royal Astronomical Society*, 410(1):94–110.

[Habets and Heintze, 1981] Habets, G. and Heintze, J. (1981). Empirical bolometric corrections for the main-sequence. *Astronomy and Astrophysics Supplement Series*, 46:193–237.

[Hajian, 2007] Hajian, A. (2007). Efficient cosmological parameter estimation with hamiltonian monte carlo technique. *Physical Review D*, 75(8):083525.

[Halbwachs et al., 2016] Halbwachs, J.-L., Boffin, H., Le Bouquin, J.-B., Kiefer, F., Famaey, B., Salomon, J.-B., Arenou, F., Pourbaix, D., Anthonioz, F., Grellmann, R., et al. (2016). Masses of the components of sb2s observed with gaia–ii. masses derived from pionier interferometric observations for gaia validation. *Monthly Notices of the Royal Astronomical Society*, 455(3):3303–3311.

[Hastings, 1970] Hastings, W. K. (1970). Monte carlo sampling methods using markov chains and their applications.

[Heintz, 1984] Heintz, W. (1984). The triple star zeta aquarii. *The Astrophysical Journal*, 284:806–809.

[Heintz, 1975] Heintz, W. D. (1975). Double stars. In *Astronomy a Handbook*, pages 472–486. Springer.

[Hoffman et al., 2014] Hoffman, M. D., Gelman, A., et al. (2014). The no-u-turn sampler: adaptively setting path lengths in hamiltonian monte carlo. *J. Mach. Learn. Res.*, 15(1):1593–1623.

[Hou et al., 2012] Hou, F., Goodman, J., Hogg, D. W., Weare, J., and Schwab, C. (2012). An affine-invariant sampler for exoplanet fitting and discovery in radial velocity data. *The Astrophysical Journal*, 745(2):198.

[Hummel et al., 1994] Hummel, C., Armstrong, J., Quirrenbach, A., Buscher, D., Mozurkewich, D., Elias, N., and Wilson, R. (1994). Very high precision orbit of capella by long baseline interferometry. *The Astronomical Journal*, 107:1859–1867.

[Ji et al., 2017] Ji, Y., Banks, T., Budding, E., and Rhodes, M. (2017). An investigation into exoplanet transits and uncertainties. *Astrophysics and Space Science*, 362(6):1–12.

[Köhler et al., 2012] Köhler, R., Ratzka, T., and Leinert, C. (2012). Orbits and masses in the multiple system lhs 1070. *Astronomy & Astrophysics*, 541:A29.

[Kozachenko and Leonenko, 1987] Kozachenko, L. and Leonenko, N. N. (1987). Sample estimate of the entropy of a random vector. *Problemy Peredachi Informatsii*, 23(2):9–16.

[Lewis and Catlett, 1994] Lewis, D. D. and Catlett, J. (1994). Heterogeneous uncertainty sampling for supervised learning. In *Machine learning proceedings 1994*, pages 148–156. Elsevier.

[Liu and Nocedal, 1989] Liu, D. C. and Nocedal, J. (1989). On the limited memory bfgs method for large scale optimization. *Mathematical programming*, 45(1):503–528.

[Lucy, 2014] Lucy, L. (2014). Mass estimates for visual binaries with incomplete orbits. *Astronomy & Astrophysics*, 563:A126.

[Lucy, 2018] Lucy, L. (2018). Binary orbits from combined astrometric and spectroscopic data. *Astronomy & Astrophysics*, 618:A100.

[MacKnight and Horch, 2004] MacKnight, M. and Horch, E. (2004). Calculating visual binary star orbits with the downhill simplex algorithm (amoeba). *AAS*, 204:07–19.

[Mendez et al., 2017] Mendez, R. A., Claveria, R. M., Orchard, M. E., and Silva, J. F. (2017). Orbits for 18 visual binaries and two double-line spectroscopic binaries observed with hrcam on the ctio soar 4 m telescope, using a new bayesian orbit code based on markov chain monte carlo. *The Astronomical Journal*, 154(5):187.

[Metropolis et al., 1953] Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E. (1953). Equation of state calculations by fast computing machines. *The journal of chemical physics*, 21(6):1087–1092.

[Morbey, 1975] Morbey, C. (1975). A synthesis of the solutions of spectroscopic and visual binary orbits. *Publications of the Astronomical Society of the Pacific*, 87(519):689.

[Moré, 1978] Moré, J. J. (1978). The levenberg-marquardt algorithm: implementation and theory. In *Numerical analysis*, pages 105–116. Springer.

[Muterspaugh et al., 2010] Muterspaugh, M. W., Hartkopf, W. I., Lane, B. F., O'Connell, J., Williamson, M., Kulkarni, S., Konacki, M., Burke, B. F., Colavita, M., Shao, M., et al. (2010). The phases differential astrometry data archive. ii. updated binary star orbits and a long period eclipsing binary. *The Astronomical Journal*, 140(6):1623.

[Neal et al., 2011] Neal, R. M. et al. (2011). Mcmc using hamiltonian dynamics. *Handbook of markov chain monte carlo*, 2(11):2.

[Nelson et al., 2013] Nelson, B., Ford, E. B., and Payne, M. J. (2013). Run dmc: an efficient, parallel code for analyzing radial velocity observations using n-body integrations and differential evolution markov chain monte carlo. *The Astrophysical Journal Supplement Series*, 210(1):11.

[Nguyen et al., 2019] Nguyen, V.-L., Destercke, S., and Hüllermeier, E. (2019). Epistemic uncertainty sampling. In *International Conference on Discovery Science*, pages 72–86. Springer.

[Norris, 1998] Norris, J. R. (1998). *Markov chains*. Number 2. Cambridge university press.

[Pourbaix, 1994] Pourbaix, D. (1994). A trial-and-error approach to the determination of the orbital parameters of visual binaries. *Astronomy and Astrophysics*, 290:682–691.

[Pourbaix, 1998] Pourbaix, D. (1998). Simultaneous least-squares adjustment of visual and spectroscopic observations of binary stars. *Astronomy and Astrophysics Supplement Series*, 131(2):377–382.

[Prusti et al., 2016] Prusti, T., De Bruijne, J., Brown, A. G., Vallenari, A., Babusiaux, C., Bailer-Jones, C., Bastian, U., Biermann, M., Evans, D., Eyer, L., et al. (2016). The gaia mission. *Astronomy & Astrophysics*, 595:A1.

[Ross et al., 1996] Ross, S. M., Kelly, J. J., Sullivan, R. J., Perry, W. J., Mercer, D., Davis, R. M., Washburn, T. D., Sager, E. V., Boyce, J. B., and Bristow, V. L. (1996). *Stochastic processes*, volume 2. Wiley New York.

[Sahlmann et al., 2013] Sahlmann, J., Lazorenko, P., Ségransan, D., Martín, E., Queloz, D., Mayor, M., and Udry, S. (2013). Astrometric orbit of a low-mass companion to an ultracool dwarf. *Astronomy & Astrophysics*, 556:A133.

[Salvatier et al., 2016] Salvatier, J., Wieckiâ, T. V., and Fonnesbeck, C. (2016). Pymc3: Python probabilistic programming framework. *Astrophysics Source Code Library*, pages ascl–1610.

[Sebastiani and Wynn, 2000] Sebastiani, P. and Wynn, H. P. (2000). Maximum entropy sampling and optimal bayesian experimental design. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 62(1):145–157.

[Shabram et al., 2020] Shabram, M. I., Batalha, N., Thompson, S. E., Hsu, D. C., Ford, E. B., Christiansen, J. L., Huber, D., Berger, T., Catanzarite, J., Nelson, B. E., et al. (2020). Sensitivity analyses of exoplanet occurrence rates from kepler and gaia. *The Astronomical Journal*, 160(1):16.

[Sharma and Bilgic, 2017] Sharma, M. and Bilgic, M. (2017). Evidence-based uncertainty sampling for active learning. *Data Mining and Knowledge Discovery*, 31(1):164–202.

[Shewry and Wynn, 1987] Shewry, M. C. and Wynn, H. P. (1987). Maximum entropy sampling. *Journal of applied statistics*, 14(2):165–170.

[Straižys and Kuriliene, 1981] Straižys, V. and Kuriliene, G. (1981). Fundamental stellar parameters derived from the evolutionary tracks. *Astrophysics and Space Science*, 80(2):353–368.

[Thiele, 1883] Thiele, T. N. (1883). Neue methode zur berechung von doppelsternbahnen. *Astronomische Nachrichten*, 104:245.

[Tokovinin, 1992] Tokovinin, A. (1992). The frequency of low-mass companions to k and m stars in the solar neighbourhood. *Astronomy and Astrophysics*, 256:121–132.

[Tokovinin, 2016] Tokovinin, A. (2016). The triple system zeta aquarii. *The Astrophysical Journal*, 831(2):151.

[Tokovinin, 2018] Tokovinin, A. (2018). Dancing twins: Stellar hierarchies that formed sequentially? *The Astronomical Journal*, 155(4):160.

[Tokovinin and Latham, 2017] Tokovinin, A. and Latham, D. W. (2017). Relative orbit orientation in several resolved multiple systems. *The Astrophysical Journal*, 838(1):54.

[Tokovinin et al., 2015] Tokovinin, A., Latham, D. W., and Mason, B. D. (2015). The unusual quadruple system hd 91962 with a "planetary" architecture. *The Astronomical Journal*, 149(6):195.

[Tran et al., 2016] Tran, D., Kucukelbir, A., Dieng, A. B., Rudolph, M., Liang, D., and Blei, D. M. (2016). Edward: A library for probabilistic modeling, inference, and criticism. *arXiv preprint arXiv:1610.09787*.

[Villegas et al., 2021] Villegas, C., Mendez, R. A., Silva, J. F., and Orchard, M. E. (2021). Bayes-based orbital elements estimation in triple hierarchical stellar systems. *arXiv preprint arXiv:2105.05258*.

[Wang et al., 2009] Wang, Q., Kulkarni, S. R., and Verdú, S. (2009). Divergence estimation for multidimensional densities via k-nearest-neighbor distances. *IEEE Transactions on Information Theory*, 55(5):2392–2405.

[Wenger et al., 2000] Wenger, M., Ochsenbein, F., Egret, D., Dubois, P., Bonnarel, F., Borde, S., Genova, F., Jasniewicz, G., Laloë, S., Lesteven, S., and Monier, R. (2000). The SIMBAD astronomical database. The CDS reference database for astronomical objects. , 143:9–22.

[Yang et al., 2015] Yang, Y., Ma, Z., Nie, F., Chang, X., and Hauptmann, A. G. (2015). Multi-class active learning by uncertainty sampling with diversity maximization. *International Journal of Computer Vision*, 113(2):113–127.

[Ypma, 1995] Ypma, T. J. (1995). Historical development of the newton–raphson method. *SIAM review*, 37(4):531–551.

[Zhu et al., 2008] Zhu, J., Wang, H., Yao, T., and Tsou, B. K. (2008). Active learning with sampling by uncertainty and density for word sense disambiguation and text classification. In *Proceedings of the 22nd International Conference on Computational Linguistics (Coling 2008)*, pages 1137–1144.

Appendix A

Implementation Details

A.1 Gradient Computation

The implementation of the proposed Bayesian inference methodology based on the No-U-Turn sampler algorithm requires computing the gradient of the posterior function, which can be analytically derived from the Keplerian model formulae presented in Section 2.1 by taking the partial derivatives relative to each orbital parameter that characterizes the binary/hierarchical stellar system. A special consideration must be taken with the partial derivatives of the eccentric anomaly E, since it is not analytically calculated but numerically approximated. The non-zero partial derivatives of the eccentric anomaly can be expressed as a function of the variable itself as follows:

$$\begin{aligned} \frac{\partial E}{\partial e} &= \frac{\sin(E)}{1 - e\cos(E)}, \\ \frac{\partial E}{\partial T} &= -\frac{2\pi}{P} \cdot \frac{1}{1 - e\cos(E)}, \\ \frac{\partial E}{\partial P} &= -\frac{2\pi(t-T)}{P^2} \cdot \frac{1}{1 - e\cos(E)}, \end{aligned} \tag{A.1}$$

with the value of E previously approximated by any numerical method (such as the Newton-Raphson method [Ypma, 1995]).

A.2 Programming Language and Code

Many probabilistic programming languages already implement the No-U-Turn sampler algorithm [Tran et al., 2016, Salvatier et al., 2016, Abadi et al., 2016, Bingham et al., 2019], but the probabilistic programming language used in the current work for the implementation of the proposed Bayesian inference methodology in binary and hierarchical stellar systems was the *Stan* language [Carpenter et al., 2017], since implements the No-U-Turn sampler algorithm efficiently through automatic differentiation, only requiring to specify the probabilistic model on which the inference it is made.

The implementation of the probabilistic models in *Stan* language of the double-lined visual-spectroscopic binary system, single-lined visual-spectroscopic binary system and triple hierarchical system used in the current work are presented respectively in Listing A.1, Listing A.2 and Listing A.3, while the Kepler equation implementation within its analytical partial derivatives are presented in Listing A.4.

```
functions {
    real kepler(real M, real e);
    vector[] orbit(int N, vector t, real T, real P, real e, real a, real w, real Omega, real i) {
        // Variables declaration
        real M; real A; real B; real F; real G;
        vector[N] E; vector[N] x; vector[N] y; vector[N] pos[2];
        // Iterate over epochs
        for (j in 1:N) {
            // Mean anomaly
            M = 2 * pi() * (t[j] - T) / P;
            // Eccentric anomaly
            E[j] = kepler(M, e);
        }
        // Auxiliary normalized coordinates
        x = cos(E) - e;
        y = sin(E) * sqrt(1 - e^2);
        // Thiele-Innes constants
        A = a * (cos(w) * cos(Omega) - sin(w) * sin(Omega) * cos(i));
        B = a * (cos(w) * sin(Omega) + sin(w) * cos(Omega) * cos(i));
        F = a * (-sin(w) * cos(Omega) - cos(w) * sin(Omega) * cos(i));
        G = a * (-sin(w) * sin(Omega) + cos(w) * cos(Omega) * cos(i));
        // Apparent orbit
        pos[1] = A * x + F * y;
        pos[2] = B * x + G * y;
        return pos;
    }
    vector radial_velocity(int N, vector t, real T, real P, real e, real a, real w, real i, real V0, real plx,
                           real q, int primary) {
        // Variables declaration
        real M; real K; vector[N] E; vector[N] nu; vector[N] V;
        real convCoeff = 149597870.660 / (365.25 * 86400.0); // [AU] to [arcsec]
        // Iterate over epochs
        for (j in 1:N) {
            // Mean anomaly
            M = 2 * pi() * (t[j] - T) / P;
            // Eccentric anomaly
            E[j] = kepler(M, e);
        }
        // True anomaly
        nu = 2 * atan(sqrt((1 + e) / (1 - e)) * tan(E / 2));
        // Primary object radial velocity
        if (primary) {
            // Amplitude
            K = 2 * pi() * (a / plx) * (q / (1+q)) * convCoeff * sin(i) / (P * sqrt(1 - e^2));
            // Radial velocity
            V = V0 + K * (cos(w + nu) + e * cos(w));
        }
        // Companion object radial velocity
        else {
            // Amplitude
            K = 2 * pi() * (a / plx) * (1 / (1+q)) * convCoeff * sin(i) / (P * sqrt(1 - e^2));
            // Radial velocity
            V = V0 - K * (cos(w + nu) + e * cos(w));
        }
        return V;
    }
}
data {
    int<lower=0> N;          // number of samples
    int<lower=0> N1;         // number of samples
    int<lower=0> N2;         // number of samples
    vector[N]    t_as;       // epochs (years)
    vector[N1]   t_rv1;      // rv1 epochs (years)
    vector[N2]   t_rv2;      // rv2 epochs (years)
    vector[N]    x_obs;      // x coordinate observation
    vector[N]    y_obs;      // y coordinate observation
    vector[N]    pos_err;    // observation uncertainties
    vector[N1]   rv1_obs;    // rv1 observations
    vector[N1]   rv1_err;    // rv1 observation uncertainties
    vector[N2]   rv2_obs;    // rv2 observations
    vector[N2]   rv2_err;    // rv2 observation uncertainties
}
transformed data {
    real min_t     = fmin(min(t_as), fmin(min(t_rv1), min(t_rv2)));   // first epoch
    vector[N] t0   = t_as - min_t;                                    // normalized rv1 epochs
    vector[N1] t10 = t_rv1 - min_t;                                   // normalized rv1 epochs
    vector[N2] t20 = t_rv2 - min_t;                                   // normalized rv2 epochs
}
parameters {
    real<lower=0, upper=1>        T0;       // normalized time of periastron passage (T - t0) / P
    real                          log_P;    // ln(period)
    real<lower=0, upper=1>        e;        // eccentricity
    real<lower=0>                 a;        // semimajor axis (secods of arc)
    real<lower=0, upper=2*pi()>   w;        // argument of periapsis (rad)
    real<lower=0, upper=2*pi()>   Omega;    // longitude of the ascending node (rad)
    real<lower=0, upper=pi()>     i;        // inclination (rad)
    real                          V0;       // velocity of center of mass (km/s)
    real<lower=0>                 plx;      // parallax (seconds of arc)
    real<lower=0, upper=1>        q;        // mass ratio
}
transformed parameters {
    real P = exp(log_P);   // period
}
model {
    // Variables declaration
```

```
    real T; vector[N] pos[2]; vector[N1] V1; vector[N2] V2;
    // Period
    T = T0 * P;
    // Orbit
    pos = orbit(N, t0, T, P, e, a, w, Omega, i);
    x_obs ~ normal(pos[1], pos_err);
    y_obs ~ normal(pos[2], pos_err);
    // Primary object radial velocity
    V1 = radial_velocity(N1, t10, T, P, e, a, w, i, V0, plx, q, 1);
    rv1_obs ~ normal(V1, rv1_err);
    // Companion object radial velocity
    V2 = radial_velocity(N2, t20, T, P, e, a, w, i, V0, plx, q, 0);
    rv2_obs ~ normal(V2, rv2_err);
}
generated quantities {
    real w_deg     = w * 180 / pi();                    // argument of periapsis (degrees)
    real Omega_deg = Omega * 180 / pi();                // longitude of the ascending node (degrees)
    real i_deg     = i * 180 / pi();                    // inclination (degrees)
    real T         = T0 * P + min_t;                    // unnormalized time of periastron passage
    real m1        = (a / plx)^3 * 1 / (P^2 * (1 + q)); // primary object mass (solar masses)
    real m2        = q * m1;                            // companion object mass (solar masses)
    real plx_mas   = plx * 1000;                        // parallax (miliseconds of arc)
    real alpha     = (1 / plx) * q / (1 + q);
}
```

Listing A.1: Double-lined visual-spectroscopic binary probabilistic model in Stan.

```
functions {
    real kepler(real M, real e);
    vector[] orbit(int N, vector t, real T, real P, real e, real a, real w, real Omega, real i) {
        // Variables declaration
        real M; real A; real B; real F; real G;
        vector[N] E; vector[N] x; vector[N] y; vector[N] pos[2];
        // Iterate over epochs
        for (j in 1:N) {
            // Mean anomaly
            M = 2 * pi() * (t[j] - T) / P;
            // Eccentric anomaly
            E[j] = kepler(M, e);
        }
        // Auxiliary normalized coordinates
        x = cos(E) - e;
        y = sin(E) * sqrt(1 - e^2);
        // Thiele-Innes constants
        A = a * (cos(w) * cos(Omega) - sin(w) * sin(Omega) * cos(i));
        B = a * (cos(w) * sin(Omega) + sin(w) * cos(Omega) * cos(i));
        F = a * (-sin(w) * cos(Omega) - cos(w) * sin(Omega) * cos(i));
        G = a * (-sin(w) * sin(Omega) + cos(w) * cos(Omega) * cos(i));
        // Apparent orbit
        pos[1] = A * x + F * y;
        pos[2] = B * x + G * y;
        return pos;
    }
    vector radial_velocity(int N, vector t, real T, real P, real e, real a, real w, real i, real V0, real alpha) {
        // Variables declaration
        real M; real K; vector[N] E; vector[N] nu; vector[N] V;
        real convCoeff = 149597870.660 / (365.25 * 86400.0); // [AU/yr] to [km/s]
        // Iterate over epochs
        for (j in 1:N) {
            // Mean anomaly
            M = 2 * pi() * (t[j] - T) / P;
            // Eccentric anomaly
            E[j] = kepler(M, e);
        }
        // True anomaly
        nu = 2 * atan(sqrt((1 + e) / (1 - e)) * tan(E / 2));
        // Amplitude
        K = 2 * pi() * a * alpha * convCoeff * sin(i) / (P * sqrt(1 - e^2));
        // Radial velocity
        V = V0 + K * (cos(w + nu) + e * cos(w));
        return V;
    }
}
data {
    int<lower=0> N;          // number of samples
    int<lower=0> N1;         // number of samples
    vector[N]    t_as;       // epochs (years)
    vector[N1]   t_rv1;      // epochs (years)
    vector[N]    x_obs;      // x coordinate observation
    vector[N]    y_obs;      // y coordinate observation
    vector[N]    pos_err;    // observation uncertainties
    vector[N1]   rv1_obs;    // rv observations
    vector[N1]   rv1_err;    // rv observation uncertainties
    real<lower=0> plx_obs;   // parallax observation
    real<lower=0> plx_err;   // parallax observation uncertainty
    real<lower=0> m1_obs;    // m1 observation
    real<lower=0> m1_err;    // m1 observation uncertainty
}
transformed data {
    real min_t     = fmin(min(t_as), min(t_rv1));   // first epoch
    vector[N] t0   = t_as - min_t;                  // normalized as epochs
    vector[N1] t10 = t_rv1 - min_t;                 // normalized rv1 epochs
}
parameters {
```

```
    real<lower=0, upper=1>          T0;       // normalized time of periastron passage (T - t0) / P
    real                            log_P;    // ln(period)
    real<lower=0, upper=1>          e;        // eccentricity
    real<lower=0>                   a;        // semimajor axis (secods of arc)
    real<lower=0, upper=2*pi()>     w;        // argument of periapsis (rad)
    real<lower=0, upper=2*pi()>     Omega;    // longitude of the ascending node (rad)
    real<lower=0, upper=pi()>       i;        // inclination (rad)
    real                            V0;       // velocity of center of mass (km/s)
    real<lower=0, upper=1>          q;        // mass ratio
    real<lower=0, upper=200>        alpha;    // ---
}
transformed parameters {
    real P   = exp(log_P);                              // period (years)
    real plx = q / (1 + q) * 1 / alpha;                 // parallax (seconds of arc)
    real m1  = (a / plx)^3 * 1 / (P^2 * (1 + q));       // primary object mass (solar masses)
}
model {
    // Variables declaration
    real T; vector[N] pos[2]; vector[N1] V1;
    // Period
    T = T0 * P;
    // Orbit
    pos = orbit(N, t0, T, P, e, a, w, Omega, i);
    x_obs ~ normal(pos[1], pos_err);
    y_obs ~ normal(pos[2], pos_err);
    // Primary object radial velocity
    V1 = radial_velocity(N1, t10, T, P, e, a, w, i, V0, alpha);
    rv1_obs ~ normal(V1, rv1_err);
    // Parallax prior
    if (plx_obs > 0) {
        plx_obs ~ normal(plx, plx_err);
    }
    // Primary object mass prior
    if (m1_obs > 0) {
        m1_obs ~ normal(m1, m1_err);
    }
}
generated quantities {
    real w_deg     = w * 180 / pi();        // argument of periapsis (degrees)
    real Omega_deg = Omega * 180 / pi();    // longitude of the ascending node (degrees)
    real i_deg     = i * 180 / pi();        // inclination (degrees)
    real T         = T0 * P + min_t;        // unnormalized time of periastron passage
    real m2        = q * m1;                // companion object mass (solar masses)
    real plx_mas   = plx * 1000;            // parallax (miliseconds of arc)
}
```

Listing A.2: Single-lined visual-spectroscopic binary probabilistic model in Stan.

```
functions {
    real kepler(real M, real e);

    vector[] orbit(int N, vector t, real T, real P, real e, real a, real w, real Omega, real i) {
        // Variables declaration
        real M; real A; real B; real F; real G;
        vector[N] E; vector[N] x; vector[N] y; vector[N] pos[2];
        // Iterate over epochs
        for (j in 1:N) {
            // Mean anomaly
            M = 2 * pi() * (t[j] - T) / P;
            // Eccentric anomaly
            E[j] = kepler(M, e);
        }
        // Auxiliary normalized coordinates
        x = cos(E) - e;
        y = sin(E) * sqrt(1 - e^2);
        // Thiele-Innes constants
        A = a * (cos(w) * cos(Omega) - sin(w) * sin(Omega) * cos(i));
        B = a * (cos(w) * sin(Omega) + sin(w) * cos(Omega) * cos(i));
        F = a * (-sin(w) * cos(Omega) - cos(w) * sin(Omega) * cos(i));
        G = a * (-sin(w) * sin(Omega) + cos(w) * cos(Omega) * cos(i));
        // Apparent orbit
        pos[1] = A * x + F * y;
        pos[2] = B * x + G * y;
        return pos;
    }

    vector radial_velocity(int N, vector t, real T, real P, real e, real a, real w, real i, real V0, real plx,
                           real q, int primary) {
        // Variables declaration
        real M; real K; vector[N] E; vector[N] nu; vector[N] V;
        real convCoeff = 149597870.660 / (365.25 * 86400.0); // [AU] to [arcsec]
        // Iterate over epochs
        for (j in 1:N) {
            // Mean anomaly
            M = 2 * pi() * (t[j] - T) / P;
            // Eccentric anomaly
            E[j] = kepler(M, e);
        }
        // True anomaly
        nu = 2 * atan(sqrt((1 + e) / (1 - e)) * tan(E / 2));
        // Primary object radial velocity
        if (primary == 1) {
            // Amplitude
            K = 2 * pi() * (a / plx) * (q / (1+q)) * convCoeff * sin(i) / (P * sqrt(1 - e^2));
```

```
            // Radial velocity
            V = V0 + K * (cos(w + nu) + e * cos(w));
        }
        // Companion object radial velocity
        else {
            // Amplitude
            K = 2 * pi() * (a / plx) * (1 / (1+q)) * convCoeff * sin(i) / (P * sqrt(1 - e^2));
            // Radial velocity
            V = V0 - K * (cos(w + nu) + e * cos(w));
        }
        return V;
    }
}
data {
    int<lower=0>    N_A;          // number of samples
    int<lower=0>    Nv_Aa;        // number of samples
    int<lower=0>    Nv_Ab;        // number of samples
    int<lower=0>    N_B;          // number of samples
    int<lower=0>    Nv_B;         // number of samples

    vector[N_A]     t_A;          // epochs (years)
    vector[Nv_Aa]   t_rv_Aa;      // epochs (years)
    vector[Nv_Ab]   t_rv_Ab;      // epochs (years)
    vector[N_B]     t_B;          // epochs (years)
    vector[Nv_B]    t_rv_B;       // epochs (years)

    vector[N_A]     x_obs_A;      // x coordinate observation
    vector[N_A]     y_obs_A;      // y coordinate observation
    vector[N_A]     pos_err_A;    // observation uncertainties
    vector[N_B]     x_obs_B;      // x coordinate observation
    vector[N_B]     y_obs_B;      // y coordinate observation
    vector[N_B]     pos_err_B;    // observation uncertainties

    vector[Nv_Aa]   rv_obs_Aa;   // rv observations
    vector[Nv_Aa]   rv_err_Aa;   // rv observations uncertainties
    vector[Nv_Ab]   rv_obs_Ab;   // rv observation
    vector[Nv_Ab]   rv_err_Ab;   // rv observation uncertainties
    vector[Nv_B]    rv_obs_B;    // rv observations
    vector[Nv_B]    rv_err_B;    // rv observation uncertainties

    int<lower=-1, upper=1> sgn;  // system configuration (sgn=1: AaAb & AaB | sgn=-1: BaBb & ABa)
}
transformed data {
    real          min_t     = fmin(min(t_A), fmin(min(t_B), fmin(min(t_rv_B), fmin(min(t_rv_Aa), min(t_rv_Ab)))));
    vector[N_A]   t_A0      = t_A - min_t;                                    // normalized iner position epochs
    vector[Nv_Aa] t_rv_Aa0  = t_rv_Aa - min_t;                                // normalized inner rv1 epochs
    vector[Nv_Ab] t_rv_Ab0  = t_rv_Ab - min_t;                                // normalized inner rv2 epochs
    vector[N_B]   t_B0      = t_B - min_t;                                    // normalized outer position epochs
    vector[Nv_B]  t_rv_B0   = t_rv_B - min_t;                                 // normalized outer rv2 epochs
}
parameters {
    real<lower=0, upper=1>          T0_A;       // normalized time of periastron passage (T - t0) / P
    real                            log_P_A;    // ln(period)
    real<lower=0, upper=1>          e_A;        // eccentricity
    real<lower=0>                   a_A;        // semimajor axis (secods of arc)
    real<lower=0, upper=2*pi()>     w_A;        // argument of periapsis (rad)
    real<lower=0, upper=2*pi()>     Omega_A;    // longitude of the ascending node (rad)
    real<lower=0, upper=pi()>       i_A;        // inclination (rad)
    real<lower=0, upper=1>          q_A;        // mass ratio
    real<lower=0, upper=1>          L_A;        // light ratio

    real<lower=0, upper=1>          T0_B;       // normalized time of periastron passage (T - t0) / P
    real                            log_P_B;    // ln(period)
    real<lower=0, upper=1>          e_B;        // eccentricity
    real<lower=0, upper=2*pi()>     w_B;        // argument of periapsis (rad)
    real<lower=0, upper=2*pi()>     Omega_B;    // longitude of the ascending node (rad)
    real<lower=0, upper=pi()>       i_B;        // inclination (rad)
    real<lower=0, upper=1>          q_B;        // mass ratio

    real<lower=0>                   plx;        // parallax (seconds of arc)
    real                            V0;         // velocity of center of mass (km/s)
}
transformed parameters {
    real a_B;                    // Outer semimajor axis (secods of arc)
    real P_A = exp(log_P_A);     // period
    real P_B = exp(log_P_B);     // period
    if (sgn == 1) {
        a_B = ((a_A^3 / P_A^2) * P_B^2 * (1 + q_B)       )^(1.0/3.0);
    }
    else {
        a_B = ((a_A^3 / P_A^2) * P_B^2 * (1 + q_B) / q_B)^(1.0/3.0);
    }
}
model {
    // Variables declaration
    real T_A; real T_B;
    vector[N_A] pos_A[2]; vector[N_B] pos_BA[2]; vector[N_B] pos_B[2];
    vector[Nv_Aa] V_Aa; vector[Nv_Aa] V_AaB;
    vector[Nv_Ab] V_Ab; vector[Nv_Ab] V_AbB;
    vector[Nv_B] V_B;
    // Period
    T_A = T0_A * P_A;
    T_B = T0_B * P_B;
    // Orbit inner
    if (N_A > 0) {
        pos_A = orbit(N_A, t_A0, T_A, P_A, e_A, a_A, w_A, Omega_A, i_A);
```

```
        x_obs_A ~ normal(pos_A[1], pos_err_A);
        y_obs_A ~ normal(pos_A[2], pos_err_A);
    }
    // Orbit outer
    if (N_B > 0) {
        pos_BA = orbit(N_B, t_B0, T_A, P_A, e_A, a_A, w_A, Omega_A, i_A);
        pos_B = orbit(N_B, t_B0, T_B, P_B, e_B, a_B, w_B, Omega_B, i_B);
        pos_B[1] = pos_B[1] + sgn * ((q_A - L_A) / ((1 + q_A) * (1 + L_A))) * pos_BA[1];
        pos_B[2] = pos_B[2] + sgn * ((q_A - L_A) / ((1 + q_A) * (1 + L_A))) * pos_BA[2];
        x_obs_B ~ normal(pos_B[1], pos_err_B);
        y_obs_B ~ normal(pos_B[2], pos_err_B);
    }
    // Radial velocity outer
    if (Nv_B > 0) {
        V_B = radial_velocity(Nv_B, t_rv_B0, T_B, P_B, e_B, a_B, w_B, i_B, V0, plx, q_B, sgn * -1);
        rv_obs_B ~ normal(V_B, rv_err_B);
    }
    // Radial velocities inner
    if (Nv_Aa > 0) {
        V_AaB = radial_velocity(Nv_Aa, t_rv_Aa0, T_B, P_B, e_B, a_B, w_B, i_B, V0, plx, q_B, sgn);
        V_Aa = radial_velocity(Nv_Aa, t_rv_Aa0, T_A, P_A, e_A, a_A, w_A, i_A, 0, plx, q_A, 1);
        V_Aa += V_AaB;
        rv_obs_Aa ~ normal(V_Aa, rv_err_Aa);
    }
    if (Nv_Ab > 0) {
        V_AbB = radial_velocity(Nv_Ab, t_rv_Ab0, T_B, P_B, e_B, a_B, w_B, i_B, V0, plx, q_B, sgn);
        V_Ab = radial_velocity(Nv_Ab, t_rv_Ab0, T_A, P_A, e_A, a_A, w_A, i_A, 0, plx, q_A, 0);
        V_Ab += V_AbB;
        rv_obs_Ab ~ normal(V_Ab, rv_err_Ab);
    }
}
generated quantities {
    real w_A_deg     = w_A * 180 / pi();                                  // argument of periapsis (degrees)
    real Omega_A_deg = Omega_A * 180 / pi();                              // longitude of the ascending node (degrees)
    real i_A_deg     = i_A * 180 / pi();                                  // inclination (degrees)
    real T_A         = T0_A * P_A + min_t;                                // unnormalized time of periastron passage
    real m1_A        = (a_A / plx)^3 * 1 / (P_A^2 * (1 + q_A));           // primary object mass (solar masses)
    real m2_A        = q_A * m1_A;                                        // companion object mass (solar masses)
    real plx_mas     = plx * 1000;                                        // parallax (miliseconds of arc)
    real w_B_deg     = w_B * 180 / pi();                                  // argument of periapsis (degrees)
    real Omega_B_deg = Omega_B * 180 / pi();                              // longitude of the ascending node (degrees)
    real i_B_deg     = i_B * 180 / pi();                                  // inclination (degrees)
    real T_B         = T0_B * P_B + min_t;                                // unnormalized time of periastron passage
    real m1_B        = (a_B / plx)^3 * 1 / (P_B^2 * (1 + q_B));           // primary object mass (solar masses)
    real m2_B        = q_B * m1_B;                                        // companion object mass (solar masses)
    real phi = cos(i_A) * cos(i_B) + sin(i_A) * sin(i_B) * cos(Omega_B - Omega_A);  // Mutual incliniation
    real phi_deg = phi * 180 / pi();                                      // Mutual inclination (degrees)
}
```

Listing A.3: Triple hierarchical system probabilistic model in Stan.

```
inline double kepler (double M, double e, std::ostream* pstream) {
  double E0 = M;
  double E = M;
  double g, gp;
  for (int i = 0; i < 200; ++i) {
    g = E0 - e * sin(E0) - M;
    gp = 1.0 - e * cos(E0);
    E = E0 - g / gp;
    // Check for convergence.
    if (fabs((E - E0) / E) <= 1.234e-10) {
      return E;
    }
    E0 = E;
  }
  // If we get here, we didn't converge, but return the best estimate.
  return E;
}

inline var kepler (const var& M_var, const var& e_var, std::ostream* pstream) {
  // First, compute the value of E at the current values of M and e
  double M = M_var.val(),
         e = e_var.val(),
         E = kepler(M, e, pstream);

  // Then compute the partial derivatives:
  double dE_dM = 1.0 / (1.0 - e * cos(E)),
         dE_de = sin(E) * dE_dM;

  // Construct the autodiff wrapper:
  return var(new precomp_vv_vari(
    E,              // The _value_ of the output
    M_var.vi_,      // The input gradient wrt M
    e_var.vi_,      // The input gradient wrt e
    dE_dM,          // The partial introduced by this function wrt M
    dE_de           // The partial introduced by this function wrt e
  ));
}

inline var kepler (double M, const var& e_var, std::ostream* pstream) {
  double e = e_var.val(),
         E = kepler(M, e, pstream),
         dE_de = sin(E) / (1.0 - e * cos(E));
  return var(new precomp_v_vari(E, e_var.vi_, dE_de));
```

```
}

inline var kepler (const var& M_var, double e, std::ostream* pstream) {
  double M = M_var.val(),
         E = kepler(M, e, pstream),
         dE_dM = 1.0 / (1.0 - e * cos(E));
  return var(new precomp_v_vari(E, M_var.vi_, dE_dM));
}
```

Listing A.4: Kepler equation implementation.