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

BAYES-BASED ORBITAL PARAMETERS ESTIMATION IN TRIPLE HIERARCHICAL STELLAR SYSTEMS

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

CONSTANZA LUISA VILLEGAS MARDONES

PROFESOR GUÍA:
JORGE SILVA SÁNCHEZ

PROFESOR CO-GUÍA:
RENÉ MÉNDEZ BUSSARD

MIEMBROS DE LA COMISIÓN:
MARCOS ORCHARD CONCHA
GAËL CHAUVIN

This work has been partially funded by FONDECYT 1170854 and AC3E Basal Project

SANTIAGO DE CHILE
2020

RESUMEN DE LA TESIS PARA OPTAR
AL GRADO DE MAGÍSTER EN CIENCIAS DE LA INGENIERÍA, MENCIÓN ELÉCTRICA
POR: CONSTANZA LUISA VILLEGAS MARDONES
FECHA: JULIO 2020
PROF. GUÍA: SR. JORGE SILVA SÁNCHEZ

BAYES-BASED ORBITAL PARAMETERS ESTIMATION IN TRIPLE HIERARCHICAL STELLAR SYSTEMS

A hierarchical triple stellar system is viewed as two binary star systems combined, for estimating its orbital parameters. This inference is very challenging technically because of the large dimensions of the parameter space and the complex relationships between parameters and the observations involved (astrometry and radial velocity).

This work proposes a new methodology for this estimation using a Bayesian MCMC-based framework. In particular, graphical models are proposed for modelling the probabilistic relationship between parameters and observations in the context of isolated astrometry, isolated radial velocity, and the joint case with astrometry and radial velocity as information sources. They provide a novel way of performing the factorization of the joint distribution (of parameter and observations) in terms of conditional independent components (factors), so the estimation can be performed in a two-stage process that combines different observations sequentially.

A mathematical formalism to reduce the dimensionality in the state space for triple hierarchical stellar systems in general is also provided.

Finally, this framework is tested on three well-studied benchmark cases of triple systems, where the inner and outer orbital elements are determined, coupled with the mutual inclination of the orbits and the individual stellar masses. The results are consistent with previously reported ones.

RESUMEN DE LA TESIS PARA OPTAR
AL GRADO DE MAGÍSTER EN CIENCIAS DE LA INGENIERÍA, MENCIÓN ELÉCTRICA
POR: CONSTANZA LUISA VILLEGAS MARDONES
FECHA: JULIO 2020
PROF. GUÍA: SR. JORGE SILVA SÁNCHEZ

BAYES-BASED ORBITAL PARAMETERS ESTIMATION IN TRIPLE HIERARCHICAL STELLAR SYSTEMS

En la estimación de parámetros orbitales, un sistema estelar triple jerárquico puede ser visto como dos sistemas binarios combinados. Este proceso de inferencia es muy desafiante, dada la alta dimensionalidad del espacio de parámetros y las complejas relaciones entre éstos y las observaciones involucradas (astrometría y velocidades radiales).

Este trabajo propone una nueva metodología para realizar esta estimación, usando un framework bayesiano basado en MCMC. En particular, se proponen modelos gráficos para modelar las relaciones probabilísticas entre parámetros y observaciones en el contexto de sólo astrometría, sólo velocidades radiales, y el caso conjunto de astrometría y velocidades radiales como fuentes de información. Estos proveen una manera novedosa de realizar la factorización de la distribución conjunta (de parámetros y observaciones) en componentes condicionales independientes (factores), y la estimación se realiza combinando diferentes tipos de observaciones de manera secuencial.

También se plantea un formalismo matemático para reducir la dimensionalidad en el espacio de estados para sistemas estelares triples jerárquicos en general.

Finalmente, este framework es testeado en tres casos conocidos de benchmark, donde los elementos orbitales internos y externos están determinados, junto con la inclinación mutua de las órbitas y las masas estelares individuales. Los resultados son consistentes con los previamente reportados.

Acknowledgements

I would like to thank all the people who helped me during my Master's thesis project. Firstly, my friends, who supported me with love, understanding, and memes. I would also like to give a special emphasis to Pablo, Gabriel, and little Kida, who held me up in the worst times.

Secondly, I would like to acknowledge my committee members, each of whom has provided advice and guidance throughout the whole research process. Thank you all for your support.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

In astronomy, the knowledge of stellar masses is really important to develop models related to formation, structure and evolution of stars. In the particular case of multiple systems, there is great interest in also determining the relative orbit orientation, because it provides information about the formation and evolution of the stars and planets involved in the system [Muterspaugh et al., 2010, Tokovinin and Latham, 2017].

The only direct method to calculate them is through the analysis of the motion of stars that are bounded gravitationally and the computation of the parameters related to that movement [Czekala et al., 2017, Mendez et al., 2017, Pourbaix, 1994]. In the case of triple systems in particular, both visual and radial velocity data are required [Muterspaugh et al., 2010, Tokovinin and Latham, 2017]; and visual-only orbits coupled with parallax measurements, can be used to measure the total mass of the system.

As a Bayesian approach has not been explored for considering the combination of radial velocity data and astrometry, this Thesis addresses the task of estimating the orbital parameters in triple stellar systems by obtaining the conditional distribution over the full parameter space. Generative models are proposed using graphical model tools that exploit the hierarchical approximation, and the distributions are computed adopting simulation-based schemes. After computing them, the most likelihood solution of orbital parameters is obtained, as well as confidence measures.

1.1 Hypotheses

This work focuses on the study of triple hierarchical stellar systems using astrometric and spectroscopic measurements, and it aims at testing the following hypotheses:

1. Graphical model tools are a means to characterize the posterior probability density function of orbital parameters in three different observational settings:

   (a) Data sets with astrometric measurements of relative position between a primary and a companion star.

(b) Data sets containing radial velocity measurements.

(c) Data sets containing both kind of measurements.

2. The sample-based approximation of the probability density functions provided by Markov Chain Monte Carlo can be used to calculate estimates such as expected value and *maximum a posteriori*, and those estimates lead to physically feasible orbits of the stellar systems under study.

3. The availability of both astrometric and spectroscopic observations provide a estimation of the individual stellar masses and the mutual inclination of the systems included in the triple hierarchical approximation.

1.2 Objectives

1.2.1 Main objective

The main objective of this work is the implementation of a framework that analyses triple hierarchical stellar systems using Bayesian processes. This framework gives as an output an estimation of the orbital parameters in terms of sample-based approximations of the probability density distribution. This approximation allows to compute the most likely orbital parameters concerning the input data, and provides significant quantities as the stellar masses involved. In addition, it has the potential to describe the full posterior distributions that can be used to have indicators of the discrimination capacity that measurements have to estimate orbital parameters.

1.2.2 Specific objectives

The specific objectives of this work are the following:

1. Formulate analytic expressions of the kinematic equations of hierarchical triple systems, which describe the position and radial velocity of the inner and outer systems.

2. Design a Bayesian framework using graphical model tools. The following scenarios must be considered:

   (a) Using astrometric measurements alone,

   (b) Using spectroscopic measurements alone,

   (c) Using both astrometric and spectroscopic measurements.

3. Implementation of the algorithms associated to the aforementioned scenarios. This task considers the design and implementation of several Markov Chain Monte Carlo routines to sample the posterior distribution formulated in the graphical model mentioned before.

   (a) Using astrometric measurements alone,

(b) Using spectroscopic measurements alone,

(c) Using both astrometric and spectroscopic measurements.

4. Analyse Tokovinin and Latham [2017] and Tokovinin [2018a] data sets in order to test the implemented algorithms. This aims to check the hypotheses mentioned in Section 1.1.

5. Brief analysis.

It is important to mention that this work is performed as a student member of the Information and Decision Systems Laboratory[1] and it is carried out in collaboration with Professors René Méndez[2], Jorge F. Silva[3] and Marcos Orchard[4]. Also, it is funded by Fondecyt Project 1170854 - *Topics on Information and Decision with Applications to Coding and Inverse Problems in Astronomy.*

1.3 Structure

This document consists of six chapters, which are described as follows:

- Chapter 2 presents the State-of-the-Art regarding orbital parameters estimation in hierarchical systems and using Bayesian tools.

- Chapter 3 shows the theoretical subjects used to accomplish the estimation modelling: graphical models in Section 3.1, Markov Chain Monte Carlo in Section 3.2 and the Gibbs Sampler in Section 3.3.

- Chapter 4 describes the observational model for triple hierarchical systems, along with the probabilistic modelling and orbit calculation.

- Chapter 5 compares the obtained model to three well-studied systems selected from Tokovinin [2018a], Tokovinin and Latham [2017], in different scenarios that serve as benchmarks.

- Chapter 6 shows a brief discussion about the algorithm performance

- Finally, Chapter 7 presents the summary, conclusions and outlook; and Chapter 8 displays the future work.

The Appendices give full details about some relevant aspects of the methodology which are described only succinctly in the main body of this Thesis.

[1]Electrical Engineering Department, Universidad de Chile
[2]Associated Professor, Astronomy Department, Universidad de Chile
[3]Associated Professor, Electrical Engineering Department, Universidad de Chile
[4]Associated Professor, Electrical Engineering Department, Universidad de Chile

Chapter 2

State of the Art

This chapter presents a brief review of the state-of-the-art regarding fitting stellar orbits in hierarchical triple systems and hierarchical systems in general, especially focusing on works that consider a Bayesian approach. Due to the hierarchical approximation, Bayesian estimation on binary systems is also covered. To get more detail about the hierarchical triple systems structure, see Appendix A.

The only direct method to calculate the stellar masses is through the analysis of the motion of stars that are bounded gravitationally and the computation of the parameters related to that movement [Czekala et al., 2017, Mendez et al., 2017, Pourbaix, 1994]. For this reason, the problem of estimating orbital parameters in binaries has been widely studied in the literature, in particular with algorithms that involve a Bayesian methodology [eg: Blunt et al., 2017, Ford, 2005, Lucy, 2014, 2018, Mendez et al., 2017, Sahlmann et al., 2013]. Bayesian procedures have a probabilistic nature, and their final objective is a precise approximation of the conditional distribution of orbital parameter given the observations. Besides, it is important to have an indicator expressing the confidence about the estimated observational parameters after analyzing the star motion data, and the aforementioned distribution captures the uncertainties and ll the information inferred from the data. On the other hand, Bayesian orbit fitting could be useful to determine the optimal placement of future observations, and thus reduce the uncertainty in the computed distributions [Blunt et al., 2017].

Moving to multiple systems, there is great interest in also determining the relative orbit orientation, because it provides information about the formation and evolution of the stars and planets involved in the system [Muterspaugh et al., 2010, Tokovinin and Latham, 2017]. In the case of triple systems in particular, in order to derive stellar masses, luminosisites, and radii, along with determining the system's coplanarity, both visual and RV data are required [Muterspaugh et al., 2010, Tokovinin and Latham, 2017]. Nevertheless, visual-only orbits coupled with parallax measurements, can be used to measure the total mass of the system.

The hierarchical approximation[1] is in many cases useful, because it describes the whole system as two binary systems interacting between themselves, so there is no need to resort to numerical methods, because an analytic expression can be obtained. There are plenty

[1]For a definition of a hierarchical stellar system see Appendix A.

of works that exploit this approximation and handle the estimation by disconnecting the inner and the outer orbits. They treat each system as a binary case, where they perform the estimation of parameters by the optimization a function of merit or a geometric based procedure [eg: Docobo et al., 2008, Köhler et al., 2012, Tokovinin, 2018a].

In light of this, there are some works that combine the visual information with spectroscopic measurements.[Muterspaugh et al., 2010] combine both data set and minimize the χ^2 statistsic. [Czekala et al., 2017] determine the parameters using cross-correlation peaks [Torres et al., 2002] and MCMC, combining the RV data with archival astrometry, and assessing convergence using the Gelman-Rubin statistic [Gelman et al., 1992]. [Tokovinin and Latham, 2017] propose to perform the estimation in consecutive sequential stages, alternating between visual and spectroscopic data from the inner and outer systems, while minimizing the χ^2 statistic.

2.1 Contribution

However, it can be noted that a Bayesian approach has not been explored for considering the combination of RV data and astrometry. Even though those methods have been widely used in exoplanets research [e.g.: Ford and Gregory, 2006, Gregory, 2005, 2009, 2010, 2011, Retired, 2010], they only considers RV measurements, and information related to several orbital elements is present in both data sets [Lucy, 2018]. In this *combined* scenario the problem of estimating the orbital parameters becomes more complex, since the dimensions of the parameter space and the number of observation sources increases[2], in addition to the intricate relationships between them.

To cope with those challenges, the task of estimating the orbital parameters in triple hierarchical stellar systems is addressed, by obtaining the conditional distribution over the full parameter space, where visual and spectroscopic information is taken into account. Generative models that capture the distribution (probabilistic relationship) between parameters and observations are proposed. To accomplish that, graphical model tools are employed to model that probabilistic relationship considering the underlying dynamical model of triple hierarchical stellar systems. These graphical models provide a novel way of performing the factorization of the joint distribution (of parameter and observations) in terms of conditional independent components (factors). Taking into account these factorizations of the joint distribution, certain probabilistic relationships between parameters and observations get disconnected, and, as a result, the estimation of the posterior distribution can be performed in a multiple-stage process [Jordan, 1998]. This process combines different sources of observations sequentially to update the posterior distribution of parameters given the observations.

To compute the mentioned distributions, the well-known simulation-based scheme MCMC [Robert and Casella, 2004] is adopted. It has been widely exploited to provide an empirical but precise approximation of the posterior distribution when the exact expression is intractable [e.g.: Gamerman and Lopes, 2006, Liu, 2008]. In the present work a new MCMC-based code is developed to compute the orbital configuration of a triple hierarchical stellar

[2]The astrometry can be obtained for the inner and outer systems, and the radial velocity measurements from each one of the bodies involved.

system, partially constructed upon the binary case described by [Mendez et al., 2017]. After computing them, the most likelihood solution of orbital parameters (the MAP solution) can be obtained, as well as confidence measured in terms of the variance of the posterior distribution.

Chapter 3

Theoretical Framework

This chapter presents the main concepts that are necessary to understand the probabilistic modelling in this thesis.

3.1 Graphical Models

Before the interactions between the observations and the parameters are analysed, some basic concepts from the theory of Graphical Models that help to encode the conditional independence properties observed within the involved variables will be introduced. Graphical models are a powerful tool that comes from statistics, graph theory, and computer science. They seek to represent statistical relationships and dependencies between variables through a graphs representation where every node in the graph is a (scalar) random variable and the arcs captures concrete independence properties of the joint distribution of the problem. Therefore a graph is induced by a joint distribution and the structure of the graph encodes key dependencies within the variables. To illustrate, in the context of the Bayes Theorem (Equation 3.1), the prior-to-posterior inference process can be encoded as the diagram depicted in Figure 3.1, where a joint distribution (center) can be factorized in conditional components (left and right). Graphs can be directed or undirected, depending on the system being graphically represented, and the direction of the arrows represent influence.

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \tag{3.1}$$

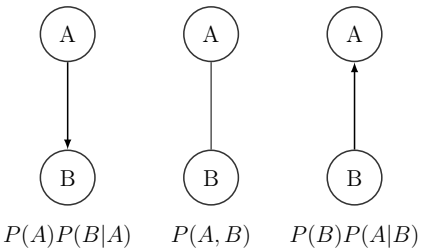

Figure 3.1: Graphical Model representation of the Bayes Theorem. Left and right graphs are directed graphs, while the center graph is an undirected one.

Bayesian networks are a specific configuration of directed and acyclic graphical models, which portray the joint distribution of a set of variables in terms of conditional and prior probabilities. They allow to simplify the whole distribution in terms of factorization, due to the following basic principles [Bishop, 2006, Jordan, 1998]:

1. The graph $Z \leftarrow X \rightarrow Y$ means that $Z \perp\!\!\!\perp Y|X$, i.e. Z is conditionally independent from Y given X.

2. Given any node X_i, let's denote by $\text{pa}(X_i)$ its parents variables from the directed graph. Then its basic conditional probability (or predictive model) is: $P(X_i|\text{pa}(X_i))$. If a particular node X_i has no parents ($\text{pa}(X_i) = \{\}$), a not conditional probability will be taken: $P(X_i|\text{pa}(X_i)) = P(X_i)$.

3. Given a graph, the joint density over the set of variables $U = \{X_i, i = 1, .., L\}$ follows a **recursive factorization**:

$$P(U) = \prod_{i=1}^{L} P(X_i|\text{pa}(X_i)) \tag{3.2}$$

   Assuming that a variable X is independent of its non-descendants given its parents : $X \perp\!\!\!\perp \text{nd}(X)|\text{pa}(X)$, where $\text{nd}(X)$ is a short-hand for all the variables not contained in $\text{pa}(X)$ and non-including X. This is called the **directed Markov property**.

Another example can be seen in Figure 3.2, where a Bayesian network with three variables is shown. In this case, the joint probability is given by $P(A, B, C) = P(C)P(A|C)P(B|C)$.

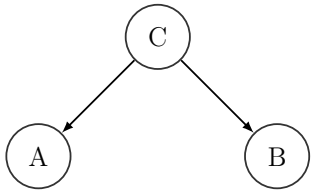

Figure 3.2: Three variable Bayesian network

Therefore, graphical models (and Bayesian Networks in particular) can answer *conditional probability query* $P(\Theta|Z = z)$, where Z is the evidence and Θ corresponds to some random

Algorithm 1 Metropolis-Hastings algorithm

Initialize $x^{(0)}$ sampling from prior distribution
for $i = 1, .., N_{\text{steps}}$ **do**
  $x' = x$ with $x \sim q(x|x^{(i-1)})$
  $u' = u$ with $u \sim \mathcal{U}(0, 1)$
  $\mathcal{A} = \min\left(1, \frac{\pi(x')\cdot q(x|x')}{\pi(x)\cdot q(x'|x)}\right)$
  if $u' < \mathcal{A}$ **then**
    Accept sample: $x^{(i)} = x'$
  else
    $x^{(i)} = x^{(i-1)}$
  end if
end for

variables in the network. As $P(\Theta|Z = z) = \frac{P(\Theta,z)}{P(z)}$, $P(\Theta, z)$ can be obtained through the factorization in conditional independent components [Koller et al., 2007]. Then, in this context, this tool provides a way to conduct the estimation in several processes that combine different observations in a sequential fashion.

3.2 Markov Chain Monte Carlo

Markov Chain Monte Carlo, also known as MCMC, is a Bayesian inference technique that is based on Markov chains properties. The iterations of the algorithm are modelled as states from an irreducible aperiodic Markov chain $\{X_n\}_{n\geq 0}$ with state space E, and the equilibrium distribution of the chain $\pi(\cdot)$ corresponds to the target distribution to be sampled from. The design of the chain grants spending most of the time in areas of high probability Andrieu et al. [2003].

There is no need to actually know the target distribution because the samples are drawn from a **proposal distribution** $q(\cdot|\cdot)$ and they are accepted or rejected considering an **acceptance probability ratio** $\mathbf{A}(x, x')$. In the context of Bayesian inference problems, the target distribution is the posterior distribution of the parameters Θ to be estimated from the data Z, $\pi = p(\Theta|Z)$. Also, the proposal distribution is based on the posteriors approximation $p(\Theta|Z)\ \alpha\ p(Z|\Theta)\cdot p(Z)$, so it is portrayed by $q(\Theta|\Theta') = p(Z|\Theta')\cdot p(Z)$.

Finally, convergence is guaranteed because for these chains, for any initial distribution μ and all $i \in E$, $\lim_{n\to\infty} P_\mu(X_n = i) = \pi(i)$, which is analogous for the continuous case Brémaud [2013].

There is a wide variety of MCMC variants, and one of them is the Metropolis-Hastings method (Algorithm 1), where:

1. The **acceptance probability ratio** consists of $\mathcal{A}(x, x') = \min\left(1, \frac{\pi(x')\cdot q(x|x')}{\pi(x)\cdot q(x'|x)}\right)$.

2. The **target density** update consists of $\pi(x_{i+1}|x_i) = q(x_{i+1}|x_i)\cdot \min\left(1, \frac{\pi(x')\cdot q(x|x')}{\pi(x)\cdot q(x'|x)}\right)$.

Algorithm 2 Gibbs sampler algorithm

Initialize $x^{(0)}$
for $i = 1, .., N_{\text{steps}}$ **do**
  Retrieve states from last iteration: $x^{(i)} = x^{(i-1)}$
  for $j = 0, .., d$ **do**
    $x_j^i \sim \pi(x_j | x_j^i, .., x_{j-1}^i, x_{j+1}^{i-1}, .., x_d^{i-1})$
  end for
end for

3.3 Gibbs Sampler

Gibbs sampling (Algorithm 2) is a special case of the Metropolis-Hastings algorithm, useful for sampling multidimensional distributions. Given a multidimensional parameter θ, this sampler splits it into blocks and samples each block separately, generating the posterior distribution by sampling through the conditional distribution of each block, conditional on the fixed current values of the other blocks.

This scheme is equivalent to draw samples from the joint target distribution itself, but it breaks down a complex high-dimensional problem into simple low-dimensional problems. Also, sequential sampling of parameters prevents the algorithm from falling in zones of near-zero probability.

The proposal distribution of Gibbs sampler is defined for each component $j \in \{1, \dots, d\}$:

$$q(x'|x^{(i)}) = \begin{cases} \pi(x'_j | x_1^{(i)}, \dots, x_{j-1}^{(i)}, x_{j+1}^{(i-1)}, \dots, x_d^{(i-1)}), & \text{if } x_k^{(i)} = x'_k, \forall k \neq j \\ 0, & \text{otherwise} \end{cases}$$

And the **acceptance probability ratio** $\mathcal{A}$ is always 1, so the proposal is always accepted.

3.3.1 Metropolis-within-Gibbs

This sort of approach is known as variable-at-a-time Metropolis-Hastings or Metropolis-within-Gibbs, which consists of including Metropolis-Hastings steps within the Gibbs sampler (Algorithm 3). Here, an auxiliary proposal distribution q'_j is defined, that proceeds only for the j-th component. This leads to a much easier way of proposing updates by sequentially sampling the parameters; nevertheless, when variables are highly correlated, it may be very difficult to change one without simultaneously changing the other.

3.4 Imputation Theory

Imputation theory addresses the problem of performing an estimation when incomplete data is considered, which contemplates missing data and/or partial measurements, by augmenting the observed data. While **single imputation techniques** rely on generating one plausible data set that completes the blank spaces, they omit any sources of uncertainty that can be associated to the imputation process, topic taken into account by the **multiple imputation** approach. These techniques generate multiple plausible data sets, which are analysed and then combined.

Algorithm 3 Metropolis-within-Gibbs algorithm

Initialize x^0
for $i = 1, .., N_{\text{steps}}$ **do**
  Retrieve states from last iteration: $x^{(i)} = x^{(i-1)}$
  for $j = 0, .., d$ **do**
    $x' = x$ with $x \sim q'_j(x|x^{(i)})$
    $u' = u$ with $u \sim \mathcal{U}(0, 1)$
    $\mathcal{A} = \min\left(1, \frac{\pi(x') \cdot q'_j(x^{(i)}|x')}{\pi(x^{(i)}) \cdot q'_j(x'|x^{(i)})}\right)$
    if $u' < \mathcal{A}$ **then**
      Accept sample: $x^{(i)} = x'$
    else
      $x^{(i)} = x^{(i-1)}$
    end if
  end for
end for

Let $\mathcal{Y}$ be a data set, which consists of observed ($\mathcal{Y}_{obs}$) and missing values ($\mathcal{Y}_{mis}$), and θ any parameter vector, the main goal of multiple imputation is the estimation of $p(\theta|\mathcal{Y}_{obs})$, considering the distribution $p(\mathcal{Y}_{mis}|\mathcal{Y}_{obs})$. Tanner and Wong [1987] propose an data augmentation algorithm that converges to $p(\theta|\mathcal{Y}_{obs})$ and can be included in any iterative estimation scheme:

1. Generate $m > 0$ samples $Y_{mis}^{(i,1)}, \ldots, Y_{mis}^{(i,m)}$ from $p_i(\mathcal{Y}_{mis}|\mathcal{Y}_{obs})$, considering the current guess $p_i(\theta|\mathcal{Y}_{obs})$. This can be achieved by drawing a sample $\theta_i \sim p_i(\theta|\mathcal{Y}_{obs})$ and then sampling the imputations from $p(Y_{mis}|\theta_i, \mathcal{Y}_{obs})$ Claveria et al. [2019].

2. Update $p_i(\theta|\mathcal{Y}_{obs})$ considering a mixture of the m augmented posteriors.

Chapter 4

Orbital Parameters Estimation

4.1 Observation Model

Triple hierarchical stellar systems consist of two binary systems bounded together, as the ones illustrated in Figure 4.1, where two configurations are feasible. The first one is composed of two bodies in an *inner* orbit (A_a and A_b) and an external body (B) which, along with the center of mass of the inner orbit (A), constitutes the *outer* orbit. The alternative configuration consists of a body (A) which, along with an inner binary (B_a and B_b) form the outer system. Then, both systems behave like Keplerian orbits and they are characterized by parameters shown in Figure 4.2, where the notation for each system and the symbols meanings are indicated. The *inner parameters* will be denoted as the parameters related to the inner orbit and the *outer parameters*, as the ones related to the outer orbit.

For each system, certain sets of measurements are available. In all cases, the observation model from Equation (4.1) will be assumed, which maps the n_θ-dimensional parameter vector into the n_z-dimensional measurement vector. Besides, $f(\theta, \tau)$ corresponds to a n_z-dimensional function, and $\varepsilon(\tau)$ to the observation noise, assumed to be additive white Gaussian noise.

$$z(\tau) = f(\theta, \tau) + \varepsilon(\tau) \tag{4.1}$$

4.2 Inference processes

The final objective in Bayesian inference is the computation of the predictive model of the parameters $P_{\Theta|Z}(\cdot|z)$. This posterior distribution captures all the information inferred from the data z, allowing to derive estimators of the parameters $\hat{\theta}$ given the evidence and taking into account the uncertainty in the estimate, instead of basing the prediction just on the most likely value. The purpose of this Section is to present some analysis of the inference problem to simplify the computation of the predictive model $P_{\Theta|Z}(\cdot|z)$.

In the context of a hierarchical stellar system, the main practical goal in this process is to obtain the relative orbit orientation along with the individual stellar masses, thus, following

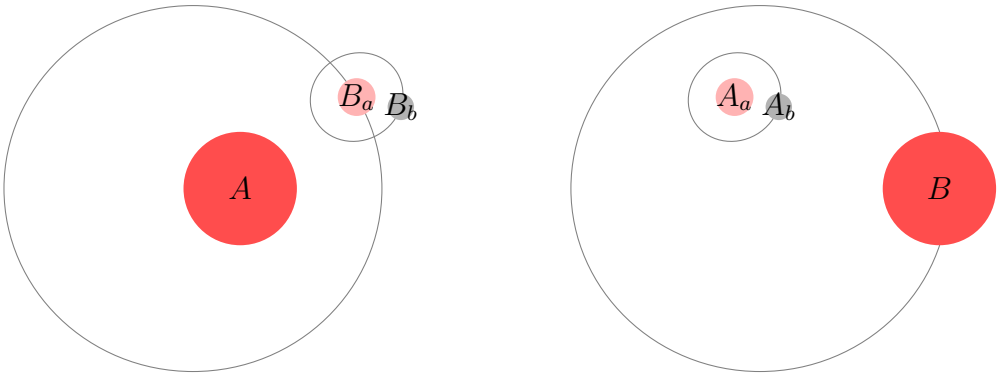

Figure 4.1: Triple hierarchical stellar systems. They consist of an internal binary (A_a and A_b), orbited by an external body (B). Following the classical convention from visual binary research, the A component is the brighter (more massive) star, while B is the the fainter (less massive) component. Of course, it could also be that the tighter binary is B (consisting thus of the external binary B_a and B_b) orbiting a single primary A.

Symbol	Meaning
$P_{A_a A_b}$	Period for inner system
$T_{A_a A_b}$	Time of periastron passage for inner system
$e_{A_a A_b}$	Eccentricity for inner system
$a_{A_a A_b}$	Semi-major axis for inner system
$\omega_{A_a A_b}$	Argument of the periapsis for inner system
$\Omega_{A_a A_b}$	Longitude of the ascending node for inner system
$i_{A_a A_b}$	Orbital inclination for inner system
K_1	Radial velocity amplitude for inner system's primary
K_2	Radial velocity amplitude for inner system's secondary
$q_{A_a A_b}$	Mass ratio for inner system
P_{AB}	Period for outer system
T_{AB}	Time of periastron passage for outer system
e_{AB}	Eccentricity for outer system
a_{AB}	Semi-major axis for outer system
ω_{AB}	Argument of the periapsis for outer system
Ω_{AB}	Longitude of the ascending node for outer system
i_{AB}	Orbital inclination for outer system
K_3	Radial velocity amplitude for outer system's primary
K_4	Radial velocity amplitude for outer system's secondary
q_{AB}	Mass ratio for outer system
v_{cm}	Velocity of system's center of mass
$\nu_{A_a A_b}$	True anomaly for the inner system
ν_{AB}	True anomaly for the outer system

Figure 4.2: Variables and symbols used in this manuscript

the procedure from Appendix A.5, the predictive model of the parameters a, P, q, i, Ω from the inner **and** the outer subsystems is needed. For both of them, astrometry and radial velocities measurements are accessible; however, they are not always available for both orbits simultaneously. Therefore, a methodology that considers all possible combinations of astrometry and radial velocities for every triple hierarchical systems does not seem reasonable. Based on the type of data that are typically available, the following scenarios have been considered: **astrometric observations alone**, **radial velocity observations alone** and **both sources combined**.

Astrometry Alone

An astrometric observation of the inner system corresponds to the relative position of the secondary A_b with respect to the primary A_a, described by a Keplerian orbit, in Cartesian coordinates. It follows the general observation model from Equation (4.1), with a highly non-linear function f_1:

$$\vec{z_1}(\tau) = f_1(T_{A_aA_b}, P_{A_aA_b}, e_{A_aA_b}, a_{A_aA_b}, \omega_{A_aA_b}, \Omega_{A_aA_b}, i_{A_aA_b}, \tau) + \varepsilon_1(\tau) \tag{4.2}$$

An astrometric observation of the outer system corresponds to the relative position of B with respect to the inner primary A_a. However, as A and B behave in a Keplerian way, and the primary A_a is moving along with A_b, it causes a *wobble*, leading the measurements also to depend on inner parameters. The position of B with respect to A_a is described in Cartesian coordinates, and the observations also follow the general observation model from Equation (4.1):

$$\begin{aligned} \vec{z_2}(\tau) &= f_2(T_{A_aA_b}, P_{A_aA_b}, e_{A_aA_b}, a_{A_aA_b}, \omega_{A_aA_b}, \Omega_{A_aA_b}, i_{A_aA_b}, q_{A_aA_b}, \\ &\quad T_{AB}, P_{AB}, e_{AB}, a_{AB}, \omega_{AB}, \Omega_{AB}, i_{AB}, \tau) + \varepsilon_2(\tau) \end{aligned} \tag{4.3}$$

However, given the form of f_2, the observation equation from Equation (4.3) can be written as:

$$\begin{aligned} \vec{z_2}(\tau) &= f_1(T_{AB}, P_{AB}, e_{AB}, a_{AB}, \omega_{AB}, \Omega_{AB}, i_{AB}, \tau) \\ &+ \frac{q_{A_aA_b}}{1+q_{A_aA_b}} \cdot f_1(T_{A_aA_b}, P_{A_aA_b}, e_{A_aA_b}, a_{A_aA_b}, \omega_{A_aA_b}, \Omega_{A_aA_b}, i_{A_aA_b}, \tau) + \varepsilon_2(\tau) \\ &= \vec{y_2}(\tau) + \vec{y_1}(\tau) + \varepsilon_2(\tau) \end{aligned} \tag{4.4}$$

Given the parameters involved in the observation model, this scenario allows to compute the mutual inclination[1] and the sum of the stellar masses[2], but not the individual stellar masses [Lane et al., 2014]. It is worth mentioning that functions f_1 and f_2 are presented in full detail in Appendix A.3. Based on the dependencies observed in the Eq. (4.2), (4.3) and (4.4), graphical models were designed expressing the relationships between parameters and the observations $\vec{z_1}$ (inner orbit) and $\vec{z_2}$ (outer orbit). They are shown in Figures 4.3 and 4.4, where N_1 and N_2 are the number of observations of $\vec{z_1}$ and $\vec{z_2}$, respectively. Those networks

[1]Ambiguously, due to the ambivalence of the Ω angle, the RV measurements are needed

[2]The parallax of the system is needed.

allows to factorize the joint distribution in conditional components and, therefore, arrange the inference process for this scenario.

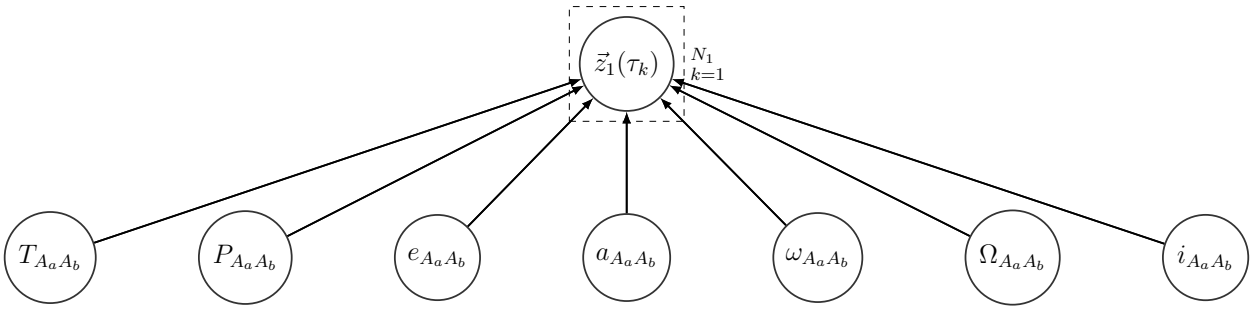

Figure 4.3: Graphical model representation of the observation model from inner astrometric measurements.

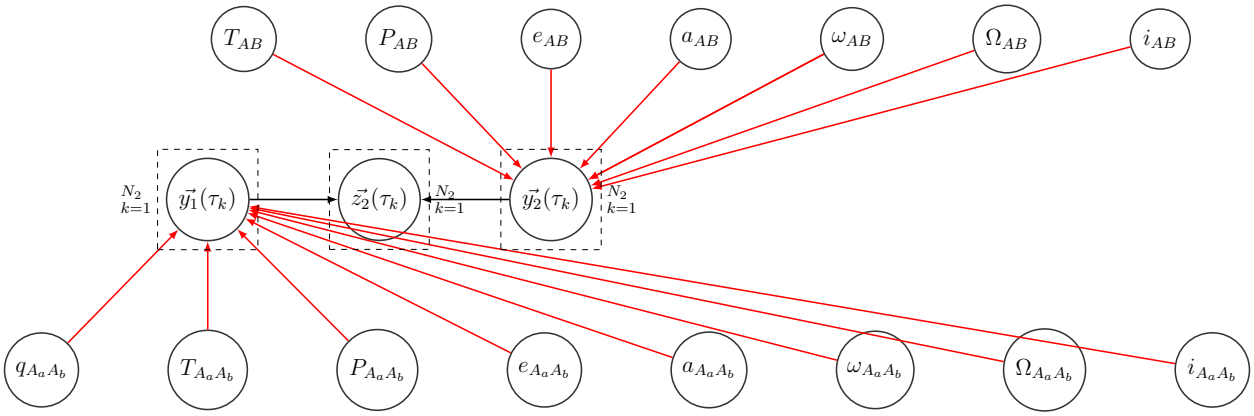

Figure 4.4: Graphical model representation of the observation model from outer astrometric measurements. Red arrows represent deterministic (not probabilistic) relationships

The graphical model of the inner system is really simple, so the factorization remains as Equation (4.5).

$$\begin{aligned} p(U) &= p(\vec{z}_1|T_{A_aA_b}, P_{A_aA_b}, e_{A_aA_b}, a_{A_aA_b}, \omega_{A_aA_b}, \Omega_{A_aA_b}, i_{A_aA_b}) \\ &\cdot\ p(T_{A_aA_b}) \cdot p(P_{A_aA_b}) \cdot p(e_{A_aA_b}) \cdot p(a_{A_aA_b}) \cdot p(\omega_{A_aA_b}) \cdot p(\Omega_{A_aA_b}) \cdot p(i_{A_aA_b}), \\ \text{where}\quad & U = \{\vec{z}_1, T_{A_aA_b}, P_{A_aA_b}, e_{A_aA_b}, a_{A_aA_b}, \omega_{A_aA_b}, \Omega_{A_aA_b}, i_{A_aA_b}\} \end{aligned} \tag{4.5}$$

In contrast, the graphical model of the outer system is more complex, because it involves more parameters and includes some virtual observations in it. Then, the factorization can be shown as:

$$\begin{aligned} p(U) &= p(\vec{z}_1|\vec{y}_1, \vec{y}_2) \cdot p(\vec{y}_1|T_{A_aA_b}, P_{A_aA_b}, e_{A_aA_b}, a_{A_aA_b}, \omega_{A_aA_b}, \Omega_{A_aA_b}, i_{A_aA_b}, q_{A_aA_b}) \\ &\cdot\ p(\vec{y}_2|T_{AB}, P_{AB}, e_{AB}, a_{AB}, \omega_{AB}, \Omega_{AB}, i_{AB}) \\ &\cdot\ p(T_{AB}) \cdot p(P_{AB}) \cdot p(e_{AB}) \cdot p(a_{AB}) \cdot p(\omega_{AB}) \cdot p(\Omega_{AB}) \cdot p(i_{AB}) \\ &\cdot\ p(T_{A_aA_b}) \cdot p(P_{A_aA_b}) \cdot p(e_{A_aA_b}) \cdot p(a_{A_aA_b}) \cdot p(\omega_{A_aA_b}) \cdot p(\Omega_{A_aA_b}) \cdot p(i_{A_aA_b}) \cdot p(q_{A_aA_b}), \\ \text{where}\quad & U = \{\vec{z}_2, \vec{y}_1, \vec{y}_2, T_{A_aA_b}, P_{A_aA_b}, e_{A_aA_b}, a_{A_aA_b}, \omega_{A_aA_b}, \Omega_{A_aA_b}, i_{A_aA_b}, q_{A_aA_b}, \\ & T_{AB}, P_{AB}, e_{AB}, a_{AB}, \omega_{AB}, \Omega_{AB}, i_{AB}\} \end{aligned} \tag{4.6}$$

In these graphical models, the red arrows in Figure (4.4) indicate that the conditional distribution is Dirac function, representing a deterministic relationship:

$$
\begin{aligned}
p(U) \quad &= \quad p(\vec{z}_1|\vec{y}_1,\vec{y}_2) \\
&\cdot \quad \delta(\vec{y}_1 - \frac{q_{A_aA_b}}{1+q_{A_aA_b}} \cdot f_1(T_{A_aA_b}, P_{A_aA_b}, e_{A_aA_b}, a_{A_aA_b}, \omega_{A_aA_b}, \Omega_{A_aA_b}, i_{A_aA_b})) \\
&\cdot \quad \delta(\vec{y}_2 - f_1(T_{AB}, P_{AB}, e_{AB}, a_{AB}, \omega_{AB}, \Omega_{AB}, i_{AB})) \\
&\cdot \quad p(T_{AB}) \cdot p(P_{AB}) \cdot p(e_{AB}) \cdot p(a_{AB}) \cdot p(\omega_{AB}) \cdot p(\Omega_{AB}) \cdot p(i_{AB}) \\
&\cdot \quad p(T_{A_aA_b}) \cdot p(P_{A_aA_b}) \cdot p(e_{A_aA_b}) \cdot p(a_{A_aA_b}) \cdot p(\omega_{A_aA_b}) \cdot p(\Omega_{A_aA_b}) \cdot p(i_{A_aA_b}) \cdot p(q_{A_aA_b}), \\
\text{where} \quad &U = \{\vec{z}_2, \vec{y}_1, \vec{y}_2, T_{A_aA_b}, P_{A_aA_b}, e_{A_aA_b}, a_{A_aA_b}, \omega_{A_aA_b}, \Omega_{A_aA_b}, i_{A_aA_b}, q_{A_aA_b}, \\
&T_{AB}, P_{AB}, e_{AB}, a_{AB}, \omega_{AB}, \Omega_{AB}, i_{AB}\}
\end{aligned}
\tag{4.7}
$$

Finally, the conditional independence structures encoded in the graphical models allow to conduct the inference in a series of sequential steps illustrated in Figure (4.5). This sequential process uses the following notation:

$$
\begin{aligned}
\Theta_1 &= \{T_{A_aA_b}, P_{A_aA_b}, e_{A_aA_b}, a_{A_aA_b}, \omega_{A_aA_b}, \Omega_{A_aA_b}, i_{A_aA_b}\} \\
\Theta_2 &= \{q_{A_aA_b}, T_{AB}, P_{AB}, e_{AB}, a_{AB}, \omega_{AB}, \Omega_{AB}, i_{AB}\}
\end{aligned}
\tag{4.8}
$$

and the following steps:

1. Compute the predictive model $P_{\Theta_1|\vec{Z}_1}(\cdot|\vec{z}_1)$ using $\{\vec{z}_1(\tau_k)\}_{k=1}^{N_1}$ in a sample-based scheme.

2. Generate empirical samples of θ_1 using the posterior distribution of Θ_1 given $\vec{Z}_1$, $P_{\Theta_1|\vec{Z}_1}(\cdot|\vec{z}_1)$.

3. Generate virtual observations $\{\vec{y}_1(\tau_k)\}_{k=1}^{N_2}$, using $P_{\Theta_1|\vec{Z}_1}(\cdot|\vec{z}_1)$, for each observation epoch from $\vec{z}_2$, in an imputations framework.

4. Compute the predictive model $P_{\Theta_2|\vec{Z}_1,\vec{Z}_2}(\cdot|\vec{z}_1,\vec{z}_2)$ using $\{\vec{y}_1(\tau_k)\}_{k=1}^{N_2}$ and $\{\vec{z}_2(\tau_k)\}_{k=1}^{N_2}$ in a sample-based scheme. At the end of this stage, obtain i.i.d. samples of the posterior distributions of the whole set of parameters $\Theta_1 \cup \Theta_2$ given the observations $\{\vec{z}_1(\tau_k)\}_{k=1}^{N_1}$ and $\{\vec{z}_2(\tau_k)\}_{k=1}^{N_2}$.

More details about each subprocess can be found in Appendix C.

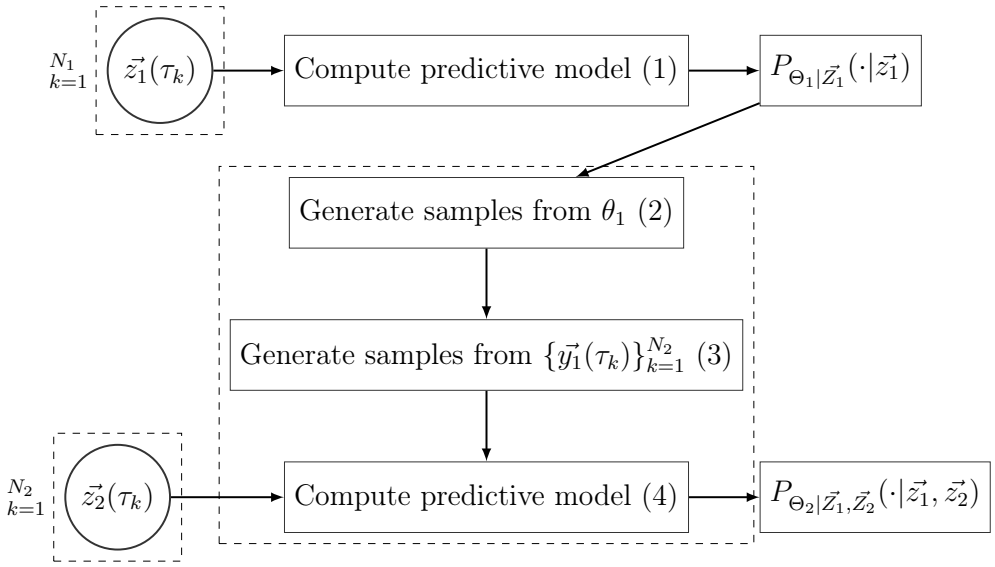

Figure 4.5: Information-flow diagram in astrometric alone scenario.

Radial Velocity Alone

Radial Velocity (RV) observations correspond to the velocity of one of the bodies involved in the triple hierarchical configuration, measured along the observer's line-of-sight. Measurements of the inner system's primary A_a (B_a) are necessary; however, measurements of A_b and B (B_b and A) are included if available. The RV measurements B only depends on outer parameters; but, as the center of mass A moves in the outer orbit, the RVs from A_a and A_b consider this movement and depend on inner and outer parameters. The observation model follows the general form from Equation (4.1), with highly non-linear functions f:

$$\begin{aligned} z_3(\tau) &= f_3(T_{A_aA_b}, P_{A_aA_b}, e_{A_aA_b}, a_{A_aA_b}, \omega_{A_aA_b}, i_{A_aA_b}, q_{A_aA_b}, \\ &\quad T_{AB}, P_{AB}, e_{AB}, a_{AB}, \omega_{AB}, \Omega_{AB}, i_{AB}, q_{AB}, \tau) + \varepsilon_3(\tau) \\ z_4(\tau) &= f_4(T_{A_aA_b}, P_{A_aA_b}, e_{A_aA_b}, a_{A_aA_b}, \omega_{A_aA_b}, i_{A_aA_b}, q_{A_aA_b}, \\ &\quad T_{AB}, P_{AB}, e_{AB}, a_{AB}, \omega_{AB}, \Omega_{AB}, i_{AB}, q_{AB}, \tau) + \varepsilon_4(\tau) \\ z_5(\tau) &= f_5(T_{AB}, P_{AB}, e_{AB}, a_{AB}, \omega_{AB}, \Omega_{AB}, i_{AB}, q_{AB}, \tau) + \varepsilon_5(\tau) \end{aligned} \tag{4.9}$$

Given the parameters involved in the observation model, this scenario allows to compute the individual stellar masses [3], but not the mutual inclination. It is worth mentioning that functions f_i are presented in full detail in Appendix A.4.

Based on the dependencies observed in Eq. (4.9), a graphical model was designed expressing the relationship between the parameters and the observations z_3 (inner system's primary), z_4 (inner system's secondary) and z_5 (outer system's primary/secondary). It is shown in Figure 4.6, where N_3, N_4 and N_5 are the number of observations of z_3, z_4 and z_5, respectively.

Even though the measurements related to the outer body (z_5) are disconnected from the inner parameters (as can be seen in Eq. (4.9) and Figure 4.6), it is not common to have a lot

[3]The parallax of the system is needed.

of observations[4]; then a preliminary "outer body" stage would involve just a few observations. On the other hand, measurements from the inner's secondary (z_4) are also uncommon, due to technical reasons mostly.

Therefore, a sequential factorization of the joint in terms of partial posteriors is not simple to develop and the modelling the complete joint distribution between parameters and observations is preferred(see Figure 4.7).

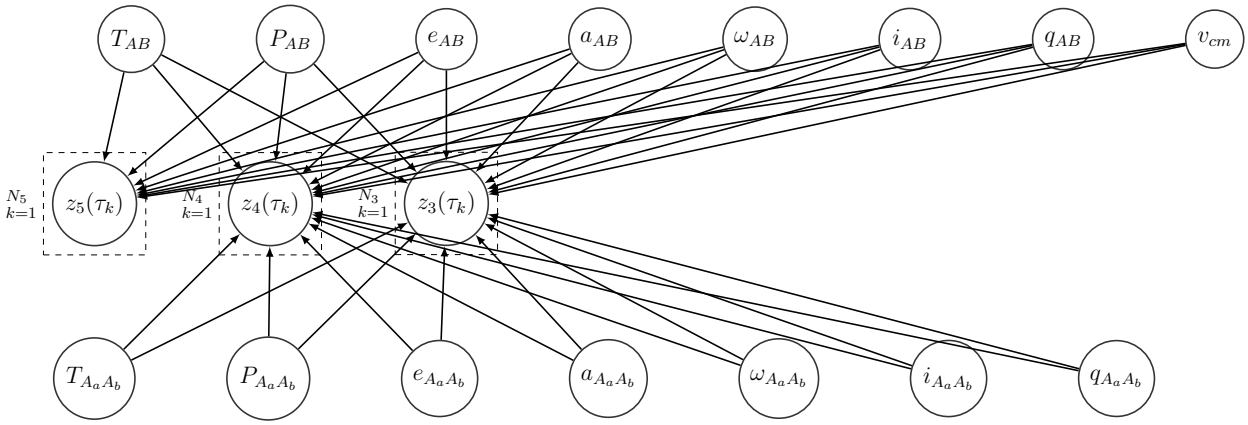

Figure 4.6: Graphical model representation of RV scenario.

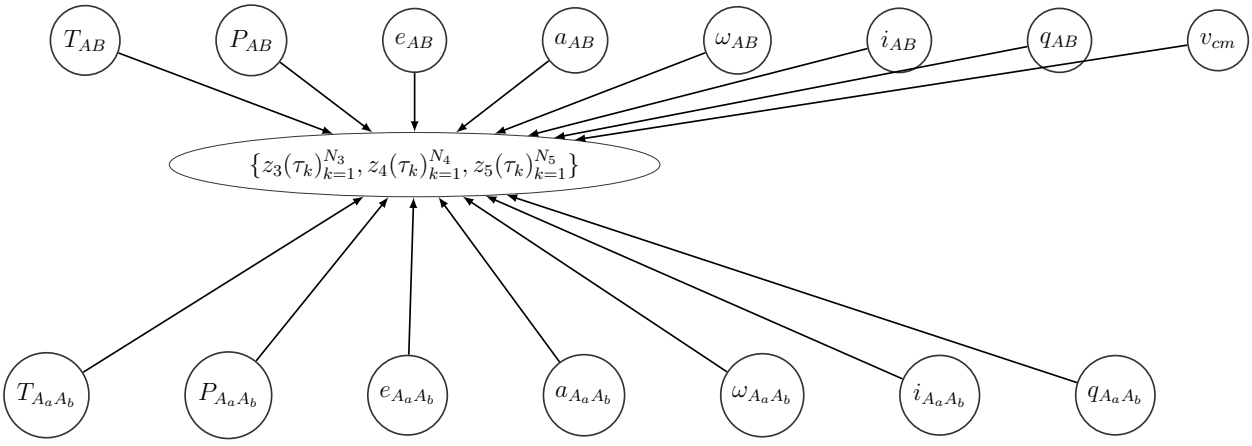

Figure 4.7: Graphical model representation of RV scenario.

That network allows to factorize the joint distribution in conditional components and, therefore, arrange the inference process for this scenario. As the graphical model is really simple, the factorization remains as Equation (4.10).

[4]Mainly due to the long outer periods.

$$
\begin{aligned}
p(U) \quad &= \quad p(Z_{RV}|T_{A_aA_b}, P_{A_aA_b}, e_{A_aA_b}, a_{A_aA_b}, \omega_{A_aA_b}, i_{A_aA_b}, q_{A_aA_b}, \\
&\qquad T_{AB}, P_{AB}, e_{AB}, a_{AB}, \omega_{AB}, i_{AB}, q_{AB}, v_{cm}) \\
&\cdot \quad p(T_{AB}) \cdot p(P_{AB}) \cdot p(e_{AB}) \cdot p(a_{AB}) \cdot p(\omega_{AB}) \cdot p(i_{AB}) \cdot p(q_{AB}) \cdot p(v_{cm}) \\
&\cdot \quad p(T_{A_aA_b}) \cdot p(P_{A_aA_b}) \cdot p(e_{A_aA_b}) \cdot p(a_{A_aA_b}) \cdot p(\Omega_{A_aA_b}) \cdot p(i_{A_aA_b}) \cdot p(q_{A_aA_b}), \\
\text{where} \quad &U = \{Z_RV, T_{A_aA_b}, P_{A_aA_b}, e_{A_aA_b}, a_{A_aA_b}, \omega_{A_aA_b}, i_{A_aA_b}, q_{A_aA_b}, \\
&\qquad T_{AB}, P_{AB}, e_{AB}, a_{AB}, \omega_{AB}, i_{AB}, q_{AB}, v_{cm}\}
\end{aligned}
\tag{4.10}
$$

In this case the inference is conducted in one step using the complete joint distribution of the problem (see Figure 4.8) and using the notation

$$
\begin{aligned}
\Theta_{RV} = \{&T_{A_aA_b}, P_{A_aA_b}, e_{A_aA_b}, a_{A_aA_b}, \omega_{A_aA_b}, i_{A_aA_b}, q_{A_aA_b}, \\
&T_{AB}, P_{AB}, e_{AB}, a_{AB}, \omega_{AB}, \Omega_{AB}, i_{AB}, q_{AB}, v_{cm}\}
\end{aligned}
$$

Concerning the predictive model $P_{\Theta_{RV}|Z_{RV}}(\cdot|z_3, z_4, z_5)$, this is approximated with samples using a sample-based scheme. More details about it can be found in Appendix C.

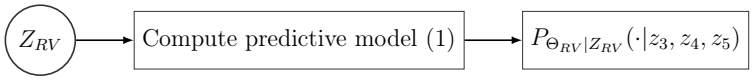

Figure 4.8: Information-flow diagram in RV alone scenario.

Combined Scenario

Finally, the scenario where both astrometric and RV measurements are available is considered. This rich scenario allows to compute the relative orbital orientation **and** the stellar masses[5]. Due to the dimensionality of this setting, the estimation is complex computationally and analytically. On the modelling side, there are several complex interdependencies between the parameters and the observations. A major effort was made in this work to encode this relationship by the graphical model presented in Figure 4.9. Consequently, a factorization as the one presented on previous simpler (unimodal) models is difficult to illustrate in a simple diagram.

For the inference, an approach that builds on the aforementioned simpler scenarios is proposed. Interestingly, the steps resemble the ones presented by [Tokovinin and Latham, 2017]. The procedure is shown in Figure 4.10 and consider the following steps:

1. Compute the predictive model $P_{\Theta_{RV}|Z_{RV}}(\cdot|z_3, z_4, z_5)$ using $\{\vec{z_3}(\tau_k)\}_{k=1}^{N_3}$, $\{\vec{z_4}(\tau_k)\}_{k=1}^{N_4}$ and $\{\vec{z_5}(\tau_k)\}_{k=1}^{N_5}$ in a sample-based scheme (1).

2. Compute the predictive model $P_{\Theta_1|\vec{Z_1}}(\cdot|\vec{z_1})$ in a sample-based scheme, using the observations $\{\vec{z_1}(\tau_k)\}_{k=1}^{N_1}$ and the posteriors from last stage as priors (2).

3. Generate empirical samples of θ_1 using the posterior distribution of Θ_1 given $\vec{Z_1}$, $P_{\Theta_1|\vec{Z_1}}(\cdot|\vec{z_1})$ (3).

[5]The parallax of the system is needed.

4. Generate virtual observations $\{\vec{y_1}(\tau_k)\}_{k=1}^{N_2}$, using $P_{\Theta_1|\vec{Z}_1}(\cdot|\vec{z_1})$, for each observation epoch from $\vec{z_2}$, in an imputations framework (4).

5. Compute the predictive model $P_{\Theta_2|\vec{Z}_1,\vec{Z}_2}(\cdot|\vec{z_1},\vec{z_2})$ using $\{\vec{y_1}(\tau_k)\}_{k=1}^{N_2}$ and $\{\vec{z_2}(\tau_k)\}_{k=1}^{N_2}$ in a sample-based scheme (5).

6. Return to (1) until a stopping criterion or the maximum amount of iterations are reached.

At the end of each stage i.i.d. samples of the posterior distribution of the parameters are obtained, partially conditioned on the set of observations associated to that stage ($\{\vec{z_1}(\tau_k)\}_{k=1}^{N_1}$, $\{\vec{z_2}(\tau_k)\}_{k=1}^{N_2}$ or Z_{RV}). Besides, with each new stage the target distribution is reached, given **all** the available observations ($\{\vec{z_1}(\tau_k)\}_{k=1}^{N_1}$, $\{\vec{z_2}(\tau_k)\}_{k=1}^{N_2}$ and Z_{RV}).

More details about each subprocess can be found in Appendix C.

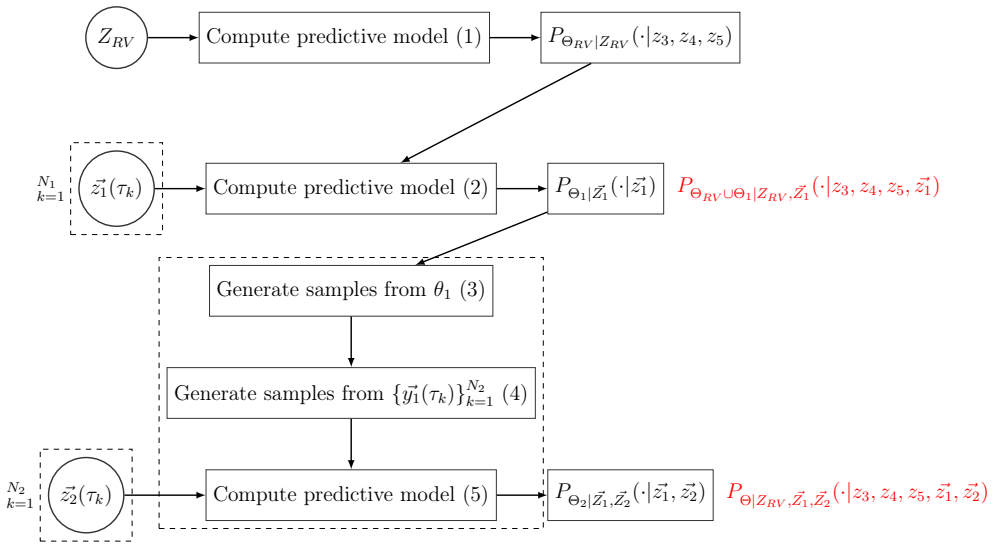

Figure 4.10: Information-flow diagram in the combined scenario.

Finally, given that the **target distribution** is unknown, and it is obtained through sampling-based methods, it is not possible to establish convergence by computing the distance between this distribution and the current estimation. This issue leads to consider as stopping criterion the difference between the current distribution $p_{S_i}(\Theta)$ and the one from the last step $p_{S_{i-1}}(\Theta)$. The selected distance measure is the **conditional variance**.

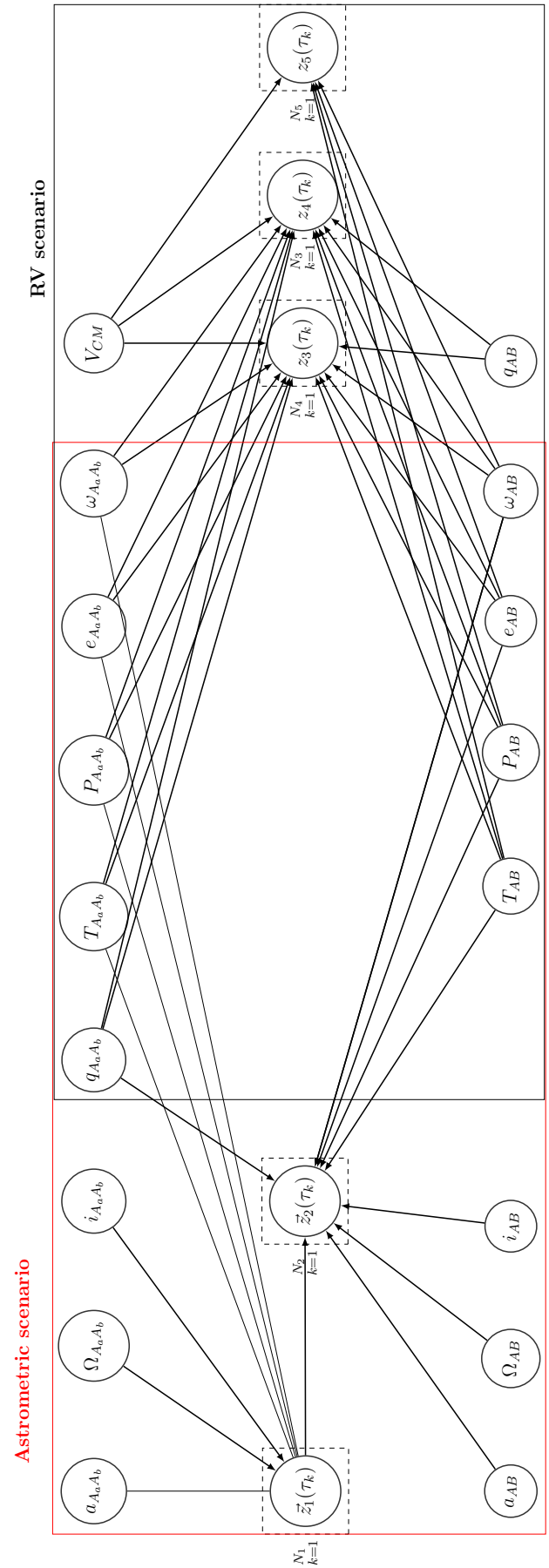

Figure 4.9: Graphical model representation considering astrometric and RV measurements at once.

Chapter 5

Results

In order to test the methodology presented in the previous sections, three well studied triple hierarchical stellar systems published by Tokovinin and collaborators [Tokovinin, 2018a, Tokovinin and Latham, 2017] are selected, and kindly provided by him upon request, then regarded as "benchmark" system from the point of view of the estimation algorithm. This section exhibits the main results of these comparisons, and a brief discussion regarding the estimation processes performed in all scenarios (astrometry only, radial velocity only, and combined scenario).

Best parameters In this work, the "best solution" is the one obtained from the *maximum a posteriori (MAP)* estimator, derived in turn from the posterior joint distribution. It is accompanied by a confidence interval obtained using "modified" quartiles over the marginal distributions: the lower (Q_1) and upper (Q_3) quartiles were computed using the MAP estimator as the median value (Q_2).

In the RV alone scenario, given the one-stage Bayesian methodology implemented in this work, MAP is equivalent to the *maximum likelihood (ML)* estimator, i.e., the particles that maximize the likelihood and, thus, minimizes the χ^2 statistic[1].

Finally, it is worth mentioning that the orbits and radial velocity curves are shown for each data set. Besides the MAP estimator, those plots take into account the 1000 most likely solutions.

Derived quantities Based on the best parameters, the mutual inclination and the sum/individual masses were obtained. It has been mentioned previously that the parallax is needed to obtain relevant physical parameters of the system, such as the sum of the masses (astrometry observations alone) or the individual ones (combined astrometry plus radial velocity observations). All the parallaxes considered in these calculations have been obtained from the Multiple Star Catalogue (MSC), a catalogue of hierarchical multiple stellar systems with three or more components[2] [Tokovinin, 2018b]. These results are shown in Figure 5.1.

[1] See [Mendez et al., 2017]

[2] http://www.ctio.noao.edu/∼atokovin/stars

Outer system's plots It is important to note a subtle issue regarding the orbit of the outer system: Its "wobble" (due to the presence of inner system) implies that its orbit, as measured from the primary, is not necessarily a closed orbit (see Appendix A1, Equations (A.11) and (A.12)). For this reason, when plotting the outer orbit the time t is established as the independent variable and not the (outer) *true anomaly* ν. The curves obtained for the following range of epochs are consecutively plotted:

- $[t - P, t]$ in green,
- $[t, t + P]$ in red: this is the curve that minimizes the O-C residual.
- $[t + P, t + 2P]$ in black.

5.1 LHS 1070 (00247-2653)

The triple system LHS 1070, also known as GJ 2005 and LP 881-64, consists of a binary (B_a, B_b) accompanied by a distant star (A). There are astrometric measurements available for the inner and outer subsystems [Köhler et al., 2012, Tokovinin, 2018a], so the first scenario from the methodology was applied.

Inner orbit: The period was known to be ~ 17 years, so a uniform prior between 10 and 100 years was set, and the same boundaries were determined to limit the exploration of the state space. The normalized periastron passage T and the eccentricity e were restricted only between their physical boundaries, $[0, 1]$ and $[0, 0.85]$, respectively. A number of 250000 iterations and a burn-in of 25000 iterations were chosen.

The best 1000 orbits can be seen in Figure 5.2 (left) and the marginal histograms in Figure 5.3.

Outer orbit: [Tokovinin, 2018a] estimated a period ~ 77 years, so a uniform prior between 70 and 90 years was set. Regarding the eccentricity, the uniform prior was bounded in $[0.001, 0.05]$, based on the value estimated by [Tokovinin, 2018a] of ~ 0.039. The lower bound was necessary because the algorithm was inclined to small eccentricities (under $10\hat{}-4$). The angles Ω and i were known ~ 13.9 and ~ 62.5 degrees, respectively, so both uniform priors were set between $[0, 100]$. On the other hand the ω prior was set between $[100, 360]$ degrees.

With respect to the physical boundaries, only the period was constrained $[0.001, 0.05]$ years. A number of 5000000 iterations and a burn-in of 3000000 iterations were chosen. The best 1000 orbits can be seen in Figure 5.2 (right) and the marginal histograms in Figure 5.5; however, Figure 5.4 shows in detail the orbits obtained for the following range of epochs:

- First row shows $[t - P, t]$ (green).
- Second row shows $[t, t + P]$ (red). This is the curve that minimizes the O-C residual.
- Third row shows $[t + P, t + 2P]$ (black).
- Finally, the last row shows the orbit obtained by [Tokovinin, 2018a].

Parameter	This work	[Tokovinin, 2018a]
T [yrs]	$2007.694^{+0.280}_{-0.459}$	$2006.440^{+0.007}_{-0.007}$
P [yrs]	$17.146^{+0.041}_{-0.043}$	$17.247^{+0.016}_{-0.016}$
e	$0.023^{+0.0034}_{-0.0031}$	$0.0172^{+0.0008}_{-0.0008}$
a (")	$0.460^{+0.0008}_{-0.0007}$	$0.4598^{+0.0007}_{-0.0007}$
ω (°)	$229.648^{+6.023}_{-9.860}$	202.53 fixed
Ω (°)	$14.745^{+0.0681}_{-0.0689}$	$14.82^{+0.12}_{-0.12}$
i (°)	$62.225^{+0.107}_{-0.101}$	$62.04^{+0.11}_{-0.11}$
q	$0.952^{+0.023}_{-0.041}$	0.942

Table 5.1: Estimated inner parameters of LHS1070

Parameter	This work	[Tokovinin, 2018a]
T [yrs]	$2049.750^{+1.720}_{-2.326}$	$2049.67^{+1.32}_{-1.32}$
P [yrs]	$77.972^{+2.101}_{-3.499}$	$77.62^{+2.10}_{-2.10}$
e	$0.037^{+0.039}_{-0.018}$	$0.039^{+0.02}_{-0.02}$
a (")	$1.531^{+0.025}_{-0.049}$	$1.528^{+0.112}_{-0.112}$
ω (°)	$212.974^{+1.826}_{-3.946}$	$210.7^{+6.4}_{-6.4}$
Ω (°)	$14.016^{+0.496}_{-0.988}$	$13.9^{+0.7}_{-0.7}$
i (°)	$62.362^{+0.467}_{-0.784}$	$62.5^{+0.4}_{-0.4}$

Table 5.2: Estimated outer parameters of LHS1070

5.2 HIP 101955 (20396+0458)

This system, also known as HD 196795 and GJ 795AB, consists of an inner subsystem A_a, A_b and a distant object B, forming the outer subsystem known as KUI 99 AB. Astrometry is available for the inner and outer orbits, and there are also radial velocity measurements for A_a, A_b and B; then, the combined scenario is performed.

Inner system: The inner system is known to be highly eccentric [Malogolovets et al., 2007, Tokovinin and Latham, 2017], then eccentricity's prior was bounded in $[0.5, 0.8]$. On

Parameter	This work	[Tokovinin and Latham, 2017]
T [yrs]	$1935.227^{+0.040}_{-0.044}$	$2000.518^{+0.004}_{-0.004}$
P [yrs]	$2.510^{+0.001}_{-0.0009}$	$2.51013^{+0.00052}_{-0.00052}$
e	$0.619^{+0.0141}_{-0.0148}$	$0.6170^{+0.0047}_{-0.0047}$
a (")	$0.124^{+0.0026}_{-0.0022}$	$0.1242^{+0.0011}_{-0.0011}$
ω (°)	$105.017^{+13.225}_{-4.538}$	$109.7^{+1.8}_{-1.8}$
Ω (°)	$151.284^{+5.639}_{-12.887}$	$147.1^{+1.8}_{-1.8}$
i (°)	$24.831^{+3.473}_{-5.184}$	$24.1^{+1.7}_{-1.7}$
q	$0.805^{+0.042}_{-0.036}$	0.84 (astrometric) 0.45 (spectroscopic)

Table 5.3: Estimated inner parameters of KUI99

the other hand, since [Duquennoy, 1987] study it is known that the period is ~ 2.5, so P's prior was bounded in $[1.5, 3.5]$.

Due to the inconsistencies in the data, [Tokovinin and Latham, 2017] found different q regarding the wobble's astrometric information and the one from spectroscopy, that is 0.84 and 0.45. As this method does not consider the amplitudes K_1 and K_2 as independent variables, q was set between $[0.7, 0.95]$, which is consistent with the fractional mass f in $[0.42, 0.48]$ and the value of 0.8 from [Malogolovets et al., 2007].

On the other side, all the speckle measurements before 1981 were considered, but the associated error was increased considerably with respect of the other measurements. Besides, there is an issue about the blending of the A_a and B components in the reported RVs, so it must be taken into account in the results. The best 1000 inner orbits can be seen in Figure 5.6 (left) and the MAP's inner RV curve in Figure 5.7 (left).

Outer system: The period was known to be ~ 40 years, with small eccentricity and highly inclined [Baize, 1981, Heintz, 1984, Malogolovets et al., 2007]. Then, the priors from those parameters were set in $[35, 45]$, $[0, 0.25]$ and $[0, 180]$, respectively.

The best 1000 outer orbits can be seen in Figure 5.6 (right) and Figure 5.8 shows in detail the orbits obtained for the following range of epochs:

- First row shows $[t - P, t]$ (green).
- Second row shows $[t, t + P]$ (red). This is the curve that minimizes the O-C residual.
- Third row shows $[t + P, t + 2P]$ (black).

Parameter	This work	[Tokovinin and Latham, 2017]
T [yrs]	$1938.605^{+0.129}_{-0.266}$	$2016.110^{+1.32}_{-1.32}$
P [yrs]	$38.641^{+0.055}_{-0.025}$	$38.6790^{+0.031}_{-0.031}$
e	$0.114^{+0.003}_{-0.004}$	$0.118^{+0.016}_{-0.016}$
a (")	$0.855^{+0.002}_{-0.003}$	$0.855^{+0.110}_{-0.110}$
ω (°)	$231.432^{+0.972}_{-1.909}$	$233.4^{+0.5}_{-0.5}$
Ω (°)	$127.632^{+0.161}_{-0.172}$	$127.6^{+0.08}_{-0.08}$
i (°)	$87.477^{+0.096}_{-0.106}$	$87.40^{+0.05}_{-0.05}$
q	$0.445^{+0.277}_{-0.207}$	
v_{cm} $(\frac{km}{s})$	$-41.081^{+1.635}_{-1.301}$	$-41.11^{+0.08}_{-0.08}$

Table 5.4: Estimated outer parameters of KUI99

- Finally, the last row shows the orbit obtained by [Tokovinin and Latham, 2017].

The MAP's outer RV curve in Figure 5.7 (first row, right).

5.3 HIP 111805 (22388+4419)

This system, also known as HD 214608, consists of an inner subsystem B_a, B_b and a distant object A. Astrometry is available for the inner and outer orbits, and there are also radial velocity measurements for B_a, B_b and A; then, the combined scenario is performed.

Inner system: The inner system is also known as HO 265 and it was recognized as highly inclined and with a small eccentricity [Balega et al., 2002, Duquennoy, 1987, Tokovinin and Latham, 2017], so the priors were set in $[0.01, 0.08]$ and $[80, 100]$. The period was known around 551 days, so the starting value was fixed to 1.5 years, but set free in $[0.1, 10]$.

On the other hand, [Tokovinin and Latham, 2017] indicated that the RV measurements presented a blend and denoted them as noisy. They also found different mass ratios q for the visual and spectroscopic data. Based on previous results, the prior was set in $[0.5, 0.9]$.

The best 1000 inner orbits can be seen in Figure 5.9 (left) and the MAP's inner RV curve in Figure 5.10 (left). Due to the visual inner information, the algorithm was inclined to small eccentricities (under 10^{-2}), which disagreed with the spectroscopic information. That issue can be seen in the fit in the radial velocity curves.

Parameter	This work	[Tokovinin and Latham, 2017]
T [yrs]	$1901.497^{+0.566}_{-0.284}$	$1986.093^{+0.093}_{-0.093}$
P [yrs]	$1.498^{+0.106}_{-0.014}$	$1.5012^{+0.0004}_{-0.0004}$
e	$0.010^{+0.034}_{-0.0005}$	$0.022^{+0.011}_{-0.011}$
a (")	$0.041^{+0.004}_{-0.002}$	$0.0385^{+0.0010}_{-0.0010}$
ω (°)	$257.039^{+51.659}_{-128.699}$	$232.9^{+22.3}_{-22.3}$
Ω (°)	$154.208^{+1.679}_{-0.268}$	$334.5^{+1.0}_{-1.0}$
i (°)	$89.565^{+2.503}_{-0.667}$	$87.40^{+0.05}_{-0.05}$
q	$0.620^{+0.055}_{-0.061}$	0.60 (astrometric) 0.68 (spectroscopic)

Table 5.5: Estimated inner parameters of HIP111805

Outer system: The outer system is also known as HDO 295 or ADS16138, with a period of $\sim$ 30 years [Duquennoy, 1987, Hough, 1890, Tokovinin and Latham, 2017]. Then, the starting value for P was fixed to 30 years, but set free in $[10, 100]$, and the prior for i was set in $[80, 100]$. The eccentricity was known to be ~ 0.3, so the prior was in $[0.2, 0.5]$. Finally, the mass ratio q prior was set in $[0.6, 0.8]$.

The best 1000 outer orbits can be seen in Figure 5.9 (right) and Figure 5.11 shows in detail the orbits obtained for the following range of epochs:

- First row shows $[t - P, t]$ (green).
- Second row shows $[t, t + P]$ (red). This is the curve that minimizes the O-C residual.
- Third row shows $[t + P, t + 2P]$ (black).
- Finally, the last row shows the orbit obtained by [Tokovinin and Latham, 2017].

The MAP's outer RV curve in Figure 5.10 (right).

Parameter	This work	[Tokovinin and Latham, 2017]
T [yrs]	$1915.827^{+0.218}_{-0.343}$	$2010.179^{+0.073}_{-0.073}$
P [yrs]	$31.551^{+0.011}_{-0.159}$	$30.127^{+0.031}_{-0.031}$
e	$0.328^{+0.004}_{-0.007}$	$0.324^{+0.004}_{-0.004}$
a (")	$0.335^{+0.001}_{-0.0004}$	$0.3361^{+0.0015}_{-0.0015}$
ω ($^\circ$)	$80.887^{+0.295}_{-0.527}$	$84.92^{+0.18}_{-0.18}$
Ω ($^\circ$)	$154.170^{+0.205}_{-0.090}$	$154.25^{+0.09}_{-0.09}$
i ($^\circ$)	$88.281^{+0.154}_{-0.127}$	$88.28^{+0.10}_{-0.10}$
q	$0.661^{+0.066}_{-0.052}$	0.70465116279
v_{cm} ($\frac{km}{s}$)	$-22.524^{+0.144}_{-0.140}$	$-22.58^{+0.08}_{-0.08}$

Table 5.6: Estimated outer parameters of HIP111805

Object	$\bar{\omega}$	Φ	M_{A_a}	M_{A_b}	M_A	M_B	M_T
WDS	($''$)	($^\circ$)	($M_\odot$)	($M_\odot$)	($M_\odot$)	($M_\odot$)	($M_\odot$)
00247-2653	0.129	0.6599714262898655	0.07948461203853474	0.07571122422674546	0.11985618686	0.1551958362652802	0.2750520231292031
20396+0458	0.0598	64.89736574784581	0.7964185854545015	0.6414285078727199	1.35933489706355	0.6049111174394143	1.9642460145
22388+4419	0.0232	1.2839248762091087	1.5483373981680584	0.9610113575600063	1.8298207990538304	1.2106659784250304	3.04048677748

Figure 5.1: Derived quantities estimators.

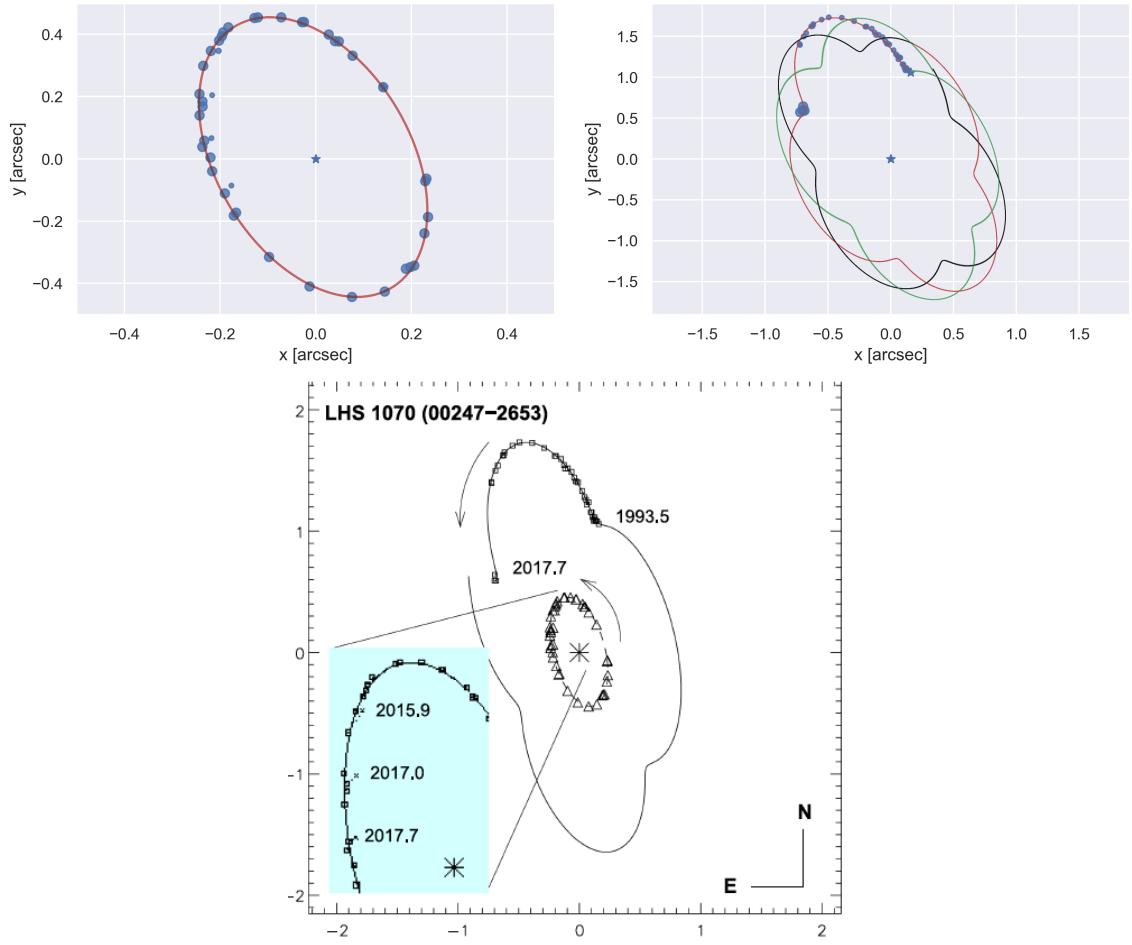

Figure 5.2: The first row shows the inner (left) and outer (right) orbits of LHS1070. The second row shows the results obtained by [Tokovinin, 2018a].

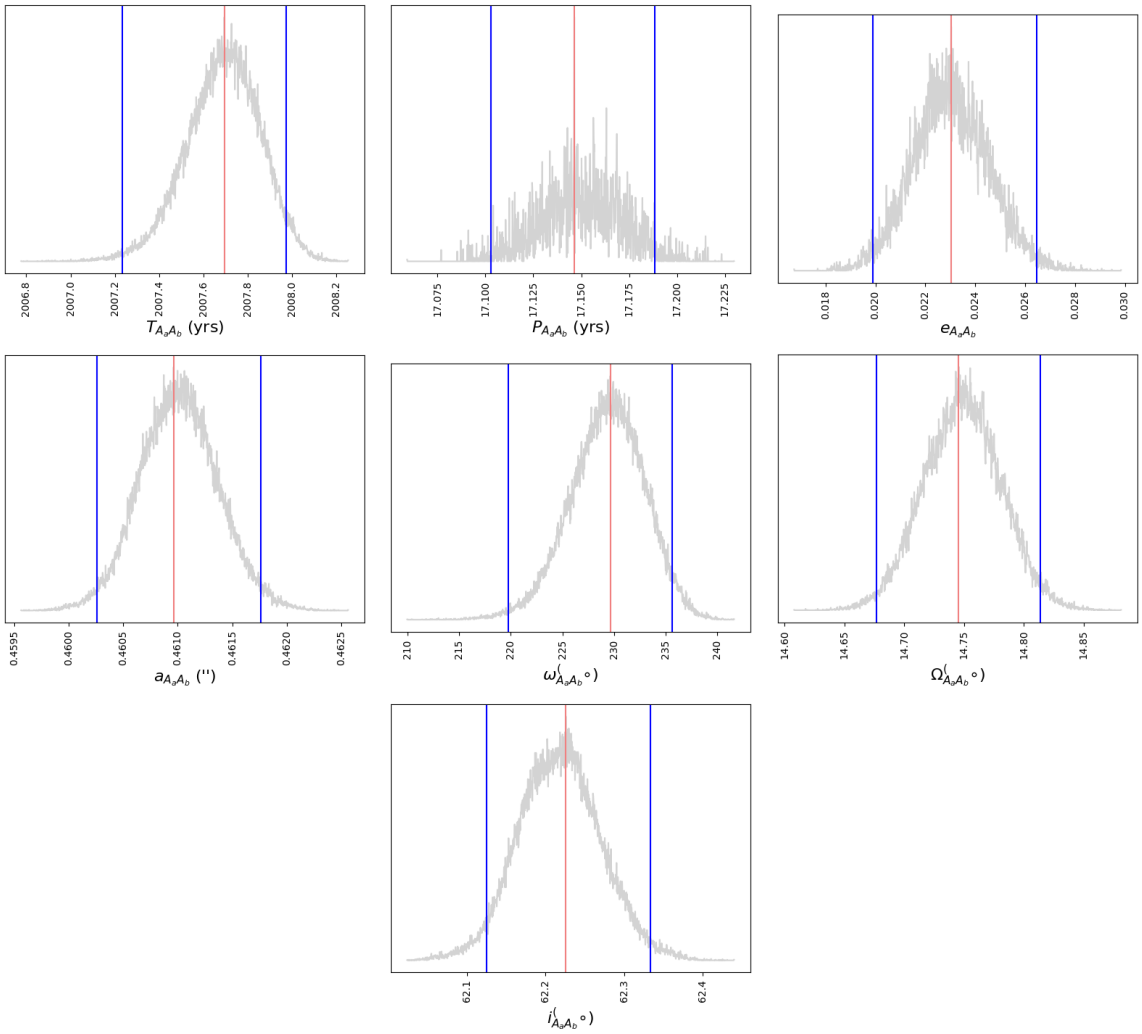

Figure 5.3: Marginal empirical distributions obtained after performing the first stage (inner orbit) for LHS1070. The MAP estimator (from the joint distribution) is indicated in red and the lower and upper quartiles are shown in blue.

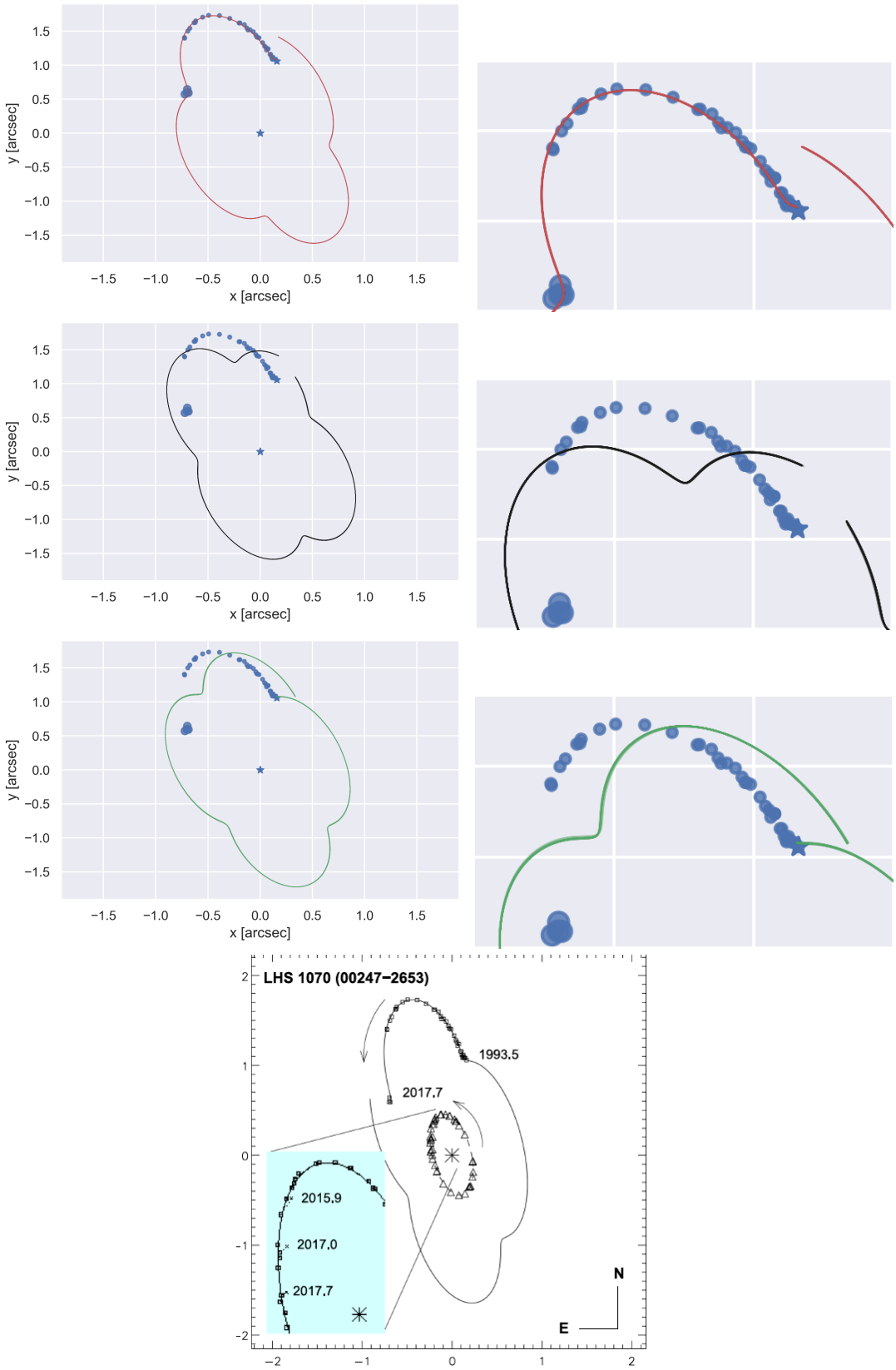

Figure 5.4: Detail of the outer astrometric orbit of LHS1070

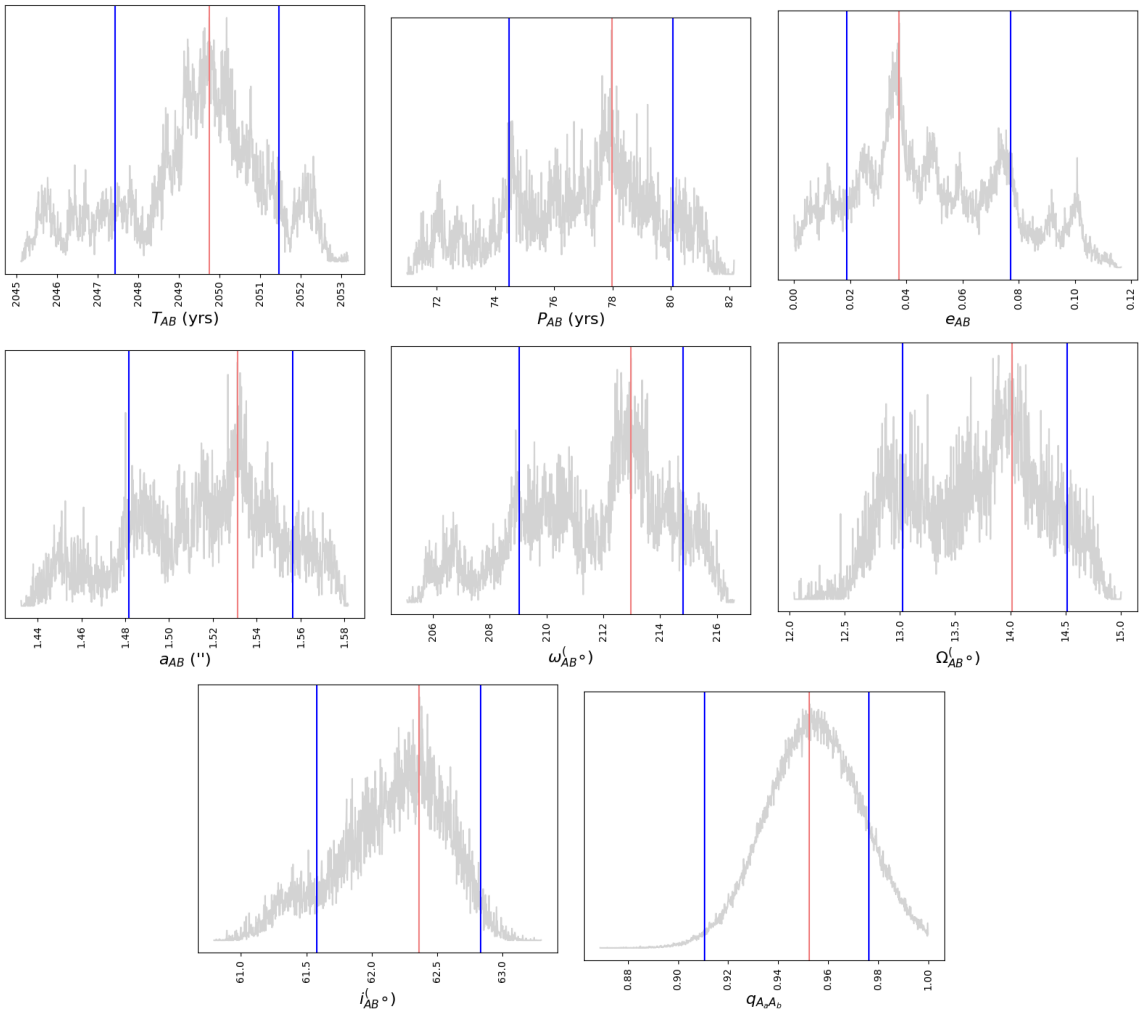

Figure 5.5: Marginal empirical distributions obtained after performing the second stage (outer orbit) for LHS1070. The MAP estimator (from the joint distribution) is indicated in red and the lower and upper quartiles are shown in blue.

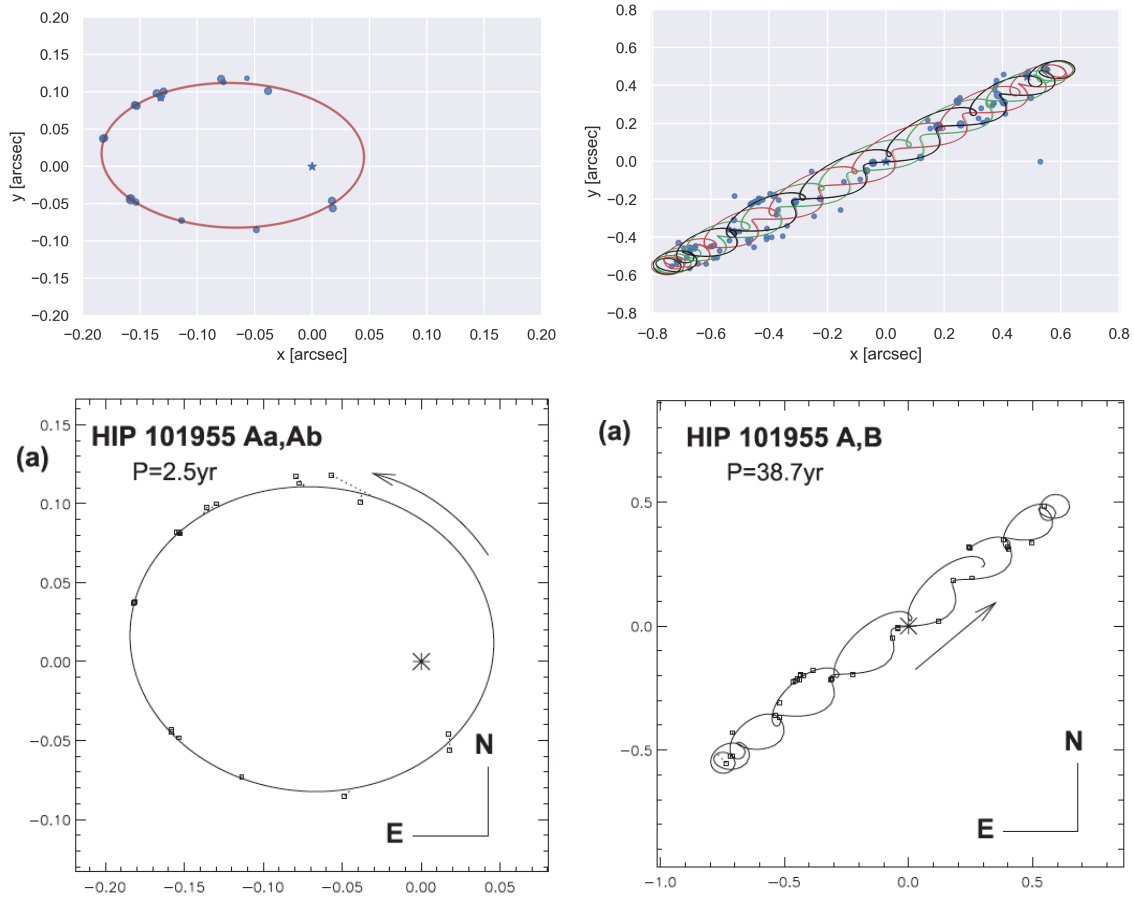

Figure 5.6: The first row shows the inner (left) and outer (right) orbits of HIP101955. The second row shows the results obtained by [Tokovinin and Latham, 2017].

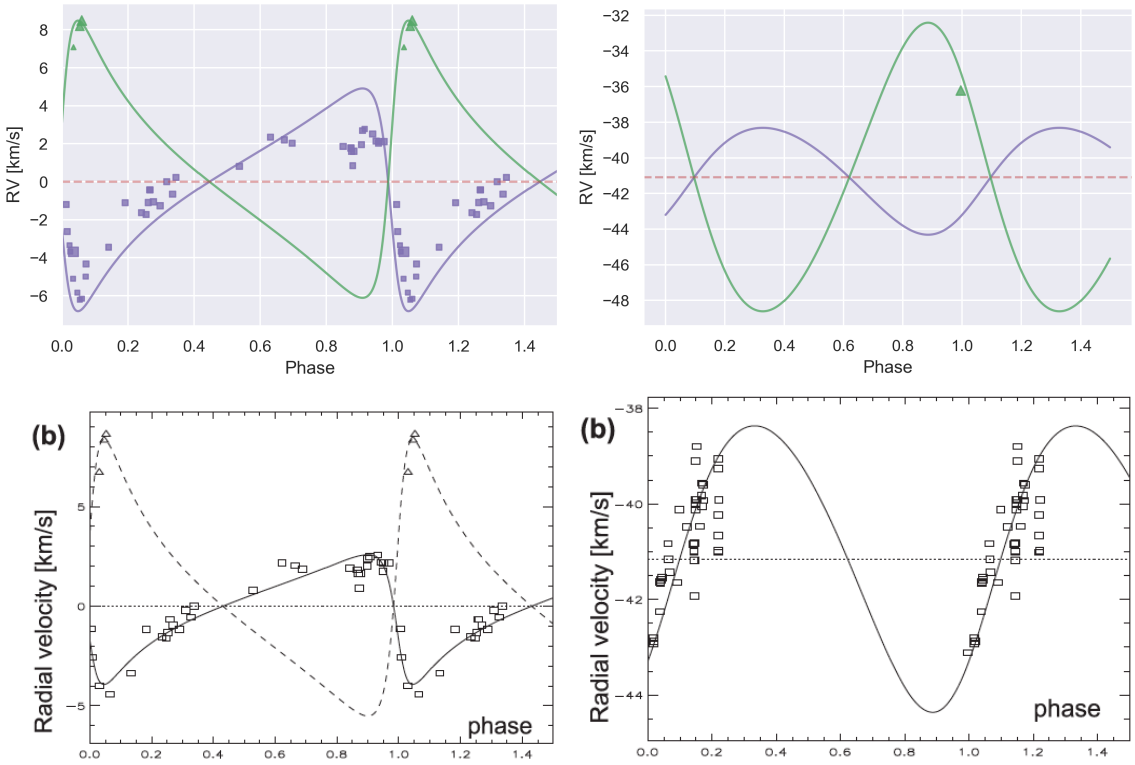

Figure 5.7: The first row shows the inner (left) and outer (right) RV curves of HIP101955. The second row shows the results obtained by [Tokovinin and Latham, 2017].

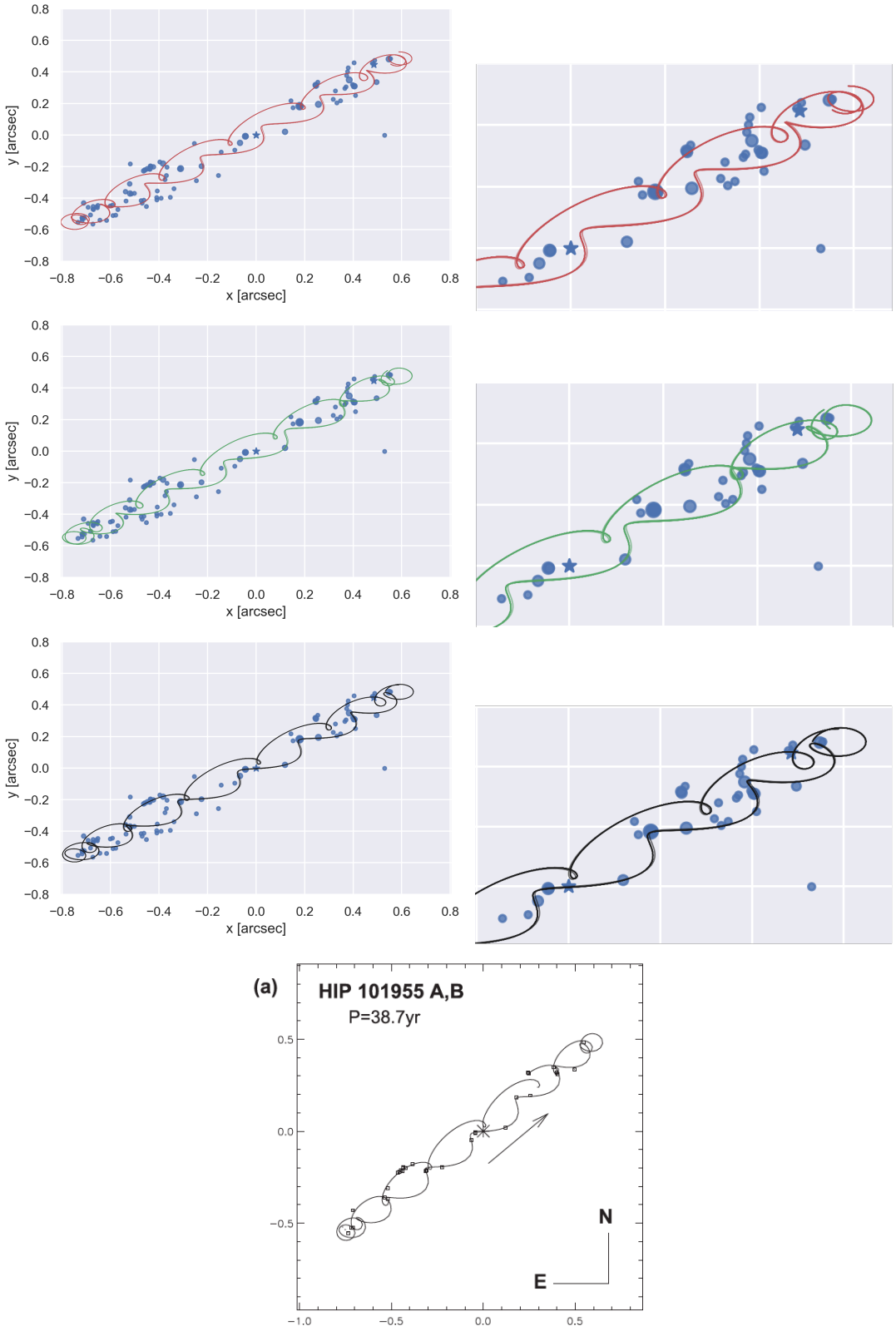

Figure 5.8: Detail of the outer astrometric orbit of HIP101955

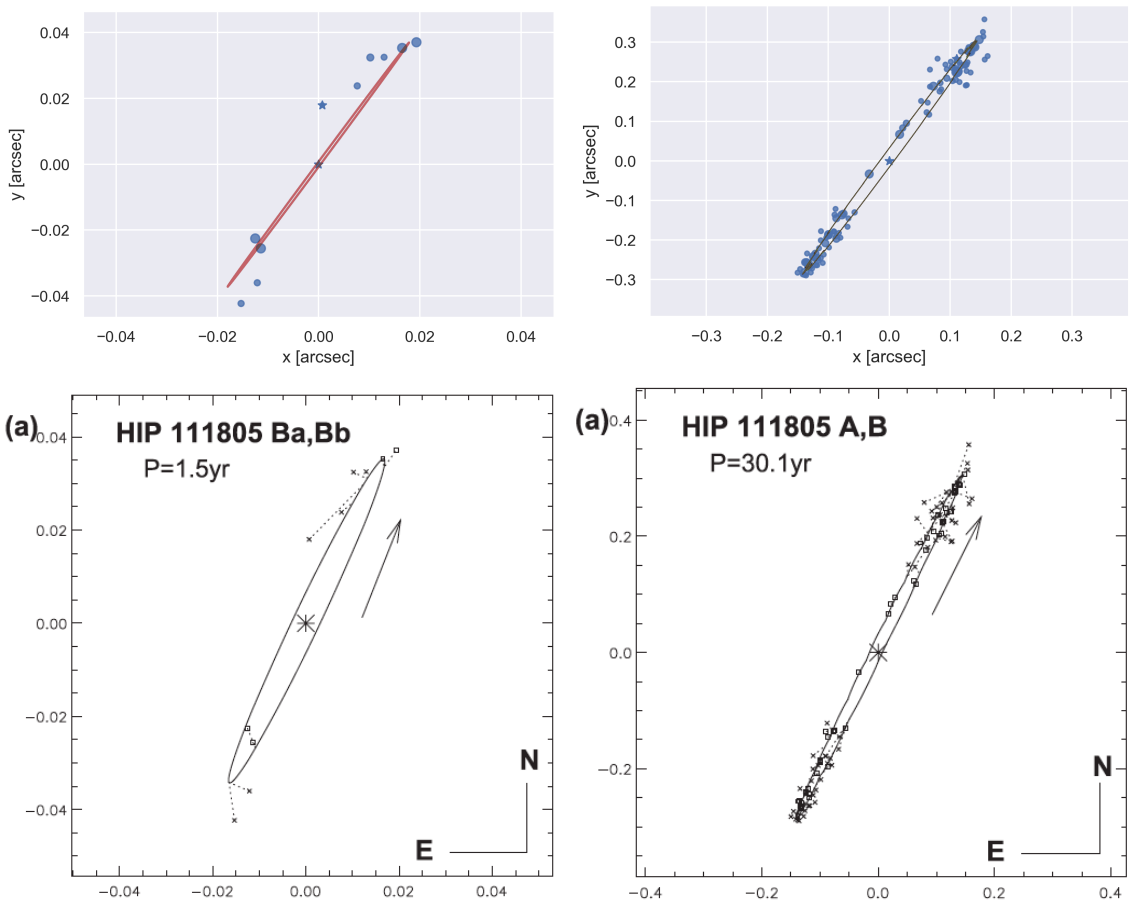

Figure 5.9: The first row shows the inner (left) and outer (right) orbits of HIP111805. The second row shows the results obtained by [Tokovinin and Latham, 2017].

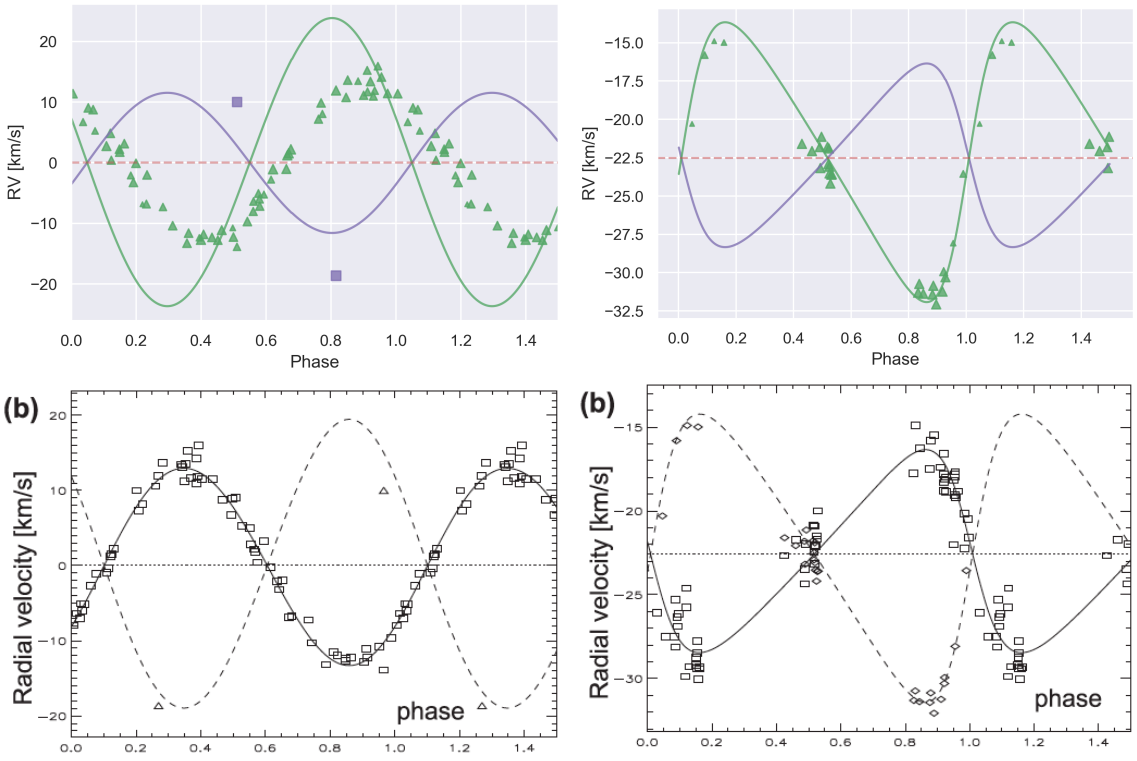

Figure 5.10: The first row shoes the Inner (left) and outer (right) RV curves of HIP111805. The second row shows the results obtained by [Tokovinin and Latham, 2017].

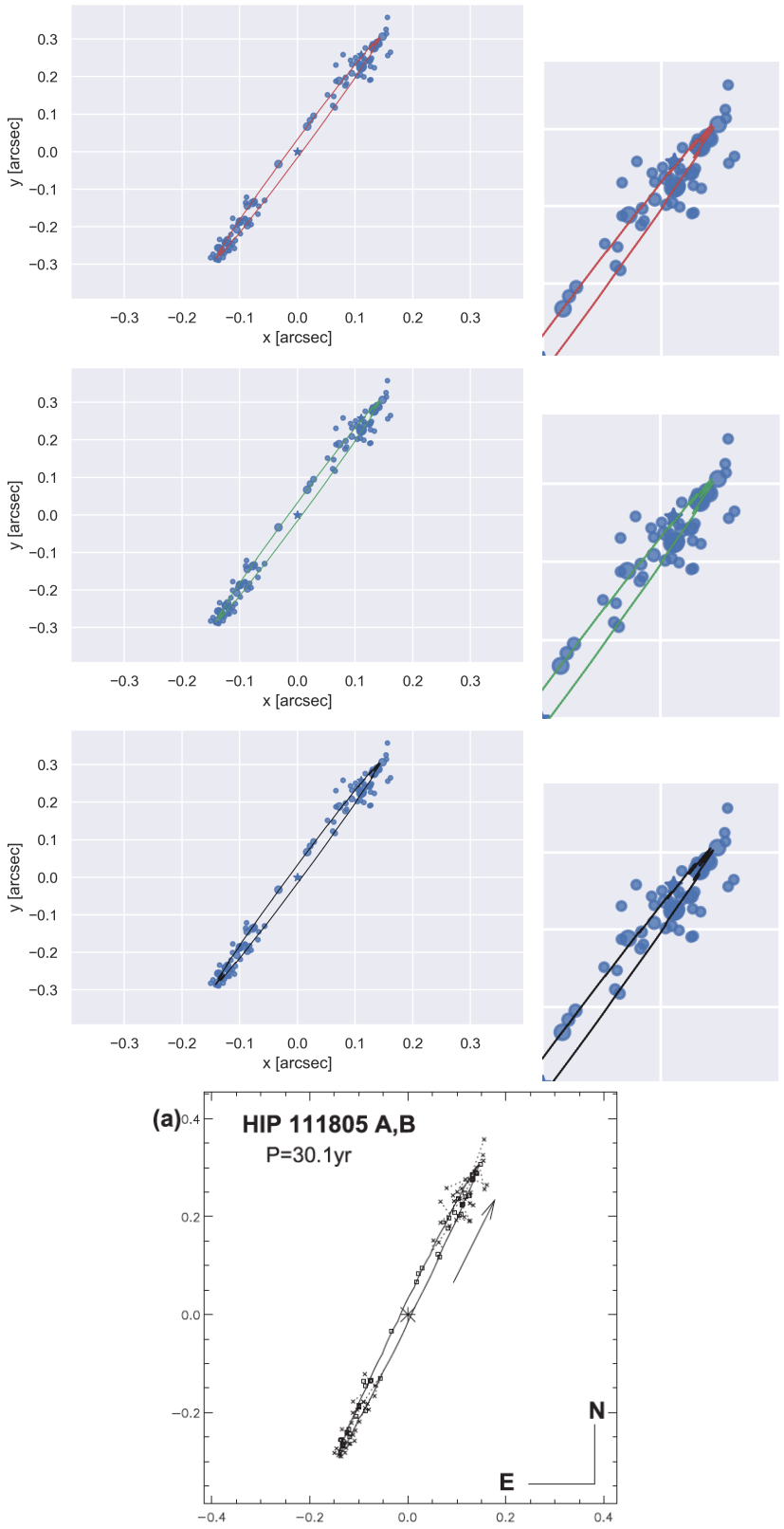

Figure 5.11: Detail of the outer astrometric orbit of HIP111805

Chapter 6

Discussion

It is challenging to work with sample-based schemes in high dimensions in terms of convergence and fine-tuning of the algorithm hyperparameters. In this work, the estimation is performed using a hybrid between MCMC and the Gibbs sampler, exploring the state space using a random walk. The estimations were successfully performed in all scenarios; however, a few considerations must be taken.

Convergence Even though the algorithm reaches the target distribution in the limit of long runs, it gets too many iterations to converge. For that reason, other options should be evaluated, like simulated tempering or parallel tempering [Earl and Deem, 2005, Marinari and Parisi, 1992]. On the other hand, methods that employ the gradient of the prior and the likelihood could be explored, such as MALA [Robert and Casella, 2004, Roberts et al., 1996], Hamiltonian Monte Carlo [Neal et al., 2011], proximal MCMC algorithms [Combettes and Pesquet, 2011, Parikh et al., 2014] or diffusion-based MCMC algorithms [Herbei et al., 2017], so local optima must be taken into account. Finally, Bayesian methods different to MCMC could be considered, as the rejection-sampling method from [Blunt et al., 2017].

Fine tuning of proposals and priors On the other hand, the hand-tuning of the priors and the proposals is a demanding task, and it is important because those hyperparameters are directly related to the algorithm's convergence.

Regarding the proposals, the variance for each one of the dimensions involved must be chosen. If it is too wide, a lot of particles where $\pi = 0$ could be chosen; if it is too small, most of the particles are accepted, so the chain moves slowly. Besides, the problem could worsen because of possible correlations between the parameters [Sharma, 2017], which are not considered at the moment. Despite rules of thumb could be applied (as $c = \sigma_2 s_d (J^T J)^{-1}$), the Jacobian must be computed, and that is not easy for highly non-linear functions as the one seen in this work. Hence, other alternatives as adaptive MCMC methods could be used. Even though those algorithms are not Markovian anymore, they do not rely on one pre-set variance and the new particle is chosen based on the earlier history of the chain [Haario et al., 2004, 2005, 2001].

On the other hand, concerning the priors, lots of particles were rejected due to physical

restrictions. This could be known beforehand by considering this information when adjusting the priors, and stop rejecting samples due to infeasible masses. Later on, this could be applied in unstable zones or another type of physical restriction.

Physical restrictions There are many restrictions related to the dynamics of triple hierarchical star systems that are included in the implemented algorithms. First of all, all of the parameters are bounded within a certain range, which could make the code to stick in the boundaries. Thus, the state space was adapted as a circular one, keeping the sampling procedure as there were no restrictions.

Despite that, the problem arises when dimensionality reductions were performed. The minimization of χ^2 statistic was conducted just using weighted least squares, so the particles were rejected after that process, which increases the computational cost and keeps the chain from converging. This issue could be fixed by solving a non-linear optimization problem with restrictions.

Besides, the imputations framework also rejects particles that violate the restrictions. That could be considered before sampling the posterior from the last step (the astrometric inner orbit) and thus prevent the situation.

Inconsistent Data Finally, inconsistencies between the astrometry and the radial velocities must be addressed. This could be solved by just not considering the ambiguous information or by adjusting the weights associated with those observations.

Chapter 7

Conclusions

The purpose of this work has been to apply a Bayesian MCMC-based methodology to the problem of estimating the orbital parameters in triple hierarchical stellar systems, that already have a measured parallax. Graphical models were employed for modelling the probabilistic relationship between parameters and observations in the astrometry-alone, radial-velocity-alone and the combined scenarios. Thus, the joint distribution is factorized in terms of independent blocks and then performed the estimation in a two-stage process, combining different sets of observations sequentially.

The framework provides MAP estimates along with the full joint posteriors of the parameters, given the observations, allowing to assess the uncertainties robustly. It requires prior knowledge about the system, although non-informative priors[1] could be used to get good results.

Regarding the radial velocities-alone scenario, a mathematical formalism is introduced, motivated by the works of [Wright and Howard, 2009] and [Mendez et al., 2017] in the context of exoplanets and visual-and-spectroscopic binaries, respectively, and adapted to the specific case of triple systems. It consists of a dimensionality reduction (15 to 10 parameters) using weighted-least squares, which allows to sample from a subset of the parameter space, hence reducing the computational cost. This is applied for systems of the form A_a, A_b - B and B_a, B_b - A.

On the other hand, the methodology is useful for outer long periods, because, even though there are just a few (or no) measurements from the distant body B (or A), the algorithm allows to constrain the mass ratio of the outer system q_{AB}.

This scheme is tested with real measurements of astrometry and radial velocities, where the inner and outer orbital elements are determined. By utilizing both kinds of measurements, these results allow to determine the mutual inclination of the orbits and the individual stellar masses, solutions that are consistent with former reported results.

[1] For example, a uniform prior

Chapter 8

Future work

Parallax and derived quantities uncertainty The parallax can be considered as a gaussian random variable, taking into account the means and variances indicated in the catalog from [Tokovinin, 2018b]. Then, through a procedure based on transformations of random variables, a posterior distribution for the individual/sum of masses and the mutual inclination could be obtained.

Radial velocities amplitude Even though the radial velocity amplitudes (K_1, K_2, K_3 and K_4) can be described in terms of some orbital parameters (details in Appendix A.4), some works as [Tokovinin and Latham, 2017, Wright and Howard, 2009] rather choose to consider them as independent parameters.

Thus, in the combined data scenario (astrometry and radial velocity information) two different inner mass ratios $q_{A_a A_b}$ are obtained:

- One from the wobble $f_{A_a A_b}$ computed using the outer astrometric information.
- One from the amplitudes K_1 and K_2, computed using the inner radial velocity information.

Although an ambiguity is presented, this variant of the algorithm should be implemented to offer a fair comparison of the estimated orbital parameters.

Belief propagation In this work, graphical models were mainly used to **show** the relationships between parameters and observations, and the estimation of the parameters was performed detaching the different observation sources, supplemented with virtual sensors techniques (imputation theory). However, more advanced belief propagation techniques can be used, exploiting directly the structure from the graphical model.

Bibliography

Andrieu, C., De Freitas, N., Doucet, A., and Jordan, M. I. (2003). An introduction to mcmc for machine learning. *Machine learning*, 50(1-2):5–43.

Baize, P. (1981). The orbital elements of 43 visual binary stars. *Astronomy and Astrophysics Supplement Series*, 44:199–208.

Balega, I., Balega, Y. Y., Hofmann, K.-H., Maksimov, A., Pluzhnik, E., Schertl, D., Shkhagosheva, Z., and Weigelt, G. (2002). Speckle interferometry of nearby multiple stars. *Astronomy & Astrophysics*, 385(1):87–93.

Bishop, C. M. (2006). *Pattern recognition and machine learning*. springer.

Blunt, S., Nielsen, E. L., De Rosa, R. J., Konopacky, Q. M., Ryan, D., Wang, J. J., Pueyo, L., Rameau, J., Marois, C., Marchis, F., et al. (2017). Orbits for the impatient: A bayesian rejection-sampling method for quickly fitting the orbits of long-period exoplanets. *The Astronomical Journal*, 153(5):229.

Brémaud, P. (2013). *Markov chains: Gibbs fields, Monte Carlo simulation, and queues*, volume 31. Springer Science & Business Media.

Claveria, R. M., Mendez, R. A., Silva, J. F., and Orchard, M. E. (2019). Visual binary stars with partially missing data: Introducing multiple imputation in astrometric analysis. *Publications of the Astronomical Society of the Pacific*, 131(1002):084502.

Combettes, P. L. and Pesquet, J.-C. (2011). Proximal splitting methods in signal processing. In *Fixed-point algorithms for inverse problems in science and engineering*, pages 185–212. Springer.

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, 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.

Duquennoy, A. (1987). A study of multiple stellar systems with coravel. *Astronomy and Astrophysics*, 178:114–130.

Earl, D. J. and Deem, M. W. (2005). Parallel tempering: Theory, applications, and new perspectives. *Physical Chemistry Chemical Physics*, 7(23):3910–3916.

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

Ford, E. B. and Gregory, P. C. (2006). Bayesian model selection and extrasolar planet detection. *arXiv preprint astro-ph/0608328*.

Gamerman, D. and Lopes, H. F. (2006). *Markov chain Monte Carlo: stochastic simulation for Bayesian inference*. Chapman and Hall/CRC.

Gelman, A., Rubin, D. B., et al. (1992). Inference from iterative simulation using multiple sequences. *Statistical science*, 7(4):457–472.

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

Gregory, P. C. (2009). Detecting extra-solar planets with a bayesian hybrid mcmc kepler periodogram. *arXiv preprint arXiv:0902.2014*.

Gregory, P. C. (2010). 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.

Gregory, P. C. (2011). Bayesian re-analysis of the gliese 581 exoplanet system. *Monthly Notices of the Royal Astronomical Society*, 415(3):2523–2545.

Haario, H., Laine, M., Lehtinen, M., Saksman, E., and Tamminen, J. (2004). Markov chain monte carlo methods for high dimensional inversion in remote sensing. *Journal of the Royal Statistical Society: series B (statistical methodology)*, 66(3):591–607.

Haario, H., Saksman, E., and Tamminen, J. (2005). Componentwise adaptation for high dimensional mcmc. *Computational Statistics*, 20(2):265–273.

Haario, H., Saksman, E., Tamminen, J., et al. (2001). An adaptive metropolis algorithm. *Bernoulli*, 7(2):223–242.

Heintz, W. (1978). *Double stars*. Geophysics & Astrophysics Monographs: No. 15. D. Reidel Publishing Company.

Heintz, W. (1984). Orbits of 15 visual binaries. *Astronomy and Astrophysics Supplement Series*, 56:5–10.

Herbei, R., Paul, R., and Berliner, L. M. (2017). Applying diffusion-based markov chain monte carlo. *PloS one*, 12(3).

Hough, G. (1890). Catalogue of 94 new double stars and measures of 107 double stars. *Astronomische Nachrichten*, 125:1.

Jordan, M. I. (1998). *Learning in graphical models*, volume 89. Springer Science & Business Media.

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

Koller, D., Friedman, N., Getoor, L., and Taskar, B. (2007). Graphical models in a nutshell. *Introduction to statistical relational learning*, 43.

Lane, B. F., Muterspaugh, M. W., Griffin, R., Scarfe, C., Fekel, F. C., Williamson, M. H., Eaton, J. A., Shao, M., Colavita, M., and Konacki, M. (2014). The orbits of the triple-star system 1 geminorum from phases differential astrometry and spectroscopy. *The Astrophysical Journal*, 783(1):3.

Leonard, P. (2000). Multiple stellar systems: Types and stability. *Encyclopedia of Astronomy and Astrophysics*.

Liu, J. S. (2008). *Monte Carlo strategies in scientific computing*. Springer Science & Business Media.

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

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

Malogolovets, E. V., Balega, Y. Y., and Rastegaev, D. (2007). Nearby low-mass triple system gj 795. *Astrophysical Bulletin*, 62(2):111–116.

Marinari, E. and Parisi, G. (1992). Simulated tempering: a new monte carlo scheme. *EPL (Europhysics Letters)*, 19(6):451.

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.

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.

Naoz, S. (2016). The eccentric kozai-lidov effect and its applications. *Annual Review of Astronomy and Astrophysics*, 54:441–489.

Neal, R. M. et al. (2011). Mcmc using hamiltonian dynamics. *Handbook of markov chain*

monte carlo, 2(11):2.

Parikh, N., Boyd, S., et al. (2014). Proximal algorithms. *Foundations and Trends® in Optimization*, 1(3):127–239.

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

Retired, A. (2010). Stars and their companions. iv. seven jovian exoplanets from keck observatory john asher johnson et al. *Publications of the Astronomical Society of the Pacific*, 122:701.

Robert, C. and Casella, G. (2004). Monte carlo statistical methods springer-verlag. *New York*.

Roberts, G. O., Tweedie, R. L., et al. (1996). Exponential convergence of langevin distributions and their discrete approximations. *Bernoulli*, 2(4):341–363.

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.

Sharma, S. (2017). Markov chain monte carlo methods for bayesian data analysis in astronomy. *Annual Review of Astronomy and Astrophysics*, 55:213–259.

Steves, B., Hendry, M., and Cameron, A. C. (2010). *Extra-Solar Planets: The Detection, Formation, Evolution and Dynamics of Planetary Systems*. CRC Press.

Tanner, M. A. and Wong, W. H. (1987). The calculation of posterior distributions by data augmentation. *Journal of the American statistical Association*, 82(398):528–540.

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

Tokovinin, A. (2018b). The updated multiple star catalog. *The Astrophysical Journal Supplement Series*, 235(1):6.

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

Torres, G., Neuhäuser, R., and Guenther, E. W. (2002). Spectroscopic binaries in a sample of rosat x-ray sources south of the taurus molecular clouds. *The Astronomical Journal*, 123(3):1701.

Wright, J. and Howard, A. (2009). Efficient fitting of multiplanet keplerian models to radial velocity and astrometry data. *The Astrophysical Journal Supplement Series*, 182(1):205.

Appendix A

Triple Hierarchical Stellar Systems Model Equations

Hierarchical stellar systems are a particular case of the general n-body problem, because it can be separated into $(n-1)$ subgroups, where each hierarchy level can be treated as a binary system separately [Leonard, 2000]. Thus, triple hierarchical stellar systems are approximated with two Keplerian orbits on top of each other; where one represents the motion of the wide system and other that of the inner/tighter system.

More precisely, those systems consist of an (inner) binary (A_a and A_b, with its center of mass denoted by A) orbited by an external body B. It is also possible to have a star A orbited by an (outer) binary (B_a and B_b). It is worth mentioning that the dynamical interaction between the inner and outer systems constantly change both orbits, the inclination and eccentricity are free to evolve in time [Steves et al., 2010] and, under certain conditions, the argument of the pericenter of the orbit oscillates around a constant value, which leads to a periodic exchange between its eccentricity (e) and its inclination (i), known as Kozai-Lidov cycles [Naoz, 2016]. However, the time scale of that evolution is much longer than the time span of the observations, so the orbital parameters can be considered constant [Tokovinin and Latham, 2017], which is one of the basic assumptions adopted in this paper.

A.1 General Dynamics

Inner system: First of all, the bodies A_a and A_b keep Newton motion laws:

$$\ddot{\vec{r}}_{A_a} = \frac{Gm_{A_b}}{r^2_{A_aA_b}}\hat{r}_{A_aA_b} \tag{A.1}$$

$$\ddot{\vec{r}}_{A_b} = \frac{Gm_{A_a}}{r^2_{A_aA_b}}\hat{r}_{A_aA_b} \tag{A.2}$$

Subtracting Equations A.1 and A.2:

$$\ddot{\vec{r}}_{A_aA_b} = \ddot{\vec{r}}_{A_b} - \ddot{\vec{r}}_{A_a} = -\frac{G}{r^2_{A_aA_b}}(m_{A_a} + m_{A_b})\hat{r}_{A_aA_b} \tag{A.3}$$

This represents the movement of the secondary around the primary, which is what the astrometric measurement portrays.

Outer system: Repeating the procedure for the outer system, the movement of B around (the center of mass of) A (considered as a single object) is represented by:

$$\ddot{\vec{r}}_{AB} = -\frac{G}{r_{AB}^2}(m_A + m_B)\hat{r}_{AB} \tag{A.4}$$

However, as for the inner system, it is a matter of interest to obtain the movement of B around the primary A_a, given that this is what what is usually measured in differential astrometry of visual binaries (see Figure 4.1 for details). Therefore, considering Equation (A.5), the relationship between A_a and the center of mass A Equation (A.6), and the position of the center of mass A Equation (A.7), the vector $\vec{r}_{A_aA}$ can be rewritten as shown in Equation (A.8).

$$\vec{r}_{A_aB} = \vec{r}_{A_aA} + \vec{r}_{AB} \tag{A.5}$$

$$\vec{r}_{A_aA} = \vec{r}_A - \vec{r}_{A_a} \tag{A.6}$$

$$\vec{r}_A = \frac{m_{A_a}\vec{r}_{A_a} + m_{A_b}\vec{r}_{A_b}}{m_{A_a} + m_{A_b}} \tag{A.7}$$

$$\vec{r}_{A_aA} = \frac{m_{A_a}\vec{r}_{A_a} + m_{A_b}\vec{r}_{A_b}}{m_{A_a} + m_{A_b}} - \vec{r}_{A_a} = (\vec{r}_{A_b} - \vec{r}_{A_a})\frac{m_{A_b}}{m_{A_a} + m_{A_b}} \tag{A.8}$$

Defining the **mass ratio** $q_{A_aA_b} = \frac{m_{A_b}}{m_{A_a}}$ and using that $\vec{r}_{A_aA_b} = \vec{r}_{A_b} - \vec{r}_{A_a}$, Equation (A.8) can be rewritten as:

$$\vec{r}_{A_aA} = \vec{r}_{A_aA_b}\left(\frac{q_{A_aA_b}}{1 + q_{A_aA_b}}\right) \tag{A.9}$$

Thereby, Equation (A.5) can be rewritten as:

$$\vec{r}_{A_aB} = \vec{r}_{AB} + \vec{r}_{A_aA_b}\left(\frac{q_{A_aA_b}}{1 + q_{A_aA_b}}\right) \tag{A.10}$$

Defining the **wobble factor** [Tokovinin and Latham, 2017] or the **fractional mass** [Heintz, 1978] $f_{A_aA_b} = \frac{q_{A_aA_b}}{1+q_{A_aA_b}}$, it is finally obtained that the movement of B with respect to the primary A_a is given by:

$$\vec{r}_{A_aB} = \vec{r}_{AB} + f_{A_aA_b} \cdot \vec{r}_{A_aA_b} \tag{A.11}$$

It is important to note that Equation (A.11) clearly shows that if $\vec{r}_{A_aA_b}$ and $\vec{r}_{AB}$ satisfy Kepler's equations, the combined orbit $\vec{r}_{A_aB}$ is not Keplerian and that, in particular, it is not a closed orbit. On the other hand, when the tight binary corresponds to B and it is formed by B_a, B_b, it is easy to show that a negative wobble factor is obtained [Lane et al., 2014, Tokovinin, 2018a, Tokovinin and Latham, 2017]:

$$\vec{r}_{B_aA} = \vec{r}_{BA} - f_{B_aB_b} \cdot \vec{r}_{B_aB_b} \tag{A.12}$$

A.2 True Anomaly

Having obtained the orbital parameters, the value of the relative position between stars can be computed at any given epoch τ. Let be T the **time of periastron passage**, the epoch when the separation between primary and its companion reaches its minimum value. Then, Kepler's equation can be written:

$$M = \frac{2\pi(\tau - T)}{P} = E - e\sin(E) \tag{A.13}$$

Where the terms M and E are the **mean anomaly** and **eccentric anomaly**, respectively. As Equation (A.13) has no analytic solution, it must be solved using numerical methods. Once E is obtained, the **true anomaly** ν can be computed through:

$$\tan\left(\frac{\nu}{2}\right) = \sqrt{\frac{1+e}{1-e}}\tan\left(\frac{E}{2}\right) \tag{A.14}$$

The **true anomaly** corresponds to the angle between the main focus of the ellipse and the companion star, provided that the periastron is aligned with the X axis, and the primary star occupies the main focus of the ellipse. As the quadrants for E and ν are the same, i.e., Equation (A.14) allows to compute ν without any ambiguity.

A.3 Cartesian Coordinates

Inner System Regarding the inner system, the movement of the secondary A_b around the primary A_a can be described in Cartesian coordinates by:

$$\begin{bmatrix} X_{A_aA_b}(\tau) \\ Y_{A_aA_b}(\tau) \end{bmatrix} = \begin{bmatrix} r_{A_aA_b}(\tau)\cdot\cos(\nu_{A_aA_b}(\tau)) \\ r_{A_aA_b}(\tau)\cdot\sin(\nu_{A_aA_b}(\tau)) \end{bmatrix} \tag{A.15}$$

Then, to project the orbit in the plane of the sky, the Thiele-Innes constants are used

$$\{A_{A_aA_b}, B_{A_aA_b}, F_{A_aA_b}, G_{A_aA_b}\}$$

which are a function of the orbital parameters

$$\{a_{A_aA_b}, \omega_{A_aA_b}, \Omega_{A_aA_b}, i_{A_aA_b}\}$$

$$\begin{bmatrix} x_{A_aA_b}(\tau) \\ y_{A_aA_b}(\tau) \end{bmatrix} = \begin{bmatrix} A_{A_aA_b}X_{A_aA_b}(\tau) + F_{A_aA_b}Y_{A_aA_b}(\tau) \\ B_{A_aA_b}X_{A_aA_b}(\tau) + G_{A_aA_b}Y_{A_aA_b}(\tau) \end{bmatrix} \tag{A.16}$$

Outer System On the other hand, concerning the outer system, the movement of the secondary B around the *fictional* primary A can be described in Cartesian coordinates by:

$$\begin{bmatrix} X_{AB}(\tau) \\ Y_{AB}(\tau) \end{bmatrix} = \begin{bmatrix} r_{AB}(\tau)\cdot\cos(\nu_{AB}(\tau)) \\ r_{AB}(\tau)\cdot\sin(\nu_{AB}(\tau)) \end{bmatrix} \tag{A.17}$$

Then, to project the orbit in the plane of the sky, the Thiele-Innes constants $\{A_{AB}, B_{AB}, F_{AB}, G_{AB}\}$ are used, which are a function of the orbital parameters $\{a_{AB}, \omega_{AB}, \Omega_{AB}, i_{AB}\}$:

$$\begin{bmatrix} x_{AB}(\tau) \\ y_{AB}(\tau) \end{bmatrix} = \begin{bmatrix} A_{AB}X_{AB}(\tau) + F_{AB}Y_{AB}(\tau) \\ B_{AB}X_{AB}(\tau) + G_{AB}Y_{AB}(\tau) \end{bmatrix} \tag{A.18}$$

Finally, Equation (A.11) is rewritten in Cartesian coordinates:

$$\begin{aligned} \vec{r}_{A_aB} &= \vec{r}_{AB} + f_{A_aA_b} \cdot \vec{r}_{A_aA_b} \\ \begin{bmatrix} X_{A_aB}(\tau) \\ Y_{A_aB}(\tau) \end{bmatrix} &= \begin{bmatrix} X_{AB}(\tau) \\ Y_{AB}(\tau) \end{bmatrix} + f_{A_aA_b} \cdot \begin{bmatrix} X_{A_aA_b}(\tau) \\ Y_{A_aA_b}(\tau) \end{bmatrix} \end{aligned} \tag{A.19}$$

As the projection in the plane of the sky can be represented as consecutive rotation matrices, it is a linear transformation. So, the projection of a weighted sum is the weighted sum of the projections. Then, the Cartesian coordinates of the position of B with respect to A_a can be written as:

$$\begin{bmatrix} x_{A_aB}(\tau) \\ y_{A_aB}(\tau) \end{bmatrix} = \begin{bmatrix} x_{AB}(\tau) \\ y_{AB}(\tau) \end{bmatrix} + f_{A_aA_b} \cdot \begin{bmatrix} x_{A_aA_b}(\tau) \\ y_{A_aA_b}(\tau) \end{bmatrix} \tag{A.20}$$

The procedure is analogous when the binary corresponds to B, and it is formed by B_a, B_b.

$$\begin{bmatrix} x_{B_aA}(\tau) \\ y_{B_aA}(\tau) \end{bmatrix} = \begin{bmatrix} x_{AB}(\tau) \\ y_{AB}(\tau) \end{bmatrix} - f_{B_aB_b} \cdot \begin{bmatrix} x_{B_aB_b}(\tau) \\ y_{B_aB_b}(\tau) \end{bmatrix} \tag{A.21}$$

A.4 Radial Velocity

Conversely, it can be noted that the velocity of A_b with respect to the center of mass A can be written as:

$$u_{A,A_b}(\tau) = u_{A_b}(\tau) - u_A(\tau) \tag{A.22}$$

Considering that A and B are moving in an elliptic orbit around the center of mass of the system, denoted by **cm**. So, the velocity of A can be described as:

$$u_A(\tau) = v_{cm} + K_3 \left(e_{AB}\cos(\omega_{AB}) + \cos(\omega_{AB} + \nu_{AB}(\tau))\right) \tag{A.23}$$

$$u_B(\tau) = v_{cm} - K_4 \left(e_{AB}\cos(\omega_{AB}) + \cos(\omega_{AB} + \nu_{AB}(\tau))\right) \tag{A.24}$$

If the mass ratio of the outer system $q_{AB} = \frac{m_B}{m_A}$ is introduced, Equations A.23 and A.24 can be rewritten as:

$$u_A(\tau) = v_{cm} + q_{AB} \cdot K_4 \left(e_{AB}\cos(\omega_{AB}) + \cos(\omega_{AB} + \nu_{AB}(\tau))\right) \tag{A.25}$$

$$u_B(\tau) = v_{cm} - K_4 \left(e_{AB}\cos(\omega_{AB}) + \cos(\omega_{AB} + \nu_{AB}(\tau))\right) \tag{A.26}$$

Besides, A_a and A_b are also moving in an elliptic orbit around A, and therefore:

$$u_{A_a}(\tau) = u_A + K_1 \left(e_{A_a A_b} \cos(\omega_{A_a A_b}) + \cos(\omega_{A_a A_b} + \nu_{A_a A_b}(\tau)\right) \quad (A.27)$$
$$u_{A_b}(\tau) = u_A - K_2 \left(e_{A_a A_b} \cos(\omega_{A_a A_b}) + \cos(\omega_{A_a A_b} + \nu_{A_a A_b}(\tau)\right) \quad (A.28)$$

If the mass ratio $q_{A_a A_b}$ is introduced, Equations A.27 and A.28 can be rewritten as:

$$u_{A_a}(\tau) = u_A + K_1 \left(e_{A_a A_b} \cos(\omega_{A_a A_b}) + \cos(\omega_{A_a A_b} + \nu_{A_a A_b}(\tau)\right) \quad (A.29)$$
$$u_{A_b}(\tau) = u_A - \frac{K_1}{q_{A_a A_b}} \left(e_{A_a A_b} \cos(\omega_{A_a A_b}) + \cos(\omega_{A_a A_b} + \nu_{A_a A_b}(\tau)\right) \quad (A.30)$$

The above Equations are expressed in terms of K_1 and K_4, because it is more likely to obtain measurements from the inner primary A_a than from the inner secondary A_b. In addition, the object A is fictional, so only B's RV measurements could be available.

A.5 Other Relevant Quantities

There are some quantities that are relevant to compute after the orbital parameters have been estimated, namely, the **stellar masses** and the **mutual inclination** of the system. In the following subsections explicit expressions in terms of the orbital parameters are derived.

A.5.1 Stellar Masses

To obtain the stellar masses for each one of the three bodies involved, the relationships for binary systems are used, given the hierarchical approximation. For the inner system, given that:

$$a_{A_a A_b} = a_{A_a} + a_{A_b} \quad (A.31)$$
$$q_{A_a A_b} = \frac{a_{A_a}}{a_{A_b}} = \frac{m_{A_b}}{m_{A_a}} \quad (A.32)$$
$$m_{A_a} + m_{A_b} = \frac{1}{\bar{\omega}^3} \cdot \frac{a^3_{A_a A_b}}{P^2_{A_a A_b}} \quad (A.33)$$

Then, the individual masses correspond to:

$$m_{A_a} = \frac{a^3_{A_a A_b}}{\bar{\omega}^3 \cdot P^2_{A_a A_b}} \cdot \frac{1}{(1 + q_{A_a A_b})} \quad (A.34)$$
$$m_{A_b} = \frac{a^3_{A_a A_b}}{\bar{\omega}^3 \cdot P^2_{A_a A_b}} \cdot \frac{q_{A_a A_b}}{(1 + q_{A_a A_b})} \quad (A.35)$$

The procedure is analogous for the outer system.

A.5.2 Mutual Inclination

Given the parameters $i_{A_aA_b}$, i_{AB}, $\Omega_{A_aA_b}$ and Ω_{AB}, the mutual inclination Φ can be obtained as [Lane et al., 2014, Muterspaugh et al., 2010]:

$$\cos(\Phi) = \cos(i_{A_aA_b}) \cdot \cos(i_{AB}) + \sin i_{A_aA_b} \cdot \sin i_{AB} \cdot \cos(\Omega_{AB} - \Omega_{A_aA_b}) \tag{A.36}$$

Appendix B

Dimensionality Reduction

B.1 Preliminaries

Here, dimensionality reduction in the *radial-velocity-alone* scenario is proposed, to reduce from 20-dimensional state space to a 15-dimensional space. Even though it is just 25% reduction, in sample-based schemes like MCMC any decrease in the computational cost is appreciated.

The reduction consists on the separation of the parameter vector into two lower dimension vectors: one containing non-linear components (θ_{NL}) and the other components that are linearly dependent with respect to θ_{NL} and can be obtained through a weighted least-squares procedure (θ_L). The method is inspired by [Wright and Howard, 2009] where they reformulate the radial velocity equations in such a way that they get linear in some parameters, allowing for an analytic calculation of weighted least-square solutions.

Therefore, the search of the state space is focused on

$$\theta_{NL} = [T_{A_aA_b}, P_{A_aA_b}, e_{A_aA_b}, q_{A_aA_b}, i_{A_aA_b}, T_{AB}, P_{AB}, e_{AB}, q_{AB}, i_{AB}]$$

and then the vector of parameters $\theta_L = [a_{A_aA_b}, \omega_{A_aA_b}, a_{AB}, \omega_{AB}, v_{cm}]$ is obtained.

B.2 Method

Following the procedure described in Appendix A, the radial velocity equations can be formulated as B.1. This modelling does not consider linear trends $d(t - t_0)$, which are used to account for unmodeled noise sources and to notice the presence of massive objects in wide orbits around the star [Retired, 2010, Wright and Howard, 2009].

$$u_A(\tau) = v_{cm} + q_{AB} \cdot K_4 \left(e_{AB}\cos(\omega_{AB}) + \cos(\omega_{AB} + \nu_{AB}(\tau)\right) \quad \text{(B.1)}$$

$$u_B(\tau) = v_{cm} - K_4 \left(e_{AB}\cos(\omega_{AB}) + \cos(\omega_{AB} + \nu_{AB}(\tau)\right) \quad \text{(B.2)}$$

$$u_{A_a}(\tau) = u_A + K_1 \left(e_{A_aA_b}\cos(\omega_{A_aA_b}) + \cos(\omega_{A_aA_b} + \nu_{A_aA_b}(\tau)\right) \quad \text{(B.3)}$$

$$u_{A_b}(\tau) = u_A - \frac{K_1}{q_{A_aA_b}} \left(e_{A_aA_b}\cos(\omega_{A_aA_b}) + \cos(\omega_{A_aA_b} + \nu_{A_aA_b}(\tau)\right) \quad \text{(B.4)}$$

However, the amplitudes K_1 and K_4 can be calculated as a function of the parameters $P_{A_aA_b}$, $e_{A_aA_b}$, $a_{A_aA_b}$, $i_{A_aA_b}$, P_{AB}, e_{AB}, a_{AB} and i_{AB} as follows:

$$K_1 = \frac{2\pi \sin(i_{A_aA_b})}{P_{A_aA_b}\sqrt{(1-e^2_{A_aA_b})}} \cdot \frac{a''_{A_aA_b}}{\bar{\omega}} \cdot \frac{q_{A_aA_b}}{1+q_{A_aA_b}} \cdot \lambda \quad \text{(B.5)}$$

$$K_4 = \frac{2\pi \sin(i_{AB})}{P_{AB}\sqrt{(1-e^2_{AB})}} \cdot \frac{a''_{AB}}{\bar{\omega}} \cdot \frac{1}{1+q_{AB}} \cdot \lambda \quad \text{(B.6)}$$

Where λ is a constant to convert from $\frac{arcsec}{yr}$ to $\frac{km}{s}$. Then, if the variables K'_1, K'_4, m_1 and m_2 from Equations (B.7) to (B.10) are considered, the radial velocity equations can be reformulated as shown in Equations (B.11) to (B.14).

$$K'_1 = \frac{\sin(i_{A_aA_b})}{P_{A_aA_b}\sqrt{(1-e^2_{A_aA_b})}} \cdot \frac{q_{A_aA_b}}{1+q_{A_aA_b}} \quad \text{(B.7)}$$

$$K'_4 = \frac{\sin(i_{AB})}{P_{AB}\sqrt{(1-e^2_{AB})}} \cdot \frac{1}{1+q_{AB}} \quad \text{(B.8)}$$

$$m_1 = 2\pi \cdot \frac{a''_{A_aA_b}}{\bar{\omega}} \cdot \lambda \quad \text{(B.9)}$$

$$m_2 = 2\pi \cdot \frac{a''_{AB}}{\bar{\omega}} \cdot \lambda \quad \text{(B.10)}$$

$$u_A(\tau) = v_{cm} + q_{AB} \cdot K'_4 \cdot m_2 \cdot (e_{AB}\cos(\omega_{AB}) + \cos(\omega_{AB} + \nu_{AB}(\tau)) \quad \text{(B.11)}$$

$$u_B(\tau) = v_{cm} - K'_4 \cdot m_2 \cdot (e_{AB}\cos(\omega_{AB}) + \cos(\omega_{AB} + \nu_{AB}(\tau)) \quad \text{(B.12)}$$

$$u_{A_a}(\tau) = u_A + K'_1 \cdot m_1 \cdot (e_{A_aA_b}\cos(\omega_{A_aA_b}) + \cos(\omega_{A_aA_b} + \nu_{A_aA_b}(\tau)) \quad \text{(B.13)}$$

$$u_{A_b}(\tau) = u_A - \frac{K'_1}{q_{A_aA_b}} \cdot m_1 \cdot (e_{A_aA_b}\cos(\omega_{A_aA_b}) + \cos(\omega_{A_aA_b} + \nu_{A_aA_b}(\tau)) \quad \text{(B.14)}$$

Finally, considering the auxiliary variables α_1, β_1, α_2 and β_2:

$$\alpha_1 = m_1 \cdot \cos(\omega_{A_aA_b}) \quad \text{(B.15)}$$

$$\beta_1 = -m_1 \cdot \sin(\omega_{A_aA_b}) \quad \text{(B.16)}$$

$$\alpha_2 = m_2 \cdot \cos(\omega_{AB}) \quad \text{(B.17)}$$

$$\beta_2 = -m_2 \cdot \sin(\omega_{AB}) \quad \text{(B.18)}$$

the following equations are obtained:

$$\vec{u}_A(\tau) = \vec{v}_{cm} + q_{AB} \cdot K'_4 \cdot \alpha_2 \cdot (e_{AB} + \cos(\nu_{AB})) + q_{AB} \cdot K'_4 \cdot \alpha_2 \cdot \sin(\nu_{AB}(\tau)) \quad \text{(B.19)}$$

$$\vec{u}_B(\tau) = \vec{v}_{cm} - K'_4 \cdot \alpha_2 \cdot (e_{AB} + \cos(\nu_{AB})) - K'_4 \cdot \beta_2 \cdot \cos(\nu_{AB}(\tau)) \quad \text{(B.20)}$$

$$\vec{u}_{A_a}(\tau) = \vec{u}_A + K'_1 \cdot \alpha_1 \cdot (e_{A_aA_b} + \cos(\nu_{A_aA_b})) + K'_1 \cdot \beta_1 \cdot \sin(\nu_{A_aA_b}(\tau)) \quad \text{(B.21)}$$

$$\vec{u}_{A_b}(\tau) = \vec{u}_A - \frac{K'_1}{q_{A_aA_b}} \cdot \alpha_1 \cdot (e_{A_aA_b} + \cos(\nu_{A_aA_b})) - \frac{K'_1}{q_{A_aA_b}} \cdot \beta_1 \cdot \sin(\nu_{A_aA_b}(\tau)) \quad \text{(B.22)}$$

The radial velocity equations depicted above can be represented in a matrix form if the parameter's vector $\vec{\theta} = [\alpha_1, \beta_1, \alpha_2, \beta_2, v_{cm}]^T$ is defined:

$$u_{A_a}(\tau) = \vec{\theta}^T \cdot \begin{bmatrix} K_1' \cdot (\cos(\nu_{A_a A_b}(\tau)) + e_{A_a A_b}) \\ K_1' \cdot \sin(\nu_{A_a A_b}(\tau)) \\ q_{AB} \cdot K_4' \cdot (\cos(\nu_{AB}(\tau)) + e_{AB}) \\ q_{AB} \cdot K_4' \cdot \sin(\nu_{AB}(\tau)) \\ 1 \end{bmatrix} = \vec{\theta}^T \cdot F_{A_a}(\tau) \tag{B.23}$$

$$u_{A_b}(\tau) = \vec{\theta}^T \cdot \begin{bmatrix} -\frac{K_1'}{q_{A_a A_b}} \cdot (\cos(\nu_{A_a A_b}(\tau)) + e_{A_a A_b}) \\ -\frac{K_1'}{q_{A_a A_b}} \cdot \sin(\nu_{A_a A_b}(\tau)) \\ q_{AB} \cdot K_4' \cdot (\cos(\nu_{AB}(\tau)) + e_{AB}) \\ q_{AB} \cdot K_4' \cdot \sin(\nu_{AB}(\tau)) \\ 1 \end{bmatrix} = \vec{\theta}^T \cdot F_{A_a}(\tau) \tag{B.24}$$

$$u_B(\tau) = \vec{\theta}^T \cdot \begin{bmatrix} 0 \\ 0 \\ -K_4' \cdot (\cos(\nu_{AB}(\tau)) + e_{AB}) \\ -K_4' \cdot \sin(\nu_{AB}(\tau)) \\ 1 \end{bmatrix} = \vec{\theta}^T \cdot F_B(\tau) \tag{B.25}$$

Then, considering a matrix F with the matrices F_{A_a}, F_{A_b} and F_B for all the epochs of measurement for the three bodies involved

$$F = [F_{A_a}(\tau_{r_{a(0)}}) \dots F_{A_a}(\tau_{r_{a(N_a)}}) | F_{A_b}(\tau_{r_{b(0)}}) \dots F_{A_b}(\tau_{r_{b(N_b)}}) | F_B(\tau_{r_{B(0)}}) \dots F_B(\tau_{r_{B(N_B)}})]$$

and a vector with the modelled values for all the epochs of measurement for the three bodies involved $\vec{u} = [u_{A_a}(\tau_{r_{a(0)}}) \dots u_{A_a}(\tau_{r_{a(N_a)}}), u_{A_b}(\tau_{r_{b(0)}}) \dots u_{A_b}(\tau_{r_{b(N_b)}}), u_B(\tau_{r_{B(0)}}) \dots u_B(\tau_{r_{B(N_B)}})]$, it can be written in compact form that $\vec{u} = \vec{\theta}^T \cdot F$.

Afterwards, the vector of parameters $\vec{\theta}$ can be estimated from the data directly using least-squares and the figure of merit χ^2 (Equation (B.26)). If the matrix W is defined as the diagonal with the weights associated to each observation $W_{kl} = \frac{\delta_{kl}}{\sigma_k^2}$, $\vec{v}$ a vector with all the observations from the three bodies concatenated and using weighted least-squares:

$$\chi^2 = \sum_{k=1}^{N_a} \frac{(v_k - u_{A_a}(\tau_{r_{a(k)}}))^2}{\sigma_{A_a}(\tau_{r_{a(k)}})^2} + \sum_{k=1}^{N_b} \frac{(v_k - u_{A_b}(\tau_{r_{b(k)}})^2}{\sigma_{A_b}(\tau_{r_{b(k)}})^2} + \sum_{k=1}^{N_B} \frac{(v_k - u_B(\tau_{r_{B(k)}})^2}{\sigma_B(\tau_{r_{B(k)}})^2} \tag{B.26}$$

$$\frac{\partial \chi^2}{\partial \vec{\theta}} = -2(\vec{v} - \vec{\theta}^T F) W F^T = \vec{0} \tag{B.27}$$

And the parameters' vector is obtained through $\vec{\theta} = \vec{v} W F^T (F W F^T)^{-1}$.

Finally, the original parameters can be recovered:

$$m_1 = \sqrt{\alpha_1^2 + \beta_1^2} \tag{B.28}$$

$$m_2 = \sqrt{\alpha_2^2 + \beta_2^2} \tag{B.29}$$

$$\omega_{A_a A_b} = \arctan\left(\frac{-\beta_1}{\alpha_1}\right) \tag{B.30}$$

$$\omega_{AB} = \arctan\left(\frac{-\beta_2}{\alpha_2}\right) \tag{B.31}$$

$$a''_{A_a A_b} = \frac{m_1 \cdot \bar{\omega}}{2\pi\lambda} \tag{B.32}$$

$$a''_{AB} = \frac{m_2 \cdot \bar{\omega}}{2\pi\lambda} \tag{B.33}$$

Appendix C

Computation of Predictive Models

All of the predictive distributions were numerically approximated using MCMC algorithms. Those sampling-schemes construct a Markov chain with Θ as state space and $\pi(\theta) = p(\theta|z)$ as the stationary distribution. They generate a sequence of parameter values $\theta_1, \theta_2 \cdots \theta_n$, by sampling from a proposal distribution an then accepting/rejecting the sample according to a criteria that depends on prior information and the likelihood. That empirical distribution approaches the target distribution in the limit of long runs.

Following Bayes rule, π can be written as:

$$\pi(\theta) = p(\theta|z) = \frac{p(z|\theta)p(\theta)}{\int p(z|\theta)p(\theta)d\theta} \tag{C.1}$$

However, it is not necessary to compute de denominator, as MCMC algorithms base the acceptance/rejection of a sample θ' based on the posterior ratio:

$$\frac{\pi(\theta')}{\pi(\theta')} = \frac{p(z|\theta')p(\theta')}{p(z|\theta)p(\theta)} \tag{C.2}$$

Due to the highly dimensional state space, an hybrid Gibbs sampler MCMC variant is used, which allows us to draw iteratively samples from the conditional posterior distribution for each variable given the remaining ones using an MH iteration. The state space is explored using a random walk with Gaussian proposal distributions. Besides, there is a constrained state space due to physical restrictions on the parameters, thus it is assumed to be circular for each dimension, then several iterations on the boundaries are avoided.

On the other hand, given that the observation model includes Gaussian additive noise (see Equation (4.1)), all likelihoods are proportional to $\exp(-\frac{1}{2}\chi^2)$:

$$\mathcal{L}(\theta) = p(z|\theta) \propto \exp\left(-\frac{1}{2}\sum_{k=1}^{N}\frac{\|z_k - f(\theta, \tau_k)\|^2}{\sigma_k^2}\right) \tag{C.3}$$

Finally, it is worth mentioning that, due to T can be replaced by $T \pm nP$ (n could be any

integer), $T \in (0, P)$ is chosen. Then, for simplicity the variable $T' = \frac{T}{P} \in (0, 1)$ is defined and it is used along the estimations. After finishing, the old variable T is obtained.

Here below, the details of each one of the scenarios presented in Section 4.2 are explained.

C.1 Astrometry alone

As seen in Figure 4.5, two processes are run consequently. First, the astrometric observations from the inner system $\{\vec{z_1}\}_{k=1}^{N_1}$ are used to estimate the parameters set

$$\Theta_1 = \{T_{A_aA_b}, P_{A_aA_b}, e_{A_aA_b}, a_{A_aA_b}, \omega_{A_aA_b}, \Omega_{A_aA_b}, i_{A_aA_b}\}$$

However, a dimensional reduction is performed. For each iteration of the algorithm, the parameters

$$\theta_{NL} = \{T_{A_aA_b}, P_{A_aA_b}, e_{A_aA_b}\}$$

are left free and the parameters $\theta_L = \{a_{A_aA_b}, \omega_{A_aA_b}, \Omega_{A_aA_b}, i_{A_aA_b}\}$ are obtained using a weighted least-squares procedure. The process is explained in detail in [Mendez et al., 2017].

Then, the astrometric observations from the outer system to estimate the parameters set $\Theta_2 = \{q_{A_aA_b}, T_{AB}, P_{AB}, e_{AB}, a_{AB}, \omega_{AB}, \Omega_{AB}, i_{AB}\}$ are used, employing an imputations framework within MCMC (see [Claveria et al., 2019] for more details). Here, for each iteration of the algorithm:

1. Sample Θ_2 using the proposal distribution, and obtain

$$\{\vec{y_2} = f_1(T_{AB}, P_{AB}, e_{AB}, a_{AB}, \omega_{AB}, \Omega_{AB}, i_{AB}, \tau_k)\}_{k=1}^{N_2}$$

2. Sample from the distribution $P_{\Theta_1|\vec{Z_1}}(\cdot|\vec{z_1})$, obtained in the last step, and generate $\{\vec{y_1} = \frac{q_{A_aA_b}}{1+q_{A_aA_b}} \cdot f_1(T_{A_aA_b}, P_{A_aA_b}, e_{A_aA_b}, a_{A_aA_b}, \omega_{A_aA_b}, \Omega_{A_aA_b}, i_{A_aA_b}, \tau_k)\}_{k=1}^{N_2}$.

3. Compute $\vec{z} = \vec{y_2} + \vec{y_1}$ and use it as observation to continue with the algorithm.

Later, we add a physical restrictions step. Besides checking the support for each parameter, it is necessary to check:

- The hierarchical approximation in periods $P_{A_aA_b} < P_{AB}$ and semi-major axes $a_{A_aA_b} < a_{AB}$.
- The sum of masses has sense: $m_{A_a} + m_{A_b} < m_A + m_B \Leftrightarrow \frac{a^3_{A_aA_b}}{P^2_{A_aA_b}} < \frac{a^3_{AB}}{P^2_{AB}}$.

This process is shown in detail in Algorithm 4.

C.2 Radial Velocities alone

Unlike the last scenario, all observations are used at once, achieving the estimation of the parameter set

$$\begin{aligned}\Theta_{RV} &= \{T_{A_aA_b}, P_{A_aA_b}, e_{A_aA_b}, a_{A_aA_b}, \omega_{A_aA_b}, \\ &\quad i_{A_aA_b}, q_{A_aA_b}, T_{AB}, P_{AB}, e_{AB}, a_{AB}, \omega_{AB}, i_{AB}, q_{AB}, v_{cm}\}\end{aligned}$$

in just one process, as seen in Figure 4.8. Nonetheless, since the dimension of the parameter space rises to fifteen, a dimensionality reduction is proposed.

For each iteration of the algorithm, the parameters

$$\theta_{NL} = \{T_{A_aA_b}, P_{A_aA_b}, q_{A_aA_b}, i_{A_aA_b}, T_{AB}, P_{AB}, e_{AB}, q_{AB}, i_{AB}\}$$

are left free and the parameters $\theta_L = \{a_{A_aA_b}, \omega_{A_aA_b}, a_{AB}, \omega_{AB}, v_{cm}\}$ are obtained using a weighted least-squares procedure, similar to the processes described in [Mendez et al., 2017, Wright and Howard, 2009], and explained in detail in Appendix B. This allows us to sample from a reduced parameter space θ_{NL}, while linearly deriving the rest of the parameters θ_L.

Later, we add a physical restrictions step. Besides checking the support for each parameter, it is necessary to check:

- The hierarchical approximation in periods $P_{A_aA_b} < P_{AB}$ and semi-major axes $a_{A_aA_b} < a_{AB}$.
- The sum of masses has sense: $m_A = m_{A_a} + m_{A_b} \Leftrightarrow |\frac{a_{AB}^3}{P_{AB}^2} \cdot \frac{1}{(1+q_{AB})} - \frac{a_{A_aA_b}^3}{P_{A_aA_b}^2}| < \varepsilon$.

The algorithm can be seen in detail in Algorithm 6.

Appendix D

Algorithms for parameter estimation

At last, here we show the pseudocode of the MCMC algorithms mentioned in Appendix C:

- Algorithm (4) shows the MCMC-based method to perform the parameter estimation of parameters $\{T_{A_aA_b}, P_{A_aA_b}, e_{A_aA_b}, a_{A_aA_b}, \omega_{A_aA_b}, \Omega_{A_aA_b}, i_{A_aA_b}\}$.
- Algorithm (5) shows the MCMC and imputations-based framework to perform the parameter estimation of parameters $\{q_{A_aA_b}, T_{AB}, P_{AB}, e_{AB}, a_{AB}, \omega_{AB}, \Omega_{AB}, i_{AB}\}$.
- Algorithm (6) shows the MCMC-based method to perform the parameter estimation of parameters

$$\{T_{A_aA_b}, P_{A_aA_b}, e_{A_aA_b}, a_{A_aA_b}, \omega_{A_aA_b}, i_{A_aA_b}, q_{A_aA_b}, T_{AB}, P_{AB}, e_{AB}, a_{AB}, \omega_{AB}, i_{AB}\, q_{AB}, v_{cm}\}$$

Algorithm 4 MCMC - Parameter estimation for inner parameters (astrometric alone scenario).

$\theta_{NL} = \{T_{A_aA_b}, P_{A_aA_b}, e_{A_aA_b}\}$
$\theta_L = \{a_{A_aA_b}, \omega_{A_aA_b}, \Omega_{A_aA_b}, i_{A_aA_b}\}$
Initialize $\theta^{(0)}$ sampling from priors
for $k = 1 \dots N_{steps}$ **do**
  $\theta' = \theta^{(k)}$
  for $j = 1 \dots 3$ (All the non linear parameters) **do**
    Sample $\theta'_j \sim \mathcal{N}(\theta'_j, \sigma_j^2)$ (Apply additive Gaussian perturbation on component j)
    Compute θ_L
    Compute $\mathcal{L}(\theta')$
    Sample $u' \sim \mathcal{U}(0, 1)$
    Compute acceptance ratio $\frac{\mathcal{L}(\theta')}{\mathcal{L}(\theta^{(k)})}$
    if $u <$ ratio **then**
      $\theta_j^{(k+1)} = \theta'_j$
    else
      $\theta_j^{(k+1)} = \theta_j^{(k)}$
    end if
  end for
end for

Algorithm 5 MCMC and Imputations - Parameter estimation for outer parameters (astrometric alone scenario).

$\theta_1 = \{T_{A_aA_b}, P_{A_aA_b}, e_{A_aA_b}, a_{A_aA_b}, \omega_{A_aA_b}, \Omega_{A_aA_b}, i_{A_aA_b}\}$
$\theta_2 = \{q_{A_aA_b}, T_{AB}, P_{AB}, e_{AB}, a_{AB}, \omega_{AB}, \Omega_{AB}, i_{AB}\}$
Initialize $\theta_2^{(0)}$ sampling from priors
for $k = 1 \dots N_{steps}$ **do**
  Generate imputations $\{\vec{y_1}\}_{k=1}^{N_2}$, using $P_{\Theta_1|\vec{Z_1}}(\cdot|\vec{z_1})$.
  $\theta'_2 = \theta_o^{(k)}$
  for $j = 1 \dots 8$ **do**
    Sample $\theta'_{2_j} \sim \mathcal{N}(\theta'_{2_j}, \sigma_j^2)$ (Apply additive Gaussian perturbation on component j)
    Compute physical restrictions
    Compute $\vec{y}_2(\theta'_2, \tau)$ and $\mathcal{L}(\theta'_2)$
    Sample $u' \sim \mathcal{U}(0, 1)$
    Compute acceptance ratio $\frac{\mathcal{L}(\theta')}{\mathcal{L}(\theta^{(k)})}$
    if $u <$ ratio **then**
      $\theta_{2_j}^{(k+1)} = \theta'_{2_j}$
    else
      $\theta_{2_j}^{(k+1)} = \theta_{2_j}^{(k)}$
    end if
  end for
end for

Algorithm 6 MCMC - Parameter estimation in the RV alone scenario.

$\theta_{NL} = \{T_{A_aA_b}, P_{A_aA_b}, e_{A_aA_b}, q_{A_aA_b}, i_{A_aA_b}, T_{AB}, P_{AB}, e_{AB}, q_{AB}, i_{AB}\}$
$\theta_L = \{a_{A_aA_b}, \omega_{A_aA_b}, a_{AB}, \omega_{AB}, v_{cm}\}$
Initialize $\theta^{(0)}$ sampling from priors
for $k = 1 \dots N_{steps}$ **do**
  $\theta' = \theta^{(k)}$
  for $j = 1 \dots 10$ (All the non linear parameters) **do**
    Sample $\theta'_{NL_j} \sim \mathcal{N}(\theta'_{NL_j}, \sigma_j^2)$ (Apply additive Gaussian perturbation on component j)
    Compute θ_L
    Compute physical restrictions
    Compute $\mathcal{L}(\theta')$
    Sample $u' \sim \mathcal{U}(0, 1)$
    Compute acceptance ratio $\frac{\mathcal{L}(\theta')}{\mathcal{L}(\theta^{(k)})}$:
    if $u <$ ratio **then**
      $\theta^{(k+1)} = \theta'_j$ and save θ'
    else
      $\theta^{(k+1)} = \theta^{(k)}$
    end if
  end for
end for