Draft version January 12, 2023 年使用
AASTeX631 默认样式排版

arXiv:2301.04162v1 [astro-ph.SR] 10 Jan 2023

最优　双星与多星系统的观测调度框架 *

Miguel Videla,[1] Rene A. Mendez,[2] Jorge F. Silva,[1] and Marcos E. Orchard[1]

[1] *Department of Electrical Engineering*
Information and Decision Systems Group (IDS)
Facultad de Ciencias Físicas y Matemáticas, Universidad de Chile
Beauchef 850, Santiago, Chile[2] *Department of Astronomy*
Facultad de Ciencias Físicas y Matemáticas, Universidad de Chile
*Casilla 36-D, Santiago, Chileand*European Southern Observatory
*Alonso de Córdova 3107, Vitacura*Santiago, Chile

摘要

对于在人的时间尺度上发生变化的天体物理现象，确定最优观测时刻是一项重要问题，因为这关系到如何有效利用通常十分稀缺的观测设施。本文通过基于最大熵采样原理的贝叶斯框架，探讨紧密目视双星系统的这一问题。我们提出的信息驱动方法利用双星系统的周期结构，高效地估计最优观测时间的概率分布。我们从贝叶斯意义上证明了所提出采样方法的最优性，并通过直接数值实验展示了其有效性。我们成功地将该方法应用于两个目视光谱双星和一个纯天体测量三合层级系统的研究。值得注意的是，我们的方法可应用于任何随时间演化的现象；在专项巡天时代，这一应用尤其引人关注，因为观测节奏的设定可能对实现科学目标产生至关重要的影响。

关键词：双星 (154) — 多星 (1081) — 贝叶斯统计 (1900) — 天体测量双星 (79) — 光谱双星 (1557) — 后验分布 (1926) - 马尔可夫链蒙特卡罗 (1089) – 轨道根数 (1177)

1. 引言

观测天文学领域的技术进步，使天文学界得以发现大量极具研究价值的天体，并促使人们开展众多观测活动，收集相关研究数据。遗憾的是，由于高精度观测设施数量有限，天体和物理现象的数量远远超出了我们的观测能力。这一限制使天文学家为争取宝贵的观测时间而展开激烈竞争，从而限制了可开展的观测次数（Patat &Hussain2012）[1]。因此，尤其对于那些演化时间尺度与人类寿命相当的动力学系统的研究而言，

通讯作者：Miguel Videlamiguel.videla@ug.uchile.cl

*本文部分基于在国际双子座天文台获得的观测数据。该天文台是美国国家科学基金会 NOIRLab 的一个项目，由大学天文研究协会（AURA）根据与美国国家科学基金会签订的合作协议，代表双子座天文台合作伙伴管理。双子座天文台合作伙伴包括：美国国家科学基金会（美国）、国家研究委员会（加拿大）、国家研究与发展署（智利）、科学、技术与创新部（阿根廷）、科学、技术、创新与通信部（巴西），以及韩国天文与空间科学研究院（大韩民国）。本文还部分基于在南方天体物理研究（SOAR）望远镜获得的观测数据。该望远镜是巴西联邦共和国科学、技术与创新部（MCTI）、美国国家光学天文台（NOAO）、北卡罗来纳大学教堂山分校（UNC）和密歇根州立大学（MSU）的联合项目。

[1] 我们团队可使用的一些设施，近期超额申请率的示例可见于https://www.gemini.edu/observing/science-operations-statistics、https://noirlab.edu/science/observing-noirlab/observing-ctio/observing-soar/observing-statistics或http://www.eso.org/sci/observing/phase1/p111/pressure.html。

以确定最合适的观测时刻，从而尽可能充分地刻画所研究的系统。在此背景下，观测调度问题可自然地置于一种现已成熟的理论框架中加以研究（见下文）；该理论正是为实验的最优设计而发展起来的。

本文中，我们将最优实验设计理论应用于双星与多星系统的研究，以确定能够最大程度提供有关这些系统物理参数（例如轨道根数、组分质量）信息的观测规划。具体而言，我们关注的是紧密双星。要恰当地研究它们，需要在大口径望远镜上使用（稀缺的）专用瑞利极限干涉技术（例如，散斑成像）或高角分辨率技术（例如，自适应光学），以分辨出各个组成部分并全面确定其轨道构型。不过，我们强调，我们的方法（详见第2节）可以自然地扩展到任何其他随时间演变的现象。在大型专门巡天时代，这可能是一个特别有趣的应用方向，因为观测节奏的设定是实现科学目标的关键环节（例如系外行星研究，Ford (2008)；Loredo 等人(2012）。

最优实验设计框架旨在根据某种统计准则确定最优实验，从而优化统计模型估计所涉及的成本（Pukelsheim 2006）。确定最优设计的一种常见准则是使估计量的方差最小化，而具体方法必须根据所考虑统计模型的性质进行调整。对于多变量线性统计模型，方差由协方差矩阵表示，因此，最小化估计值方差的问题可转化为根据某些最优性准则，使Fisher信息矩阵（协方差矩阵的逆矩阵）的不变变换最小化（Box 1982）。一些常用的最优性准则包括：使协方差矩阵的迹最小化（A-最优性）；使预先指定的模型参数线性组合的最佳线性无偏估计量的方差最小化（C-最优性）；使信息矩阵的行列式最大化（D-最优性）；使信息矩阵的特征值最大化（E-最优性）等。这些准则统称为字母型设计准则。

字母型设计准则最初仅适用于线性模型，且不考虑有关推断参数的先验信息，但这些限制可以放宽。贝叶斯实验设计方法通过将推断参数的先验分布纳入实验设计框架，克服了后一项限制。在这一新框架下，原问题转化为：根据反映实验目的的某种效用函数，使实验的期望效用最大化（Lindley1972；Chaloner &Verdinelli 1995）。在某些特定情境中，通过选择适当的效用函数，可以得到前述字母型设计准则的一些贝叶斯对应形式；例如，在高斯线性模型中采用香农效用函数，即可推导出贝叶斯 D-最优性准则（Bernardo 1979）。在非线性情形下，确定实验的期望效用通常需要计算复杂的积分，往往还需对问题中的某些要素进行近似，例如，对后验分布采用正态近似（James 1985，第 224 页）、通过局部最优方法将先验分布近似为单点分布（Chernoff 1953），或通过马尔可夫链蒙特卡罗方法估计期望效用（Muller & Parmigiani 1996）。

尽管最优实验设计问题已经在生物学（Balsa-Canto 等人，2008年）、心理学（Myung 与 Pitt，2009年）或材料科学（Dehghannasiri 等人，2017年）等多种领域得到广泛研究和应用，但其在天文学领域中的应用并不十分广泛。在这一领域，最优实验设计的研究主要集中于确定观测某些天体现象的最佳时刻。该领域最早的研究局限于非自适应观测计划（即与先前观测无关），例如规则周期观测计划或随机均匀观测计划（Sozzetti 等人，2002年；Ford，2004年）。后来，人们也探索了基于贝叶斯最优实验设计理论的严格统计方法。Ford (2008）和Loredo 等人 (2012）分别针对系外行星径向速度的搜寻和表征，提出了自适应观测计划方法。这两种方法均基于最大熵采样原理（Shewry 与 Wynn，1987年），旨在通过观测预测不确定性最高的区域，最大化关于所研究天体的信息量。最近，有人提出了一种基于Fisher信息矩阵的不同自适应观测计划方法（Hees 等人，2019年）。尽管后一种方法具有更高的计算效率，但其仅限于高斯问题。

对于双星系统，通常采用的观测准则只是确保轨道覆盖充分；然而，据我们所知，究竟怎样才算充分，尚未以

数学上严谨的方式进行讨论。例如，Lucy (2014) 早期的研究表明，要获得可靠的轨道，相位覆盖必须至少约占轨道的 40%；但该研究并未指出应将观测安排（或进行）在轨道的什么位置（何时）。本文将此前由 Ford (2008) 和 Loredo 等人 (2012) 为系外行星研究提出的、基于最大熵采样准则的最优观测调度方法，应用于双星和层级恒星系统的研究。所提出的方法利用双星系统后验分布的估计结果，生成观测时间的概率分布，以找出能够最大程度增加系统参数信息的观测时间。值得注意的是，这种由信息驱动的方法能够以极低的计算成本，在密集且有意义的时间范围内刻画后验预测信息增益；它利用双星系统的周期性来界定评估时间的范围，并估计最优观测时间的概率分布。该软件可在 GitHub[2] 上免费获取，采用三条款 BSD 许可证。

本文结构安排如下：第2节介绍贝叶斯实验设计问题和最大熵采样准则。第3节介绍轨道统计模型，以及我们提出的双星天体测量观测最优调度方法。第4节将所提出的方法应用于三个不同的双星系统（包括一个层级三星系统）作为测试案例，并对所得结果进行深入分析和讨论。最后，第5 节总结本文的主要结论并展望未来。

2. 最优实验设计与最大熵抽样

本节将贝叶斯实验设计问题形式化为最大熵采样准则，并阐明在某些特定条件下这两个问题的等价性。

2.1. 贝叶斯实验设计

贝叶斯实验设计（Lindley 1972）是一种决策论方法，用于处理实验设计问题；该方法能够纳入待估参数的先验知识，以及观测的不确定性。

令 $\theta \in \Theta$ 为我们希望根据观测随机向量 $Y \in \mathcal{Y}$ 估计的参数向量，并令 $\xi \in \Xi$ 为待选择的实验设计。根据观测 Y，将作出决策 $d \in D$。决策程序包括先选择实验设计 ξ，再选择最终决策 d。

因此，对于任意实验 $\xi \in \Xi$，最佳决策 $d \in D$ 的期望效用为：

$$U(\xi) \equiv \int_Y \max_{d \in D} \int_\Theta p(y|\xi) \cdot p(\theta|y,\xi) \cdot U(d,y,\theta,\xi)\, d\theta\, dy, \tag{1}$$

其中，$p(y|\xi) = \int p(\theta)p(y|\theta,\xi)d\theta$ 表示实验 ξ 在观测空间中的边缘分布；$p(y|\theta,\xi)$ 表示在实验 ξ 中给定 θ 时 Y 的似然；$p(\theta|y,\xi)$ 表示实验 ξ 中 θ 的后验分布；$p(\theta)$ 是参数的先验分布（与 ξ 无关）；$U(d,y,\theta,\xi)$ 表示一般的效用函数。效用函数是一个实值函数，用于描述在观测 $y \in \mathcal{Y}$ 和由 $\theta \in \Theta$ 索引的模型下，作出关于选择实验 $\xi \in \Xi$ 的决策 $d \in D$ 所具有的价值（基于实验目的）。$U(\xi)$ 可视为在估计 Y 中的 θ 这一任务中，选择 ξ 所获得的平均回报。

在此背景下，贝叶斯实验设计可定义为一种旨在确定实验$\xi^* \in \Xi$ 的程序，使其能够最大化最佳决策的期望效用：

$$\xi^* \equiv \arg\max_{\xi \in \Xi} U(\xi). \tag{2}$$

不同设计准则所对应的期望效用函数汇总，可参见 Chaloner &Verdinelli (1995) 的综述。

2.2. 最大信息提取

[2] `BinaryStars` 代码库：https://github.com/mvidela31/BinaryStars。

定义最佳决策的期望效用 $U(\xi)$，即方程（2的标准方法之一，是将其视为实验 $\xi \in \Xi$ 在给定已获取的观测 Θ的情况下，能够提供的关于未知参数向量 Y的信息量。更确切地说，信息的概念可以用香农熵 (Cover1999)来定义。香农熵衡量随机变量所具有的不确定性（先验）或平均信息量。令 X为一个字母表为 $\mathcal{X}$、概率质量函数为 $p(x) = \mathbb{P}(X = x)$ $x \in \mathcal{X}$的离散随机变量，则 X 其熵定义如下：

$$H(X) = \sum_{x \in \mathcal{X}} -p(x) \cdot \log p(x). \tag{3}$$

对于连续随机变量，则改用微分熵的概念（Cover 1999）。设 X 为支撑集位于 $\mathbb{X}$ 且概率密度函数为 $f(x)$ 的连续随机变量，则X 的微分熵定义如下：

$$h(X) = \mathbb{E}_{X \sim f(x)}\{-\log f(X)\} = -\int_{\mathbb{X}} f(x) \cdot \log f(x)\, dx. \tag{4}$$

其中， $\mathbb{E}\{.\}$ 表示该参数的期望。

因此，香农效用 $U_S(\xi)$ 表示从先验分布（$p(\theta)$）$_{\theta \in \Theta}$ 到后验分布（$p(\theta|y,\ \xi)$）$_{(\theta,y) \in \Theta \times \mathcal{Y}}$ (Lindley 1956）所获得的信息增益，或等价地说，不确定性降低（香农意义下）。对于这一目标，香农效用函数 $U_S(y,\ \theta,\ \xi)$ 定义为上述分布之间的对数似然比：

$$U_S(y, \theta, \xi) = \log(p(\theta|y, \xi)) - \log(p(\theta)) = \log \frac{p(\theta|y, \xi)}{p(\theta)}, \tag{5}$$

因此，实验 ξ 在方程（1）中的平均效用可写为：

$$\begin{aligned}
U_S(\xi) &= \int_Y \int_\Theta p(y|\xi) \cdot p(\theta|y, \xi) \cdot \log \frac{p(\theta|y, \xi)}{p(\theta)}\, d\theta\, dy && (6)\\
&= \int_Y \int_\Theta p(y|\xi) \cdot p(\theta|y, \xi) \cdot \log(p(\theta|y, \xi))\, d\theta\, dy - \int_Y \int_\Theta \underbrace{p(y|\xi) \cdot p(\theta|y, \xi)}_{p(\theta, y|\xi) = p(\theta)p(y|\theta, \xi)} \cdot \log(p(\theta))\, d\theta\, dy && (7)\\
&= \int_Y p(y|\xi) \int_\Theta p(\theta|y, \xi) \cdot \log(p(\theta|y, \xi))\, d\theta\, dy - \int_\Theta p(\theta) \cdot \log(p(\theta)) \int_Y p(y|\theta, \xi)\, dy\, d\theta && (8)\\
&= -\mathbb{E}_{Y \sim p(y|\xi)}\{H(\Theta|Y, \xi)\} + H(\Theta) && (9)\\
&= I(\Theta; Y|\xi)。 && (10)
\end{aligned}$$

事实上，方程（10）中的最终表达式就是著名的 互信息 ，它表示 Y 与 Θ在实验 ξ 中的互信息，记作 $I(\Theta; Y|\xi)$（Cover 1999）。因此，方程（2）中的最优贝叶斯实验设计可改写为：

$$\xi^* = \arg\max_{\xi \in \Xi} I(\Theta; Y|\xi) \tag{11}$$

由于 $H(\Theta)$ 与实验 ξ无关，因此方程（11）中优化问题的目标项可简单解释为对期望后验熵的最小化（注意方程（ − 9）中的$\mathbb{E}_{Y \sim p(y|\xi)}\{H(\Theta|Y, \xi)\}$，见方程（9）。

值得注意的是，方程（5）中定义的效用函数 $U_S(\xi)$ 还可以采用另一种表述：事实上，由方程（6）可知，该函数可视为先验分布（$p(\theta)$）$_{\theta \in \Theta}$ 与后验分布（$p(\theta|y, \xi)$）$_{(\theta,y) \in \Theta \times \mathcal{Y}}$之间的平均 库尔贝克–莱布勒散度 (Cover 1999：

$$U_S(\xi) = \mathbb{E}_{Y \sim p(y|\xi)}\left\{D_{KL}(p(\theta|Y, \xi)||p(\theta))\right\} \tag{12}$$

其中 fo 对所有 $y \in \mathcal{Y}$ 以及对所有 $\xi \in \Xi$ 库尔贝克–莱布勒散度（Cover 1999）定义为 s:

$$D_{KL}(p(\theta|y, \xi)||p(\theta)) = \int_\Theta \log \frac{p(\theta|y, \xi)}{p(\theta)} \cdot p(\theta|y, \xi)\, d\theta. \tag{13}$$

下一节将介绍最大熵抽样准则，以及它与前文所述最优贝叶斯实验设计问题之间的关系。

2.3. 最大熵抽样

最大熵抽样（Shewry 与 Wynn，1987年）是一个信息论问题，旨在选择一项实验 ξ。在此情境下，实验指的是样本，其目标是最大化从未采样位置的预测中获得的信息增益。

用信息论语言来说，最大熵采样准则旨在选择一个样本，使系统的不确定性或熵最小。设 $\Gamma_{\mathcal{S}} = (\Gamma_i)_{i\in\mathcal{S}}$ with $\mathcal{S} = \{1, ..., N\}$ 是一个由随机向量 Γ_i 描述的有限系统，其中 $i \in \mathcal{S}$ 是观测点的索引。设 s，$\bar{s}$ 是 $\mathcal{S}$ 的一个二元划分，即 $\mathcal{S} = s\dot{\cup}\bar{s}$，其中 $\Gamma_s = (\Gamma_i)_{i\in s}$ 表示用于估计整个系统 $\Gamma_{\mathcal{S}} = (\Gamma_i)_{i\in\mathcal{S}}$ 而进行的观测或测量。因此，问题的目标是确定样本集 $s \subseteq \mathcal{S}$，使在给定系统 Γ_s 的观测后，系统未观测部分 $\Gamma_{\bar{s}}$ 的平均不确定性最小，即最小化

$$\mathbb{E}_{\Gamma_s}\{H(\Gamma_{\bar{s}}|\Gamma_s)\} \tag{14}$$

根据熵的链式法则（Cover 1999），对于 Γ_S，可以得到以下表达式：

$$H(\Gamma_{\mathcal{S}}) = H(\Gamma_s) + \mathbb{E}_{\Gamma_s}\{H(\Gamma_{\bar{s}}|\Gamma_s)\}. \tag{15}$$

利用这一恒等式，并考虑到 $H(\Gamma_{\mathcal{S}})$ 是固定的（即独立于 s）且有限的，上述对 (14) 的最小化问题可归结为最大化 $H(\Gamma_s)$；也就是说，选择样本

$$s^* = \arg\max_{s\subseteq S} H(\Gamma_s)$$

中观测不确定性最高的样本。换言之，最大熵采样准则旨在选择不确定性最大的观测。为实现这一目标，给定已抽样总体 Γ_s 时，未抽样总体 $\Gamma_{\bar{s}}$ 的条件不确定性最小。

我们必须强调，这一抽样准则并不等同于方程 (11) 中的贝叶斯最优实验设计，因为后者的评价指标是最小化期望后验熵。值得注意的是，Sebastiani & Wynn (2000) 已证明，在某些特定假设下，这两个问题是等价的。按照上一节介绍的贝叶斯实验框架，设 $y \in Y$ 为系统的观测，$\theta \in \Theta$ 为其参数，$\xi \in \Xi$ 为一项实验。令 $\Gamma_s = Y|\xi$ 和 $\Gamma_{\bar{s}} = \Theta$ 代入方程 (15)，可得以下表达式：

$$H(Y, \Theta|\xi) = H(Y|\xi) + \mathbb{E}_{Y\sim p(y|\xi)}\{H(\Theta|Y, \xi)\}. \tag{16}$$

如果 $H(Y, \Theta|\xi)$ 固定，即独立于任何实验 $\xi \in \Xi$，且方程（16）中的所有项均为有限值，那么通过最大化观测空间 $H(Y|\xi)$ 中的边缘熵，即可实现对期望后验熵 $\mathbb{E}_{Y\sim p(y|\xi)}\{H(\theta|Y, \xi)\}$ 的最小化。换言之，如果联合香农熵 (Y, Θ) 不依赖于实验 ξ 本身，那么最大熵采样准则就等价于贝叶斯最优实验设计问题。然而，通常很难直接保证满足这一条件。

幸运的是，联合分布 (Y, Θ) 与实验 ξ 之间的独立性条件可以重新表述。交换方程（ Y 和 Θ ）中 16的作用，可得：

$$H(Y, \Theta|\xi) = H(\Theta) + \mathbb{E}_{\Theta\sim p(\theta)}\{H(Y|\Theta, \xi)\}. \tag{17}$$

由于 $p(\theta)$ 与实验无关（见第 2.1 节），前述关于 $H(Y, \Theta|\xi)$ 相对于 ξ 的不变条件，等价于要求项 $\mathbb{E}_\Theta\{H(Y|\theta, \xi)\}$ 独立于实验 ξ。值得注意的是，在我们针对双星系统观测的最优调度方法中，这一重新表述的条件更容易检验：即最大熵采样准则与贝叶斯最优实验设计问题之间的等价性。

3. 双星观测的最优排程

本节介绍双星系统的统计模型，并提出一种新颖的双星系统观测最优调度方法，该方法基于第2.3节所述的最大熵抽样。

3.1. 统计模型

本节介绍的双星统计模型遵循Videla 等人（2022年）提出的模型，但对符号作了少量调整，以便于后文第3.2节所述最优调度方法的表述。

首先，假设各个位置观测和径向速度观测服从高斯分布，并对模型参数采用均匀先验（除非另有特定的先验分布）。令$\{t_i, X_i, Y_i\}_{i=1}^{n}$ 表示一组 n 位置观测，即以直角坐标表示的伴星相对于主双星系统的位置。此外，令 $\{\bar{t}_i, V_{1i}\}_{i=1}^{n_1}$ 和 $\{\tilde{t}_i, V_{2i}\}_{i=1}^{n_2}$ 分别表示主星和伴星的一组 n_1 和 n_2 径向速度观测。给定轨道参数向量 θ （固定但未知），每项观测（测量值）均服从相互独立的高斯分布，其中心值为开普勒模型 $f^{*}_{\mathrm{kep}}(\theta,\ t)$在给定历元 t下得到的值，标准差等于相应的观测误差σ_i:

$$X_i \sim \mathcal{N}(\ f^{x}_{\mathrm{kep}}(\theta, t_i), \sigma_i^2),\ Y_i \sim \mathcal{N}(f^{y}_{\mathrm{kep}}(\theta, t_i), \sigma_i^2),\ V_{1i} \sim \mathcal{N}(f^{v_1}_{\mathrm{kep}}(\theta, \bar{t}_i), \sigma_i^2),\ V_{2i} \sim \mathcal{N}(f^{v_2}_{\mathrm{kep}}(\theta, \tilde{t}_i), \sigma_i^2), \tag{18}$$

其中 $f^{x}{}_{\mathrm{kep}}(\theta,\ t)$ 和 $f^{y}{}_{\mathrm{kep}}(\theta,\ t)$ 遵循方程 (A4)，而 $f^{v_1}{}_{\mathrm{kep}}(\theta,\ t)$ 和 $f^{v_2}{}_{\mathrm{kep}}(\theta,\ t)$ 分别遵循方程 (A6) 和 (A7)。[3]

记完整观测向量为 $\mathcal{D} = \{t_i, X_i, Y_i\}_{i=1}^{n} \cup \{\bar{t}_i, V_{1i}\}_{i=1}^{n_1} \cup \{\tilde{t}_i, V_{2i}\}_{i=1}^{n_2}$，则给定参数向量 θ 时， $\mathcal{D}$的对数似然为:

$$\begin{aligned}\log p(\mathcal{D}|\theta) = &\sum_{i=1}^{n} \log \mathcal{N}(X_i | f^{x}_{\mathrm{kep}}(\theta, t_i), \sigma_i^2) + \sum_{i=1}^{n} \log \mathcal{N}(Y_i | f^{y}_{\mathrm{kep}}(\theta, t_i), \sigma_i^2) \\ &+ \sum_{i=1}^{n_1} \log \mathcal{N}(V_{1i} | f^{v_1}_{\mathrm{kep}}(\theta, \bar{t}_i), \sigma_i^2) + \sum_{i=1}^{n_2} \log \mathcal{N}(V_{2i} | f^{v_2}_{\mathrm{kep}}(\theta, \tilde{t}_i), \sigma_i^2).\end{aligned} \tag{19}$$

每个标量轨道参数 θ_i 的先验分布建模为在其有效物理范围内的均匀分布，该范围记为 $\Theta_i \subset \mathbb{R}$ (Videla 等（2022），附录 A）。因此，在假设各分量相互独立的情况下，轨道参数向量的先验分布为:

$$\log p(\theta) = \sum_{i=1}^{|\theta|} \log U(\min \Theta_i, \max \Theta_i), \tag{20}$$

其中 $\theta_i \in \Theta_i,\ \forall i \in \{1, ..., |\theta|\}$，而 U是定义在 $\min \Theta_i$ 与 $\max\ \Theta_i$之间的均匀分布的概率密度函数。

在推断阶段，根据贝叶斯定理，给定证据 $\mathcal{D}$时，后验分布正比于似然与先验的乘积，即 $p(\theta|\mathcal{D}) \propto p(\mathcal{D}|\theta)p(\theta)$。因此，可以通过采用任意抽样技术获得完整的后验分布（Hastings 1970；Geman &Geman 1984；Goodman& Weare 2010；Neal 等人 2011；Hoffman 等（2014））。

对于本文研究的具体案例，双线光谱双星模型由轨道参数向量 $\theta_{SB2} = \{P, T, e, a, \omega, \Omega\ i、\ V_0、\ \varpi、\ q\}$ 表示，而单线光谱双星模型则由 $\theta_{SB1} = \{P, T, e, a, \omega, \Omega\ i、\ V_0、\ f/\varpi\}$ 表示。视觉三体层级恒星系统则由轨道参数向量 $\theta_{triple} = \{P_{A_aA_b}, T_{A_aA_b}, e_{A_aA_b}, a_{A_aA_b}, \omega_{A_aA_b}, \Omega_{A_aA_b}、\ i_{A_aA_b}\} \cup \{q_{A_aA_b}\} \cup \{P_{AB}, T_{AB}, e_{AB}, a_{AB}, \omega_{AB}, \Omega_{AB}、\ i_{AB}\}$ 表示。有关开普勒轨道模型的更多细节，请参阅附录 A和Videla 等人（2022年）。

3.2. 最优观测调度

根据贝叶斯框架，设 $y \in Y$和 $\theta \in \Theta$ 分别为某个系统的观测和参数。方程（11）中的最大信息提取问题，是寻找实验 $\xi \in \Xi$，使期望后验熵 $\mathbb{E}_Y\{H(\Theta|Y,\ \xi)\}$最小。该框架可用于确定双星系统的最优观测时间。考虑到 Y 是观测空间（由位置空间和径向速度空间组成），Θ 是轨道参数空间（用于表征恒星系统），而 ξ 是系统的观测时刻（历元）

[3]有关开普勒模型的详细介绍，请参见附录 A。

（以下简称 t ），方程（11）中的问题是寻找观测时刻$t^* \in \mathcal{T}$，使其最大限度地获取系统轨道参数的信息。为此，我们提出采用第2.3节介绍的最大熵采样准则。

根据第 3.1 节介绍的双星系统贝叶斯建模，假设每次观测均服从高斯分布，其均值由时刻 t 的开普勒轨道模型决定，方差则由相应的观测误差决定。因此，设 $Y = (Y_1, ..., Y_n)$ 为恒星系统的一组有限个可能的新观测（位置或径向速度），统计模型满足：

$$Y_i|(\theta, t_i) = f^*_{\text{kep}}(\theta, t_i) + \epsilon_i, \tag{21}$$

其中，$\epsilon_i \sim \mathcal{N}(0, \sigma_i^2)$ 是观测误差，$f^*_{\text{kep}}(\cdot)$ 是将时刻 t_i 映射到观测空间 Y_i 的函数，该映射由恒星系统的轨道参数 θ 确定。

我们注意到，根据 Sebastiani & Wynn（2000），$Y_i|(\theta, t_i)$ 是一个回归模型。因此，如果进一步假设随机变量 ϵ_i 与 $f^*_{\text{kep}}(\theta, t_i)$ 及 $\sigma_i = \sigma, \forall i \in \{1, ..., n\}$ 相互独立，那么有 $H(Y|\theta, t) = H(\epsilon)$，即 $H(Y|\theta, t)$ 在函数上独立于观测时间 t（实验）的选择。满足这一条件足以证明第 2.2 节提出的最大信息提取问题与第 2.3 节所述的最大熵采样准则等价。

值得一提的是，假设 $\sigma_i = \sigma$ 对任意 $i \in \{1, ..., n\}$ 都成立，意味着所有可能的未来观测都必须具有相同的不确定性。在观测天文学中，这一条件意味着，无论未来观测的数值为何（坐标空间中的位置或径向速度的大小），仪器（望远镜）都以相同的精度测量观测证据。当然，如果未来观测使用相同的仪器配置获取，这一假设是有效且符合实际的；但当考虑精度不同的仪器时，该假设便不再成立。

因此，利用方程（16）处理现有观测 $\mathcal{D}$（即方程（ $\Gamma_s \equiv Y|\mathcal{D}, \ t$ 中的15），并注意到所讨论的假设使项 H（Y, $\Theta|\mathcal{D}, \ t$）在函数上独立于 t的取值，那么，使关于系统轨道参数的信息最大化，即使期望后验熵最小化的观测时间，就是 $t^* \in \mathcal{T}$使观测空间的边缘熵最大化的时刻：

$$t^* = \arg\min_{t \in \mathcal{T}} \mathbb{E}_{Y \sim p(y|t)}\{H(\Theta|Y, \mathcal{D}, t)\} = \arg\max_{t \in \mathcal{T}} H(Y|\mathcal{D}, t). \tag{22}$$

3.3. 归一化调度

我们注意到，方程（22）中的调度问题存在一个实际问题：决策空间 $\mathcal{T}$没有上界，而且没有简单的解析方法可以解决这一任务。为解决这一问题，我们研究该问题的内在规律性。事实上，我们注意到，每个轨道参数 $\theta \in \Theta$ 的取值都会在观测空间中产生一条具有给定轨道周期 P的轨道。这个特定周期会使观测在时刻 t 提供的信息呈现周期模式（一个 P周期的信息模式）。借助这一重复模式，我们可以将 t 的探索范围缩小到一个有界的归一化域，从而使方程（22）中的问题变得可解。

总之，为了克服上述实现挑战，我们建议将方程（22）中的观测时间随机变量（实验）替换为归一化观测时间 $\tau = (t - t_0)/P$（系统相位），其中 P 是每个轨道参数向量 $\theta \in \Theta$ 的周期，而 t_0 是初始观测时间（固定设计变量）。因此，我们只需在闭区间$\tau \in [0, 1]$ 上计算边缘观测熵 H（$Y|\mathcal{D}, \ \tau$），即可评估完全完成的后验轨道的不确定性，从而完整探索每个系统 $\theta \in \Theta$。因此，方程（22）中的最优观测时间问题可化简为：

$$\tau^* = \arg\min_{\tau \in [0,1]} \mathbb{E}_{Y \sim p(y|\tau)}\{H(\Theta|Y, \mathcal{D}, \tau)\} = \arg\max_{\tau \in [0,1]} H(Y|\mathcal{D}, \tau). \tag{23}$$

根据归一化解 τ^*，我们可以使用随机变量 P 返回未归一化域。这会得到最优观测时间 t^* 的概率分布：

$$t^* = \tau^* P + t_0. \tag{24}$$

对原始问题的这种重新表述，方程（22）使我们不仅能够对最优观测历元 t^*进行确定性估计，还能得到其完整的分布，其中所有不确定性都由与系统周期对应的时间随机变量 P所囊括。

请注意，方程（ $\tau \in [0,1]$ 23）中的归一化时间可以解释为从起始时间 t_0 起，由后验分布在观测空间中的投影$p(\Theta|\mathcal{D})$所生成的所有轨道的周期完成度百分比（以小数表示）。例如，当 $\tau = 0.5$ 时，我们估计完成一半的轨道上各点的熵；而当 $\tau = 1$ 时，我们估计完整轨道上各点的熵。这种方法不同于方程（22）中的原始问题，因为它估计的是在确定的观测时间 $t \in \mathcal{T}$下轨道投影的熵，完全忽略了这些天体固有的周期性。

3.4. 算法实现

在实现方面，我们的最优调度方法需要求解公式（23）。这需要估计在给定归一化观测时间 $\tau = (t-t_0)_P \in [0,1]$下，后验预测分布 $p(y|\mathcal{D},t) = \int_\Theta p(y|\theta,\mathcal{D},t)p(\theta|\mathcal{D},t)d\theta$ 的熵。幸运的是，后验预测分布的样本可以通过将后验分布的样本投影到观测空间中直接获得（例如，Videla 等（2022）所采用的方法）；而后验分布的样本则可以使用任意马尔可夫链蒙特卡罗方法进行估计（例如，Gregory（2005, 2011）；Hou 等人（2012）；Nelson 等人（2014）；Mendez 等（2017）；Videla 等（2022））。最后，边缘后验熵可以使用任意熵估计算法计算（例如，Parzen（1962）；Kozachenko 与 Leonenko（1987））。最优调度方法涉及的步骤总结于算法1中。

算法1： 最优观测调度。

参数： n, m, t_0, $p(\theta|\mathcal{D})$
初始化 $\mathbf{D}_{n,m}$, $\mathbf{h}_n$,$\mathbf{p}_n$,$\mathbf{t}^*_n$
对于 $0 \le i < n$，**执行**
  从后验分布中抽取样本
  $\theta \sim p(\theta|\mathcal{D})$
  获取轨道周期
  $\mathbf{p}_i = \text{GETPERIOD}(\theta)$
  将投影样本投影到观测空间
  对 $0 \le j < m$ **执行**
    $\tau = j/(m-1)$
    $t = \tau \cdot \mathbf{p}_i + t_0$
    $\mathbf{D}_{i,j} = f^*_{\text{kep}}(\theta, t)$
估算边缘观测熵
对于 $0 \le j < m$ **do**
  小时$_j = \text{ENTROPYESTIMATOR}(\text{天}_{:,j})$
计算最优观测时间分布
$j^* = \arg\max_j \mathbf{h}_j$
$\tau^* = j^*/(m-1)$
$\mathbf{t}^* = \tau^* \cdot \mathbf{p} + t_0$
返回 $\mathbf{t}^*$

3.5. 对我们的调度方法进行分析

我们用于双星观测的最优调度方法的优势可概括为以下几点：

- **推断效率：** 贝叶斯最优实验设计方法的一种朴素实现方式，是在每个候选时刻估计纳入虚拟样本后的后验分布 τ_i, $i \in [0,1]$ 以评估这些样本对确定系统所提供的信息贡献。由于贝叶斯推断过程的计算成本高昂，这种方法只能评估少数几个候选时刻。相比之下，我们提出的最大熵抽样方法只需估计一次原始后验分布（不含虚拟样本），再利用该分布在各候选时刻向观测空间的投影，评估这些时刻的信息贡献。

• **不确定性估计效率：** 最大熵采样准则将最优观测时刻估计问题简化为计算观测空间中的熵 Y。这项计算至多涉及四个维度（考虑双星系统中主星和伴星的二维位置／天体测量观测以及径向速度／光谱观测），而不必计算轨道参数空间 Θ 中的熵；后者有几十个维度（见附录A）。此外，在层级恒星系统中，轨道参数空间会随着系统´的成员数量增加而迅速扩大。因此，在低维观测空间中估计熵，能够大幅降低计算成本；更重要的是，这还通过避免维数灾难相关现象，简化了估计问题。

• **信息代表性：** 上述方法的计算效率使我们能够在一组密集的归一化时间点上估算边缘观测熵 $\tau \in [0,1]$，从而计算边缘观测熵随时间变化的伪连续曲线（见第1、4和7 节中的图4）。这条不确定性曲线不仅使我们能够选出信息量最大的归一化时间点（通过应用方程（22）），还完整地刻画了未来观测所能带来的信息增益随时间的变化。这一宝贵的信息可用于比较不同可能观测之间的信息差距，从而确定其优先级（见图1、4和7的下方面板）。请注意，即使存在观测约束（例如由太阳或望远镜／技术约束造成），我们仍可进行相同的分析，只需从熵曲线中去除无法进行观测的时间段即可。

• **序贯决策：** 前述各点提到的所有属性均可用于实施序贯观测规划策略。如果我们有一系列观测时间可供使用，就可以逐一选择。此外，刻画未来观测的信息增益可作为停止获取新样本的判据。例如，如果最佳选择的信息增益（实施方程 (22) 中的决策后）低于预先设定的某个阈值，这可能意味着额外测量无法提高对系统轨道参数的推断，因此已无必要。

最后，值得指出的是，第 2.1 节所述的最优贝叶斯设计方法假设 θ 与尚待选择的实验索引（$\xi \in \Xi$）之间不存在直接依赖关系（第 2.3 节的方程 (17) 中采用了这一假设）。通过对问题进行归一化并使其变得可处理，我们可以从归一化时间构造看出，该量部分取决于 θ （因为它取决于轨道周期 P）。值得指出的是，这一观察并不妨碍实际采用我们的策略，该策略展现了显著的效果。理论上，如果 P 得到了精确确定，那么式 (23) 中的归一化设定与原始表述之间的差距为零。这或可解释为何在 P 估计精度较高时能得到近乎最优的结果；本研究考察的两个案例研究中就观察到了这一点。最后，在 P 估计不够精确的情况下，这种归一化策略的理论影响是一个值得研究的重要课题，我们将在未来的工作中加以探讨。

4. 实验验证

我们将所提出的最优观测时间估计方案应用于三个双星系统：双线视觉光谱双星系统 HIP 89000、单线视觉光谱双星系统 HIP 99675，以及视觉三合层级系统 LHS 1070。这些示例并非旨在穷举所有可能的应用，而是展示我们的方法的性能。以下实验仅考虑轨道空间中的不确定性（X， Y），因为现有径向速度观测具有良好的覆盖度和精度，使得径向速度后验不确定性可以忽略不计。

（轨道空间中的）边缘观测熵 $H(X,\ Y|\mathcal{D},\ \tau)$ 沿归一化时间的有效范围 $\tau \in [0,1]$ (其中 t_0 为最新现有观测的历元)，采用 k 近邻熵估计法（Kozachenko 和 Leonenko（1987））， 基于从后验分布$\theta \sim p(\Theta|\mathcal{D}$中随机选取的 1000 个样本进行计算。这些样本通过马尔可夫链蒙特卡罗算法——无转向采样器（Hoffman 等（2014））估计，该算法此前已在Videla 等人(2022年)中介绍。遵循该研究的实验设置，我们从由拟牛顿优化法 L-BFGS（Liu 和 Nocedal（1989））确定的起始点出发，模拟 4 条相互独立的马尔可夫链。

我们在五个具有代表性的归一化时间 $\tau_i, i \in \{1,2,3,4,5\}.$，比较边缘观测熵 $H(X,Y|\mathcal{D},\tau_i)$ 与后验预测分布 $p(X,Y|\mathcal{D},\tau_i)$。此外，我们在相同的代表性归一化时间加入虚拟位置观测 d_{τ_i}（并通过无转向采样算法重新进行推断），从而估计更新后的后验分布 $p(\Theta|\mathcal{D}\cup\{d_{\tau_i}\})$。每个虚拟观测对应于最大后验轨道（由最大后验轨道参数 θ_{map} 得出）在历元 $t_i^{map} = \tau_i \cdot P_{map} + t_0$处的位置，即 $d_{\tau_i} = (f^x\text{kep}(\theta_{map}, t_i^{map}) f^y\text{kep}(\theta_{map}, t_i^{map}))$，其中$\sigma_i = \sigma_{min}$，而 σ_{min} 是现有观测 $\mathcal{D}$中的最小标准差。为验证我们的方法，我们将边缘观测熵 $H(X,Y|\mathcal{D},\tau_i)$ 与相应的更新后验分布 $p(\Theta|\mathcal{D}\cup\{d_{\tau_i}\})$ 进行比较。

最后，我们给出了与边缘观测熵最高的归一化时间 $\tau^* = \text{arg max}_{\tau\in[0,1]} H(Y|\mathcal{D},\tau)$ 对应的最优观测历元 t^* 的估计分布。

4.1. *HIP 89000*

HIP 89000 系统（发现编号 YSC132AaAb）是一个 $SB2$ 双星，最近由 Videla 等人 (2022年) 提出并完成求解。现有数据包括主要集中在远星点通过附近（该双星非常紧密）的干涉观测，以及对两个组分丰富而精确的径向速度观测。天体测量观测及其误差显示于图 1 （左上面板）。估计的边缘观测熵曲线 $H(X,Y|\mathcal{D},\tau)$ $\tau \in [0,1]$、后验预测分布 $p(X,Y|\mathcal{D})$、它们在选定的五个归一化时间 $\tau_i, i \in \{1,2,3,4,5\}$处各自的投影，以及最优观测时间 t^* 的估计分布，也一并展示于图1 中。

估算的熵曲线显示出两个局部极大值，其间有一个显著的局部极小值；曲线起始处 $\tau = 0$ 和末尾处 $\tau = 1$ 也各有一个极小值。估算的熵曲线变化与投影到观测空间中的后验分布观测离散度相吻合：熵最小值对应于轨道上由高精度观测（$\tau = 0$）覆盖的区域，这些区域的投影轨道呈现为狭窄带状。随着 τ，熵随之增加；由于轨道空间该区域没有精确观测，投影轨道的离散度也随之增大（τ_1）。这种变化持续到熵达到第一个局部极大值，此时投影轨道的离散度最高（τ_2），这是因为轨道该区域的观测数量少且精度低。随后，熵和投影轨道的离散度开始下降，直至达到一个局部极小值（τ_3）。值得注意的是，即使该区域附近没有位置观测，投影轨道仍会变得更窄（不确定性更低）；这归因于在轨道其他区域进行的观测降低了某些轨道参数的不确定性。这一结果表明，认为观测总应安排在此前没有观测的位置这一朴素想法，未必是最有效的做法。估算的熵曲线再次上升，直至达到第二个局部极大值（τ_4），其略高于前一个局部极大值（位于 τ_2）。由于轨道该区域完全没有观测，投影轨道的离散度达到最高。最后，在观测周期末尾（$\tau = 1$），熵曲线开始下降，并达到与曲线起始处相近的极小值（$\tau = 0$），这是由轨道的周期性所致。

纳入每个样本 d_{τ_i} 后得到的边缘更新后验分布如图2所示，相应的观测空间投影如图3所示。对比边缘更新后验分布可见，轨道参数的不确定性（分布的离散度）程度各不相同；虚拟观测 d_{τ_i} 对轨道参数空间不确定性的降低幅度最大时，其对应的熵 $H(X,Y|\mathcal{D},\tau_i)$也较高，反之亦然。图1右上面板中估计值 $H(X,Y|\mathcal{D},\tau_i)$ 与 $H(\Theta|\mathcal{D}\cup\{d_{\tau_i}\})$ 的对比清楚地反映了这一现象，证实最大熵采样准则与贝叶斯最优设计问题等价。当虚拟测量位于不确定性最高区域时，投影后的更新后验分布显示出显著的不确定性降低。另一方面，在不确定性较低的区域进行虚拟测量时，不确定性降低可以忽略不计，这符合最大熵采样原理的目标。值得注意的是，更新后的边缘后验分布 $SB2 + d_{\tau_3}$（图2中的红色分布）在轨道参数 P、 T、 e 和 Ω 上的离散度最低，因为相应的虚拟观测 d_{τ_3}（图1左上面板中的红色星标）吸收了 τ_3 中相应后验预测分布所呈现的不确定性（图1左上面板中的红色云团）。与其他所有情况不同，这种投影不确定性沿轨道分散，主要由轨道周期不确定性导致，因此，模拟

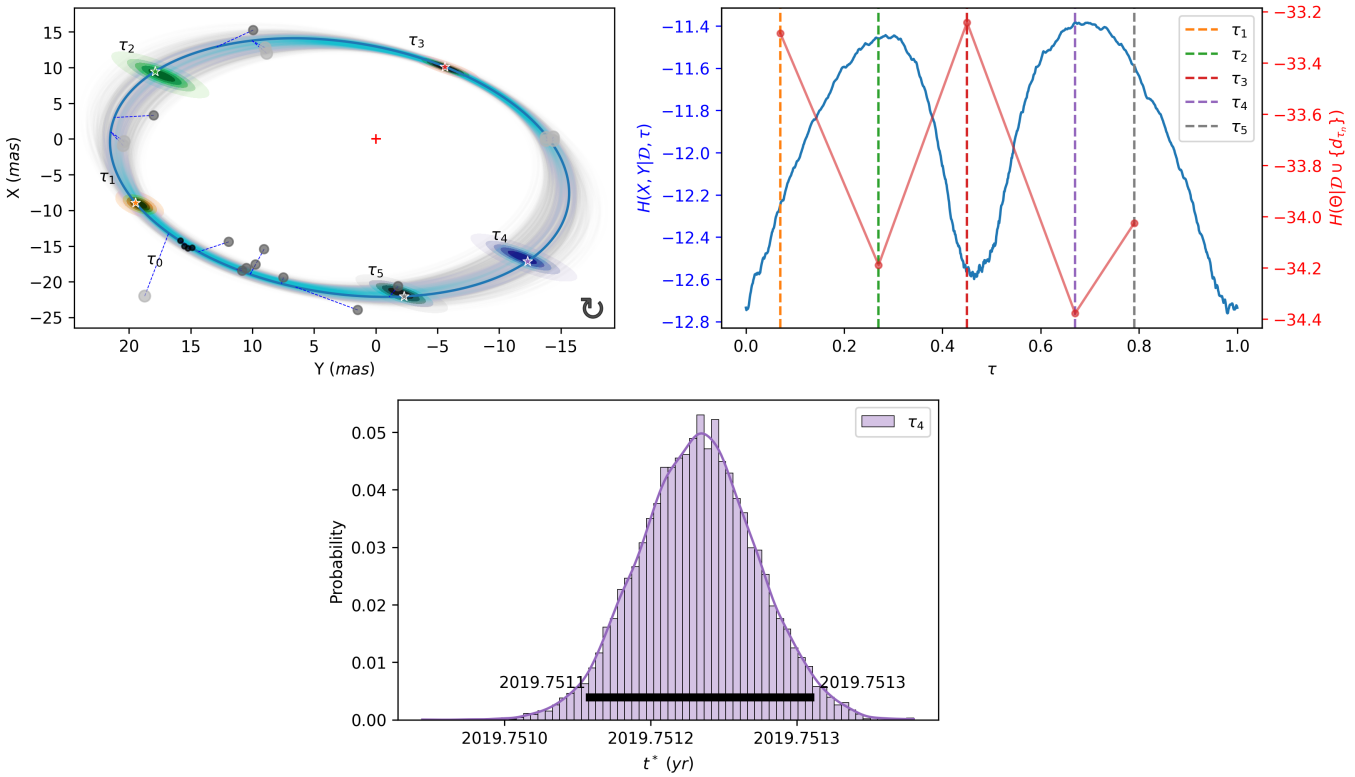

图1. 双星系统 HIP 89000 的最优观测时间分析。左上：沿 $\tau \in [0,1]$ 的后验预测分布$p(X,\ Y|\mathcal{D})$（由 1000 条青色轨道表示），并突出显示所选代表性历元$\tau_i,\ i \in \{1,2,3,4,5\}$ 处的投影（彩色云团）。每个彩色云团中的星形标记表示所选历元 τ_i 在最大后验轨道（蓝线）上的投影。实际干涉观测以灰点表示，其大小和深浅与所报告的不确定性成比例（不确定性越小，点越小、颜色越深）。这些观测点以虚线连接到其在最大后验轨道上的预测位置。右上：纳入位于五个代表性归一化时间 τ_i处的虚拟观测d_{τ_i} 时，估计的边缘观测熵曲线 $H(X,\ Y|\mathcal{D},\ \tau)$（左纵轴）和估计的更新后验分布熵 $H(\Theta|\mathcal{D} \cup \{d_{\tau_i}\})$（右纵轴）；这些时间由竖直虚线突出标示。下方：由观测空间中熵最高处对应的归一化时间 $\tau^* = \arg\max_{\tau \in [0,1]} H(X,\ Y|\mathcal{D},\ \tau)$所诱导的最优观测时间分布；本例中为 τ_4。直方图底部的黑色竖线表示 95% 高后验密度区间（HPDI），即包含后验分布 95% 概率质量（包括众数）的最窄区间。

在轨道该区域进行虚拟观测会使上述参数的不确定性降低幅度最大。然而，与原始后验分布相比，更新后的边缘后验分布 $SB2 + d_{\tau_3}$ 在其他轨道参数上也没有表现出不确定性降低 $SB2$ （分别为图2中的红色和蓝色分布），因此，主要轨道不确定区域（τ_2 和 τ_4）保持不变（图3中的红色轨道）。我们得出结论，对于所研究的系统，在 τ_3 进行观测所获得的信息增益最低，如图1右上面板中熵曲线的红色点线所示。

最后，图 t^* 下方面板展示了在观测空间不确定性最大的归一化时间（τ_4）处的最优观测时间分布。1。估计得到的分布呈现出清晰的高斯形状，离散度极低。这一结果表明，由于存在大量且精确的径向速度观测，恒星系统的周期 P 得到了很好的确定。

4.2. *HIP 99675*

系统 HIP 99675（发现编号 WRH33Aa,Ab）是一颗 $SB1$ 双星，最近由Videla 等人（2022年）介绍并求解。该天体的数据包括在轨道两个极端区域的少量天体测量观测，以及对主星分量进行的大量高精度径向速度观测。观测及其误差在图4 （左图）中展示（这也是一对需要进行干涉测量的近距双星）。估算得到的边缘观测熵曲线 H（X、$Y|\mathcal{D}$、τ） $\tau \in [0,1]$、后验预测分布

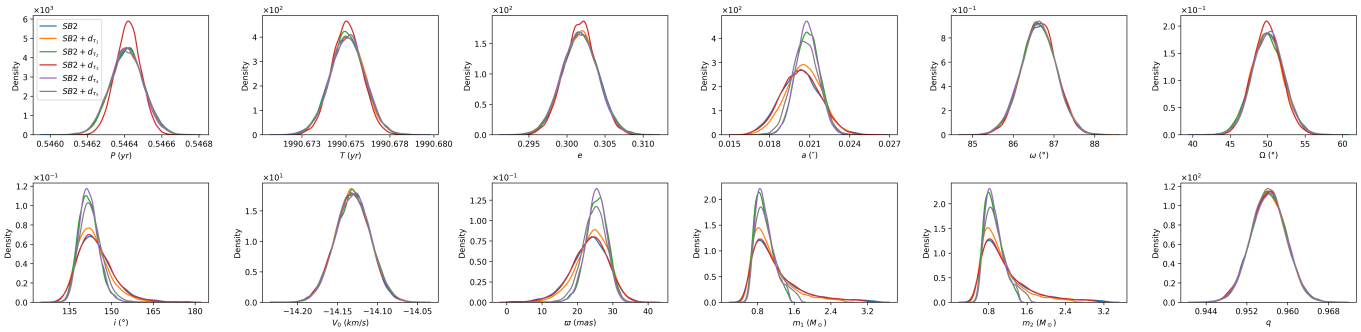

图2. 纳入虚拟观测后，双星 HIP 89000 的轨道根数及导出物理参数的边缘后验分布所受的影响 $d_{\tau_i} \sim \mathcal{N}(f_{\rm kep}(\theta_{map},\ \tau_i)\ \sigma^2)$ $i \in \{1,2,3,4,5\}$，其中 σ^2 固定为系统现有观测的最小方差 $\mathcal{D}$。这些图清晰展示了在不同历元获取数据的影响及重要性，这与图1右上面板的预测一致，即在边缘观测熵较大的区域获取观测，对降低参数不确定性的作用最大。

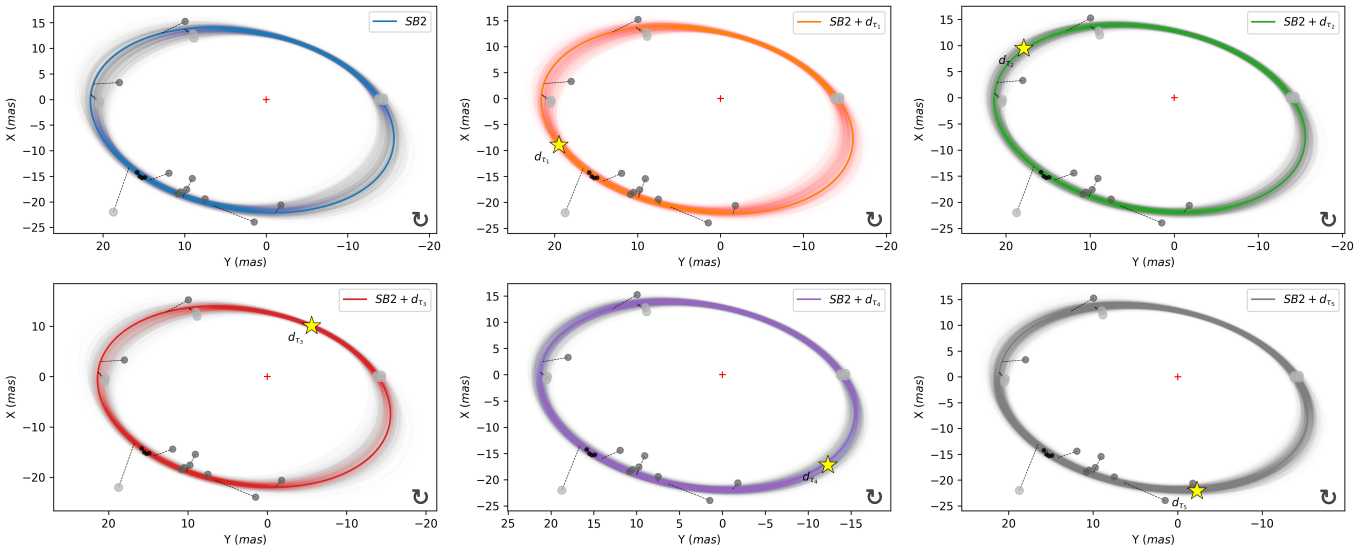

图3. 纳入历元为 d_{τ_i} 的虚拟观测后，系统 HIP 89000 更新后的后验预测分布在轨道构型上的投影 $\tau_i,\ i \in \{1,2,3,4,5\}$（以黄色星号标示）。轨道参数的数值请参见附录 1 中的表 B。

p（X、$Y|\mathcal{D}$）、它们在选定的五个归一化时间 τ_i、$i \in \{1,2,3,4,5\}$处各自的投影，以及最优观测历元的估计分布 t^* 也展示在图4中。

估算的熵曲线在最小熵点（τ_3）附近呈现准对称形态，离开该点后（$\tau < \tau_3$ 和 $\tau > \tau_3$），熵以波浪状上升，出现多个局部极小值和极大值，但整体始终保持平均上升趋势。估算的熵曲线形态显然与观测到的后验轨道无关；由于位置观测稀少，后验轨道形成了一片边界不明确的弥散线云。然而，将估算的熵曲线选出的五个归一化时间处的后验分布投影 $\tau_i,\ i \in \{1,2,3,4,5\}$ 可以看出，不确定性大多沿着与定义明确的轨道垂直的方向分布，而这些轨道彼此重叠，形成了弥散线云。轨道空间中熵最低的区域（τ_3）与大多数现有位置观测所在区域相吻合。相比之下，轨道的相对区域具有最高熵（τ_1，τ_5）。熵曲线的波浪状变化也反映在后验分布于轨道空间中的投影形状上：局部极小值对应较长但较窄的投影（τ_4），而局部极大值则对应稍短但较宽的投影（τ_2）。

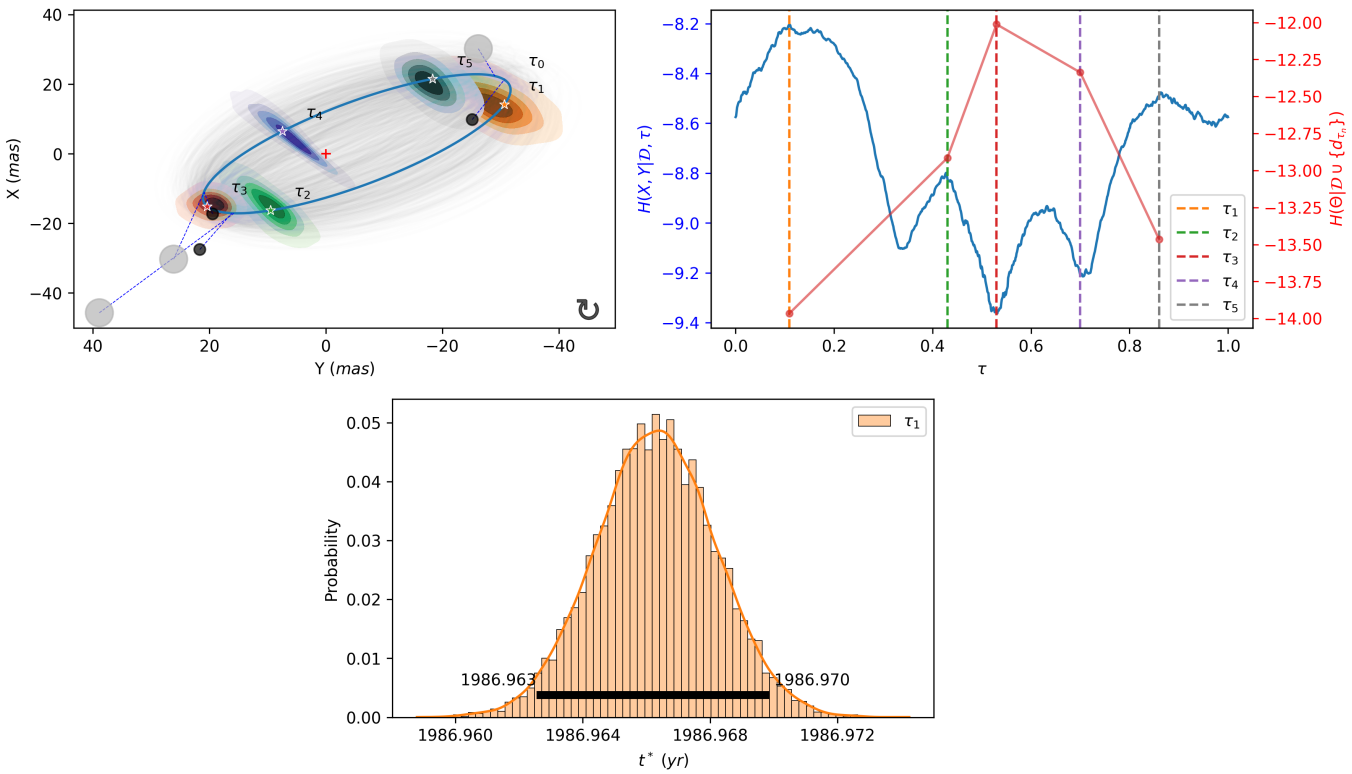

图4。 与图1相同，但对应于恒星系统 HI P 99675。

纳入每个样本后得到的边缘更新后验分布 d_{τ_i} 见图5，相应的观测空间投影见图 6。利用最低熵归一化时间处生成的虚拟观测更新后的后验分布（d_{τ_3}）使某些参数的边缘分布不确定性降低（a，Ω），但与其他更新后的后验分布相比，其他轨道参数（i）的分布仍相当分散（即不确定性没有降低）。在每种情况下，都能观察到部分轨道参数的不确定性降低与其他参数不确定性增加之间的这种权衡，从而产生不同的更新后验预测分布。正如预期，纳入的虚拟观测会降低轨道空间中靠近该测量值区域的投影不确定性。例如，d_{τ_3} 该观测大幅降低了轨道长度的不确定性，同时半长轴 a 和 Ω 的更新边缘后验分布的不确定性也有所降低。然而，该虚拟观测并未提供有关轨道倾角的任何重要信息，这一点体现为 i 的更新边缘后验分布散布范围较大（相反，如图 5所示，在 d_{τ_4} 处进行观测将有助于非常精确地确定i）。更新后后验分布的计算熵 $H(\Theta|\mathcal{D} \cup \{d_{\tau_i}\})$考虑了后验分布在不同维度上的不确定性增加与降低之间的上述权衡。

再次如 HIP 89000 的情况一样，右上面板图 $H(X,\ Y|\mathcal{D},\ \tau_i)$ 和 $H(\Theta|\mathcal{D}\cup\{d_{\tau_i}\})$ 中展示的熵对比表明，最大熵采样准则与贝叶斯最优设计问题之间的等价性得以保留。

最优观测时间的分布 t^* 位于观测空间不确定性最大的归一化时间处（τ_1），如图4下图所示。与 HIP 89000 的情况类似，估计得到的分布呈现出清晰的高斯形状，离散度极低。这反映出恒星系统的周期 P 得到了很好的确定，这归功于丰富且精确的径向速度观测；HIP 89000 的情况也是如此。

4.3. *LHS 1070*

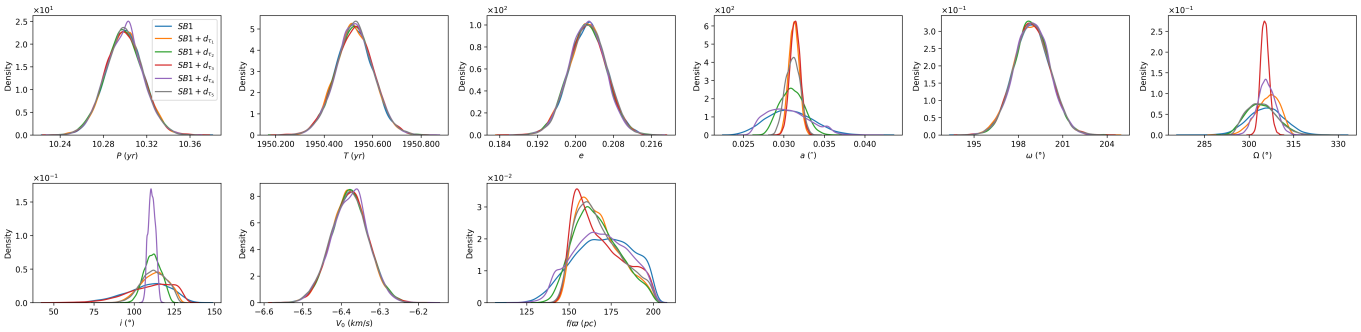

图5。 与图 2 相同，但对象为恒星系统 HIP 99675。由于这是一个单线光谱双星，我们无法像处理 HIP 89000 那样计算轨道视差和各个恒星质量，而是可以计算比值 f/ϖ，其定义见 Videla 等人(2022年)（参见该文第 2.2 和 2.4 节）。

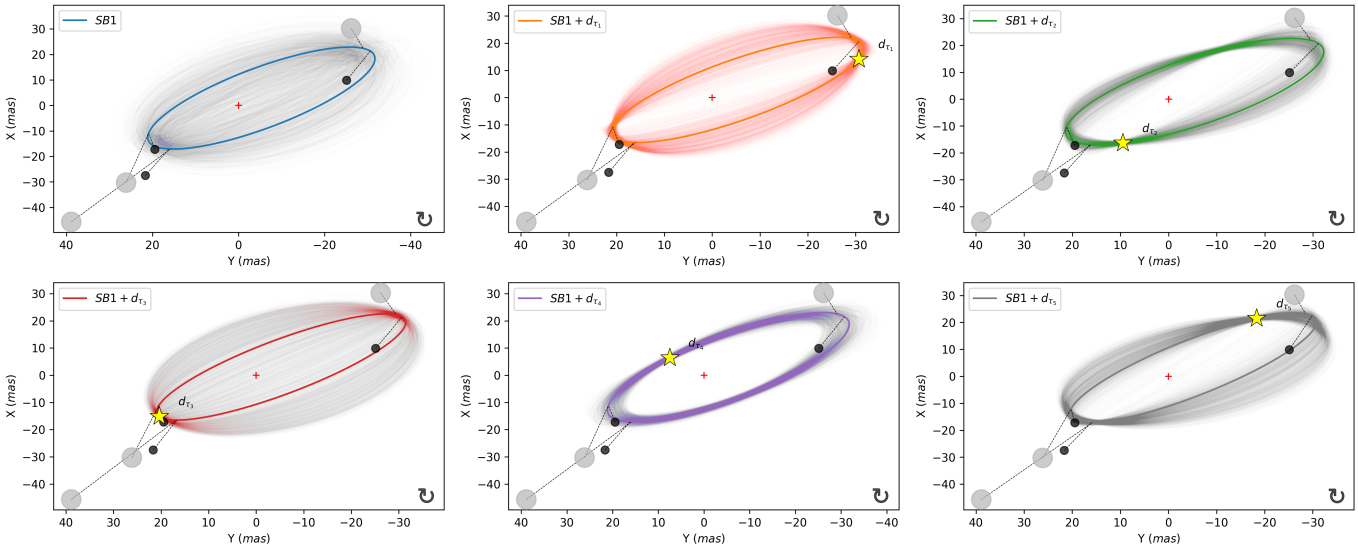

图6。 与图 3 相同，但对象为恒星系统 HIP 99675。轨道参数的数值请见附录 2中的表 B。

LHS 1070 系统是一个层级三合星系统，最近由Villegas 等人(2021年) 展示并求解。该层级系统由紧密的外层双星 BaBb 组成，其围绕较亮的主星分量 A 运行。现有数据包括大量且精确的 BaBb 双星位置观测，这些观测覆盖了整个轨道；还包括外层双星中较亮分量 Ba 相对于主星分量 A 的精确观测。由于其亮度较低（$V = 15.3$），该系统的任何分量都没有径向速度观测´。观测数据及其误差如图7 （左上）所示。估计的边缘观测熵曲线 $H(X, Y|\mathcal{D}, \tau)$ $\tau \in [0, 1]$、后验预测分布 $p(X, Y|\mathcal{D})$、它们在五个选定归一化时间处各自的投影 $\tau_i, i \in \{1, 2, 3, 4, 5\}$，以及最优观测时间的估计分布$t^*$ 如图7所示。

估计的熵曲线在起始处 $\tau = 0$呈现全局最小值，这与轨道上位置观测丰富且精确的区域相吻合。从轨道空间中的位置观测处开始，熵曲线陡增，直至在 τ_3达到最大值。曲线其余部分保持恒定，并在 τ_4之后开始略有下降。在评估的所有归一化时间 τ 上，熵曲线均呈现振荡行为。这种行为与观测空间中投影轨道的轨道瓣相吻合。在归一化时间 τ_i、$i \in \{1, 2, 3, 4, 5\}$ 下轨道空间中后验分布的投影表明，远离现有观测的区域中，后验分布的离散度增大，形成波浪状图案。这种模式可归因于投影轨道之间的偏移增大，这是由于.....的不确定性较大

纯天体测量层级系统的轨道周期（如前所述，该系统没有可用的视向速度数据）。

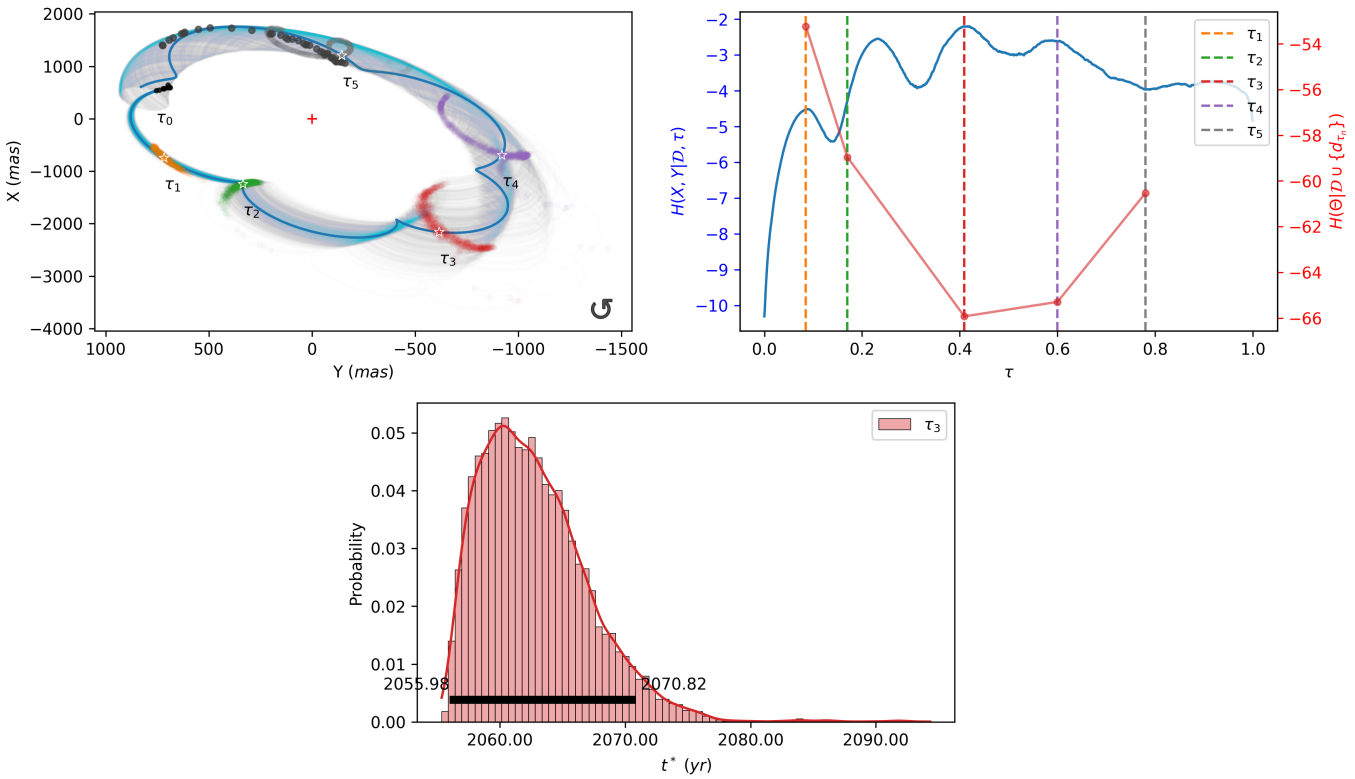

图7。 与图 1相同，但针对恒星系统 LHS 1070。

纳入每个样本 d_{τ_i} 后得到的边缘更新后验分布见图8，相应的观测空间投影见图9。对边缘更新后验分布的比较显示，外轨道AB的轨道参数具有不同程度的不确定性；其中，在轨道参数空间中实现最大不确定性降低的虚拟观测 d_{τ_i}，其对应的熵 $H(X, Y|\mathcal{D}, \tau_i)$ 最高，反之亦然。图7右上面板中对 $H(X, Y|\mathcal{D}, \tau_i)$ 和 $H(\Theta|\mathcal{D} \cup \{d_{\tau_i}\})$ 估计值的比较清楚地反映了这种相反的行为，这再次验证了最大熵采样准则与贝叶斯最优设计问题之间的等价性。更新后的后验分布显示，在不确定性最高区域获得虚拟测量时，轨道相应区段的不确定性显著降低。另一方面，在不确定性较低的区域进行虚拟测量时，不确定性降低微乎其微，这再次符合最大熵采样原理的目标。我们注意到，更新后的后验预测分布（见图9）表明，在第一个轨道瓣上纳入虚拟观测（$AS_{triple} + d_{\tau_1}$ 情况），不仅消除了该轨道瓣的不确定性，也消除了下一个轨道瓣（按旋转方向）的绝大部分不确定性。同样，在前两个轨道瓣的交点处纳入虚拟观测（$AS_{triple} + d_{\tau_2}$ 情况），会显著降低这两个轨道瓣的不确定性，并降低第三个轨道瓣相当一部分的不确定性。值得注意的是，在第三个轨道瓣上纳入虚拟观测（$AS_{triple} + d_{\tau_3}$ 情况），几乎消除了全部五个轨道瓣的所有不确定性，是整个系统中不确定性降低幅度最大的情况，与熵曲线（图7右上面板）的预测一致。此后我们观察到，在后续轨道瓣上纳入虚拟观测（$SB2 + d_{\tau_4}$ 和 $AS_{triple} + d_{\tau_5}$ 情况），能够有效消除轨道最后一段的不确定性，但与前述情况不同，无法消除最初几个轨道瓣的不确定性。我们得出结论， d_{τ_4} 和 d_{τ_5} 处的虚拟观测不如 d_{τ_3}提供的信息丰富，这与预期一致。

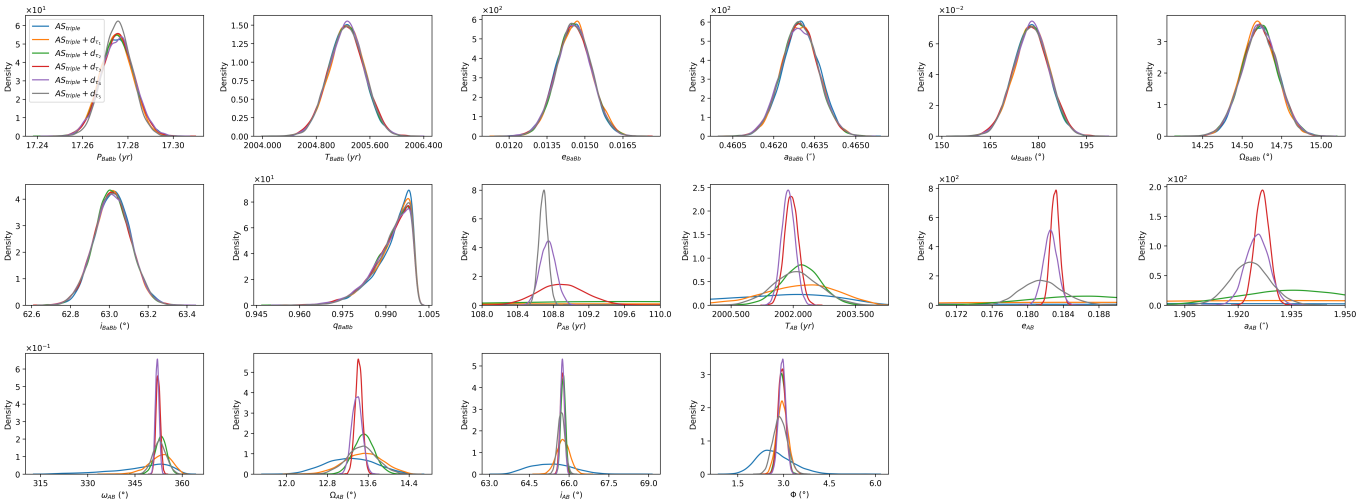

图8。 与图 2 相同，但对象为恒星系统 LHS 1070。由于这是一个三星系统，除经典轨道根数外，底部面板最右侧的图还显示了两个轨道之间的相互倾角（见方程（A13）），研究此类系统时通常会考虑这一参数。

最优观测时间的分布 t^* 位于观测空间不确定性最大的归一化时间（τ_3）的分布如图7下图所示。与此前研究的情形不同，估计得到的分布离散度很高，这反映出由于轨道空间中的位置观测覆盖不足，且缺少径向速度测量，外层系统的周期 P 存在很高的不确定性 AB 。在 P这一高不确定性状态下，我们的归一化方案仍能选出信息量很高的样本：该样本能够显著降低参数的后验熵。这表明，即使周期 P的表征不充分，导致 t^*的选择存在不确定性，我们的方法仍然适用。进而， t^* 的不确定性并不意味着从该分布中选出最可能的候选项对主要估计任务没有帮助。

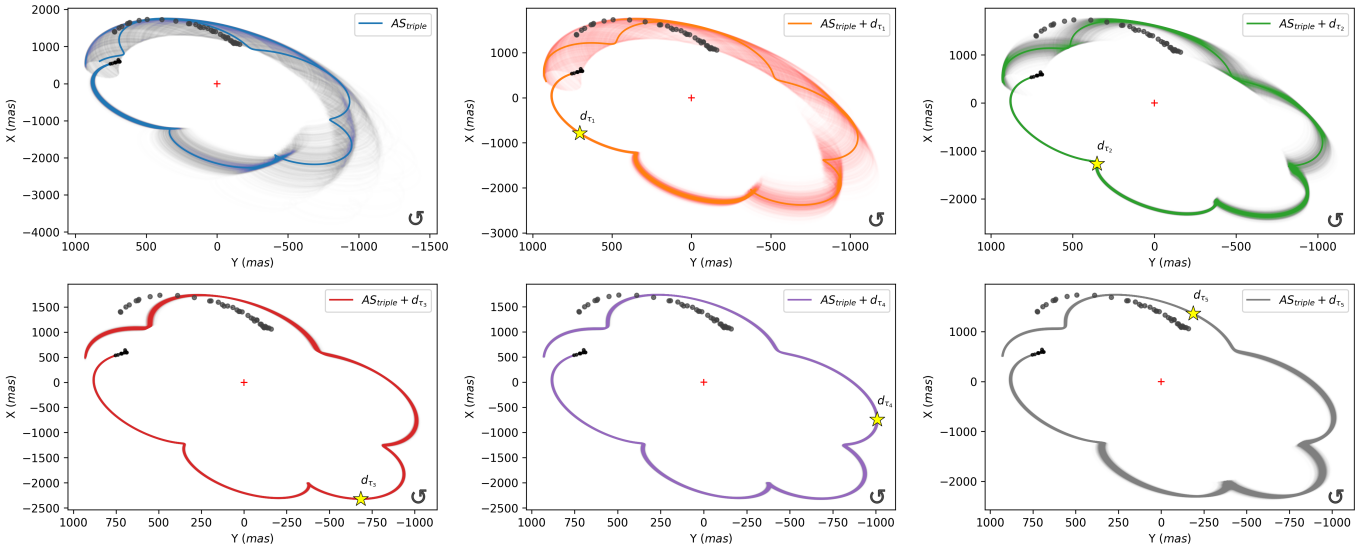

图9。 与图 3 相同，但对象为恒星系统 LHS 1070。轨道参数的数值请参见附录 B中的表 3。轨道相对于测得位置的偏离，是因为这些轨道并非闭合轨道；正如第 4节开头第二段所述，轨道是根据范围 $\tau \in [0, 1]$内的最后一次观测计算得出的。请注意，图 7 （左上面板）中也能看到这一点。

5. 结论与展望

本文介绍了如何将恒星双星和多星系统轨道根数的估计后验分布（首次由Videla 等人（2022年）提出）应用于确定最优观测时间，即最能降低表征系统的轨道参数不确定性的时刻。事实上，Videla 等人（2022年）从贝叶斯视角强调了对双星系统轨道参数后验分布进行完整表征的重要性。这种方法不仅能够进行不确定性量化（许多经典的基于优化的方法只能粗略地做到这一点），还能够直观地呈现系统在观测空间中的不确定性。后一特点可为其他统计任务提供重要输入，例如本文所探讨的最优观测问题。

我们基于最大熵采样原理提出的贝叶斯方法利用了双星系统的周期性，使我们能够全面描述其不确定性随时间的变化。随后，通过纳入一次新的观测，可以确定能够为系统带来最多或最少信息的观测时刻，并由此得到最优观测时间的概率分布。

我们已全面讨论所提出方法的理论基础与实际优势。我们表明，该方法具有较高的数值效率，能够在近乎连续的时间范围内计算信息增益曲线，从而有效刻画系统的不确定性变化。我们将所提出的框架应用于三个不同的恒星系统，展现了该方法在确定最优观测时间方面的适用性与能力。

我们还表明，为恒星系统生成的信息增益曲线使我们能够深入分析系统轨道曲线和径向速度曲线的不确定性。这些曲线为恒星系统提供了有价值且易于解读的概率描述。

如第3.4节所述，我们对系统周期性的归一化时间所作的定义，在研究对象的轨道周期尚未得到准确确定时，确实会造成理论上的不一致。然而，对高周期不确定性天体 LHS 1070 所做的实验表明，即使在上述情况下，我们提出的方法似乎仍然有效。这种归一化的理论影响是一个值得研究的重要课题，我们将在未来的工作中进一步探讨。

MV 和 RAM 感谢 FONDECYT/ANID 第 1190038 号资助。JS 和 MO 感谢 FONDECYT/ANID 第 1210315 号资助，以及先进电气与电子工程中心基础项目 FB0008 的资助。RAM 感谢智利的欧洲南方天文台在其休假期间给予的热情接待；本工作正是在此期间完成的。

我们感谢 Andrei Tokovinin（CTIO/NOIRLab）在 SOAR 上使用 HRCam 时给予的帮助与支撑集，也感谢 Ricardo Salinas（GS/ZORRO 仪器科学家）、Elise Furlan（NASA系外行星科学研究所 Caltech/IPAC 科学家）和 Steve Howell（NASA艾姆斯研究中心空间科学与天体生物学部科学家）在双子座南望远镜上使用 ZORRO 时给予的帮助与支撑集。我们也感谢 CTIO 和 GS 的所有支撑人员，在新冠肺炎疫情期间坚守岗位，保障运行。

本研究使用了由美国海军天文台维护的华盛顿双星星表，以及由斯特拉斯堡 CDS 运营的 SIMBAD数据库。我们非常感谢智利国家观测时间分配委员会通过 SOAR 项目 CN2018A-1、CN2019A-2、CN2019B-13、CN2020A-19、CN2020B-10 和 CN2021B-17 持续提供支撑集，也感谢双子座观测时间分配委员会通过项目编号 GS-2019A- Q-110、GS-2019A-Q-311、GS-2019B-Q-116、GS-2019B-Q-223、GS-2020A-Q-116 和 GS-2020B-Q-142 持续提供支撑集。

本文中的部分观测使用了高分辨率成像仪器 ZORRO。ZORRO 由 NASA 系外行星探索计划资助，并由 Steve B. Howell、Nic Scott、Elliott P. Horch 和 Emmett Quigley 在 NASA艾姆斯研究中心建造。ZORRO 安装在国际双子座天文台的双子座南望远镜上。双子座天文台是美国国家科学基金会 NOIRLab 的一个项目，由大学天文研究协会（AURA）根据与美国国家科学基金会签订的合作协议代表双子座天文台合作伙伴管理。合作伙伴包括：美国国家科学基金会（美国）、国家研究委员会（加拿大）、Agencia Nacional de Investigaci´on y Desarrollo（智利）、Ministerio de Ciencia, Tecnolog´ıa e Innovaci´on（阿根廷）、Minist´erio da Ciˆencia, Tecnologia, Inova¸c˜aes e Comunica¸c˜oes（巴西），以及韩国天文与空间科学研究院（大韩民国）。

设施：**CTIO:SOAR、双子座望远镜**

软件： Stan（Carpenter 等，2017）、NumPy（Harris 等，2020）、Matplotlib（Hunter，2007）、Seaborn（Waskom，2021）、ArviZ（Kumar 等，2019）。

附录

A. 开普勒轨道模型

A.1. 目视光谱双星系统

忽略质量转移、复杂的相对论现象以及其他天体的干扰时，双星系统的轨道由七个轨道参数描述：过近点时刻T、周期 P、轨道偏心率 e、轨道半长轴 a、近心点幅角 ω、升交点经度 Ω，以及轨道倾角 i。双星系统的轨道，即在给定时刻 t 天球平面上的位置（$X(t)$ $Y(t)$），可通过以下步骤计算：

1. 通过数值求解开普勒方程，确定某一历元 t 的所谓偏近点角 $E(t)$：

$$E(t) - e\sin E(t) = 2\pi(t-T)/P. \tag{A1}$$

2. 计算辅助归一化坐标 ($x(t)$ $y(t)$)：

$$\begin{aligned} x(t) &= \cos E(t) - e, \\ y(t) &= \sqrt{1-e^2}\sin E(t). \end{aligned} \tag{A2}$$

3. 确定蒂勒-因尼斯常数：

$$\begin{aligned} A &= a(\cos\omega\cos\Omega - \sin\omega\sin\Omega\cos i), \\ B &= a(\cos\omega\sin\Omega + \sin\omega\cos\Omega\cos i), \\ F &= a(-\sin\omega\cos\Omega - \cos\omega\sin\Omega\cos i), \\ G &= a(-\sin\omega\sin\Omega + \cos\omega\cos\Omega\cos i). \end{aligned} \tag{A3}$$

4. 计算视轨道中的位置（$X(t)$ $Y(t)$）：

$$\begin{aligned} X(t) &= Ax(t) + Fy(t), \\ Y(t) &= Bx(t) + Gy(t). \end{aligned} \tag{A4}$$

要分别计算双星系统中各个组分（主星和伴星）的径向速度（$V_1(t)$ $V_2(t)$），还需引入其他参数：视差 ϖ、各组分的质量比 $q = m_2/m_1$ 以及质心速度 V_0。因此，计算系统中各个组分的径向速度包括以下步骤：

1. 利用方程（ ν（t）确定特定时刻 t 的真近点角 E（t），该真近点角由偏近点角确定：A1）：

$$\tan\frac{\nu(t)}{2} = \sqrt{\frac{1+e}{1-e}}\tan\frac{E(t)}{2}. \tag{A5}$$

2. 计算系统′各个组分的径向速度（$V_1(t)$ $V_2(t)$）：

$$V_1(t) = V_0 + \frac{2\pi a_1\sin i}{P\sqrt{1-e^2}}[\cos(\omega+\nu(t)) + e\cos(\omega)], \tag{A6}$$

$$V_2(t) = V_0 - \frac{2\pi a_2\sin i}{P\sqrt{1-e^2}}[\cos(\omega+\nu(t)) + e\cos(\omega)], \tag{A7}$$

其中 $a_1 = a''/\varpi \cdot q/(1+q)$、 $a_2 = a''/\varpi \cdot 1/(1+q)$以及 a'' 以角秒表示的半长轴。

请注意，方程 (A5) 中真近点角 $\nu(t)$ 的确定不存在歧义，因为该参数与偏近点角 $E(t)$ 的符号相同。此外，径向速度的表达式 (A6) 和 (A7) 显式包含轨道视差 ϖ，旨在展示轨道参数之间完整的相互依存关系，避免将这些表达式中的某些参数压缩为径向速度曲线振幅的独立参数：方程 (A6) 中的 $K_1 = (2\pi a_1 \sin i)/(P\sqrt{1-e^2})$，以及方程 (A7) 中的 $K_2 = (2\pi a_1 \sin i)/(P\sqrt{1-e^2})$，如 Mendez 等（2017）所述。

如果可以获得各组分的径向速度观测（$V_1(t)$ 和 $V_2(t)$）（$SB2$ 以下称为.....），那么描述位置和径向速度观测的联合模型可由轨道参数向量 $\theta_{SB2} = \{P, T, e, a, \omega, \Omega\ i, V_0, \varpi, q\}$表征。然而，如果只能获得一个组分的径向速度观测（$SB1$ 情形），则无法同时确定参数 q 和 ϖ。此时，参数 q 和 ϖ 可合并为辅助参数 f/ϖ（Videla 等（2022）），其中 $f = q/(1+q)$ 是系统的所谓分数质量。因此，描述具有视觉轨道的单线联星的位置和径向速度观测的轨道参数向量为 $\theta_{SB1} = \{P, T, e, a, \omega, \Omega\ i, V_0, f/\varpi\}$。

A.2. 视觉三体层级恒星系统

层级系统近似依次将系统分解为两个恒星子群，其动力学遵循二体问题的解。这会形成由联星子系统构成的层级结构。由于层级结构所引起的相互作用很弱，每个组分的动力学都近似遵循绕系统质心运行的稳定开普勒轨道。

最简单的层级系统由三颗恒星或三个组分构成。由两个双星系统组成的层级结构有两种可能：外层天体绕内双星系统运行，或外层双星系统绕内层天体运行。

考虑第一种层级结构：设 A_a 和 A_b 为内双星系统（A_aA_b）的组分。第三个组分的运动可描述为绕内层系统的质心A运行的开普勒轨道，从而形成外层双星系统 AB。

根据二体问题近似，伴星天体 A_b 相对于主天体 A_a 所受的力为：

$$\ddot{\vec{r}}_{A_aA_b} = \ddot{\vec{r}}_{A_b} - \ddot{\vec{r}}_{A_a} = -\frac{G(m_{A_a} + m_{A_b})}{r^2_{A_aA_b}}\hat{r}_{A_aA_b}. \tag{A8}$$

同样，外部天体 B 相对于双星系统质心 A_aA_b 所受的力由下式给出：

$$\ddot{\vec{r}}_{AB} = \ddot{\vec{r}}_{B} - \ddot{\vec{r}}_{A} = -\frac{G(m_{A} + m_{B})}{r^2_{AB}}\hat{r}_{AB}. \tag{A9}$$

如附录 A.1所述，方程 (A8) 和 (A9) 的解是开普勒椭圆轨道，其特征由轨道参数向量 $\{P_{A_aA_b}, T_{A_aA_b}, e_{A_aA_b}, a_{A_aA_b}, \omega_{A_aA_b}, \Omega_{A_aA_b}, i_{A_aA_b}\}$ 和$\{P_{AB}, T_{AB}, e_{AB}, a_{AB}, \omega_{AB}, \Omega_{AB} i_{AB}\}$分别表示。要确定这些轨道在某一时刻 t 的位置，可按照方程 (A4) 中所述的星历公式进行计算。

内层和外层双星系统的观测通常是相对于系统中较亮且质量最大的分量（主星）测量的。因此，由于内层系统 $\vec{r}_{A_aA_b}$ 的解已是相对于 A_a而言的，因此有必要描述位置向量 $\vec{r}_{A_aB}$，以表征外层双星系统的观测。

矢量 $\vec{r}_{A_aB}$ 可表示为：

$$\vec{r}_{A_aB} = \vec{r}_{A_aA} + \vec{r}_{AB}. \tag{A10}$$

将矢量 $\vec{r}_{A_aA}$ 分解为各分量，并注意到 $\vec{r}_A$ 是内双星系统质心的位置，我们可得：

$$\begin{aligned}\vec{r}_{A_aA} &= \vec{r}_A - \vec{r}_{A_a} \\ &= \frac{m_{A_a}\vec{r}_{A_a} + m_{A_b}\vec{r}_{A_b}}{m_{A_a} + m_{A_b}} - \vec{r}_{A_a} \\ &= (\vec{r}_{A_b} - \vec{r}_{A_a})\frac{m_{A_b}}{m_{A_a} + m_{A_b}} \\ &= \vec{r}_{A_aA_b}\left(\frac{q_{A_aA_b}}{1 + q_{A_aA_b}}\right),\end{aligned} \tag{A11}$$

其中 $q_{A_aA_b} = m_{A_b}/m_{A_a}$ 表示内层系统的质量比。将方程 (A11) 代入方程 (A10) 后，外层天体相对于系统主星的位置为：

$$\vec{r}_{A_aB} = \vec{r}_{AB} + \vec{r}_{A_aA_b}\left(\frac{q_{A_aA_b}}{1 + q_{A_aA_b}}\right), \tag{A12}$$

因此，外层天体相对于内层系统主星的观测可由轨道参数向量 $\{P_{A_aA_b}, T_{A_aA_b}, e_{A_aA_b}, a_{A_aA_b}, \omega_{A_aA_b}, \Omega_{A_aA_b}$表征，

$i_{A_aA_b}\} \cup \{q_{A_aA_b}\} \cup \{P_{AB}, T_{AB}, e_{AB}, a_{AB}, \omega_{AB}, \Omega_{AB}, \ i_{AB}\}$。需要注意的是，虽然位置向量$\vec{r}_{A_aA_b}$ 和 $\vec{r}_{AB}$ 遵循开普勒轨道（因为它们分别是双星系统A_aA_b 和 AB的二体问题的解），但位置向量 $\vec{r}_{A_aB}$ 并非如此，它遵循的是非闭合轨道。一个用于表征内层系统与外层系统之间相互作用的相关派生轨道参数称为相互倾角Φ，其定义如下：

$$\cos(\Phi) = \cos(i_{A_aA_b}) \cdot \cos(i_{AB}) + \sin(i_{A_aA_b}) \cdot \sin(i_{AB}) \cdot \cos(\Omega_{AB} - \Omega_{A_aA_b}). \tag{A13}$$

现在考虑另一种层级结构：令 A 为主星， B_a, B_b 为外层双星系统的组成部分。外层系统质心的运动 B 由绕内层组成部分 A运行的开普勒轨道描述，形成内双星系统 AB。

随后，可以采用类似的方法，将外层双星系统观测表示为位置向量，这些向量是由相应轨道参数所刻画的二体问题的解。因此，描述外层系统观测的位置向量 $\vec{r}_{AB_a}$ 可表示为：

$$\begin{aligned}\vec{r}_{AB_a} &= \vec{r}_{AB} + \vec{r}_{BB_a} \\ &= \vec{r}_{AB} - \vec{r}_{B_aB} \\ &= \vec{r}_{AB} - \vec{r}_{B_aB_b}\left(\frac{q_{B_aB_b}}{1 + q_{B_aB_b}}\right),\end{aligned} \tag{A14}$$

其中，最后一个等号由方程 (A11) 中给出的结果得出。

B. COMPLEMENTARY RESULTS

参考文献

Balsa-Canto, E., Alonso, A. A., & Banga, J. R. 2008, 《IET系统生物学》, 2, 163 Bernardo, J. M. 1979, 《统计年鉴》, 686 Box, G. E. 1982, 载于《.....陆军研究与测试中的实验设计会议论文集》, 第28卷, 美国陆军研究办公室, 238 Carpenter, B., Gelman, A., Hoffman, M. D., et al. 2017, 《统计软件杂志》, 76, 1

Chaloner, K., & Verdinelli, I. 1995, 《统计科学》, 273 Chernoff, H. 1953, 《数学统计年鉴》, 586 Cover, T. M. 1999, 《信息论要素》（约翰·威立父子公司）Dehghannasiri, R., Xue, D.,Balachandran, P. V.,等 2017, 《计算材料科学》, 129, 311

表1. 纳入虚拟观测 $d_{\tau_i} \sim \mathcal{N}(f_{\rm kep}(\theta_{map}, \tau_i)\ \sigma^2)\ i \in \{1,2,3,4,5\}$ 后，双星 HIP 89000 的轨道根数和导出物理参数的边缘后验分布的最大后验估计和 95%最高后验密度区间，其中 σ^2 固定为系统现有观测 $\mathcal{D}$ 的最小方差。

θ	$SB2$	$SB2+d_{\tau_1}$	$SB2+d_{\tau_2}$	$SB2+d_{\tau_3}$	$SB2+d_{\tau_4}$	$SB2+d_{\tau_5}$
$P\ (yr)$	$0.546^{0.547}_{0.546}$	$0.546^{0.547}_{0.546}$	$0.546^{0.547}_{0.546}$	$0.546^{0.547}_{0.546}$	$0.546^{0.547}_{0.546}$	$0.546^{0.547}_{0.546}$
$T\ (yr)$	$1990.675^{1990.677}_{1990.673}$	$1990.675^{1990.677}_{1990.673}$	$1990.675^{1990.677}_{1990.673}$	$1990.675^{1990.677}_{1990.673}$	$1990.675^{1990.677}_{1990.673}$	$1990.675^{1990.677}_{1990.673}$
e	$0.302^{0.306}_{0.297}$	$0.302^{0.306}_{0.297}$	$0.302^{0.307}_{0.297}$	$0.302^{0.306}_{0.298}$	$0.302^{0.306}_{0.297}$	$0.302^{0.306}_{0.297}$
$a\ ('')$	$0.021^{0.023}_{0.017}$	$0.021^{0.023}_{0.018}$	$0.021^{0.023}_{0.019}$	$0.021^{0.023}_{0.018}$	$0.021^{0.022}_{0.019}$	$0.021^{0.023}_{0.019}$
$\omega\ (°)$	$86.775^{87.441}_{85.739}$	$86.637^{87.473}_{85.793}$	$86.736^{87.460}_{85.776}$	$86.715^{87.491}_{85.878}$	$86.621^{87.450}_{85.756}$	$86.484^{87.517}_{85.806}$
$\Omega\ (°)$	$49.611^{54.205}_{45.878}$	$50.965^{54.065}_{45.904}$	$49.904^{54.080}_{45.660}$	$49.413^{53.735}_{46.258}$	$49.872^{54.147}_{46.138}$	$49.449^{54.198}_{46.025}$
$i\ (°)$	$141.080^{157.739}_{133.203}$	$140.322^{155.190}_{133.265}$	$141.324^{148.851}_{134.968}$	$141.579^{157.203}_{133.345}$	$142.721^{148.903}_{135.587}$	$141.400^{150.454}_{135.057}$
$V_0\ (km/s)$	$-14.124^{-14.091}_{-14.174}$	$-14.130^{-14.087}_{-14.173}$	$-14.145^{-14.086}_{-14.172}$	$-14.119^{-14.089}_{-14.174}$	$-14.140^{-14.088}_{-14.173}$	$-14.137^{-14.089}_{-14.174}$
$\varpi\ (mas)$	$26.212^{33.052}_{12.424}$	$26.846^{32.800}_{14.126}$	$25.870^{31.559}_{19.709}$	$25.658^{33.083}_{13.108}$	$24.698^{30.902}_{19.557}$	$25.494^{31.334}_{18.180}$
$m_1\ (M_\odot)$	$0.894^{3.423}_{0.482}$	$0.854^{2.706}_{0.518}$	$0.909^{1.530}_{0.592}$	$0.925^{3.268}_{0.510}$	$1.004^{1.529}_{0.610}$	$0.919^{1.754}_{0.595}$
$m_2\ (M_\odot)$	$0.855^{3.287}_{0.473}$	$0.817^{2.592}_{0.498}$	$0.870^{1.460}_{0.563}$	$0.885^{3.109}_{0.476}$	$0.960^{1.461}_{0.579}$	$0.878^{1.666}_{0.560}$
q	$0.956^{0.963}_{0.950}$	$0.956^{0.963}_{0.950}$	$0.956^{0.963}_{0.949}$	$0.957^{0.963}_{0.949}$	$0.956^{0.963}_{0.949}$	$0.956^{0.963}_{0.950}$

表2. 纳入虚拟观测 $d_{\tau_i} \sim \mathcal{N}(f_{\rm kep}(\theta_{map}, \tau_i)\ \sigma^2)\ i \in \{1,2,3,4,5\}$ 后，双星 HIP 99675 的轨道根数和导出物理参数的边缘后验分布的最大后验估计和95%最高后验密度区间，其中 σ^2 固定为系统可用观测 $\mathcal{D}$ 的最小方差。

θ	$SB1$	$SB1+d_{\tau_1}$	$SB1+d_{\tau_2}$	$SB1+d_{\tau_3}$	$SB1+d_{\tau_4}$	$SB1+d_{\tau_5}$
$P\ (yr)$	$10.297^{10.332}_{10.265}$	$10.296^{10.333}_{10.266}$	$10.300^{10.332}_{10.263}$	$10.311^{10.332}_{10.264}$	$10.297^{10.333}_{10.266}$	$10.300^{10.333}_{10.265}$
$T\ (yr)$	$1950.528^{1950.675}_{1950.372}$	$1950.540^{1950.674}_{1950.374}$	$1950.526^{1950.671}_{1950.370}$	$1950.485^{1950.675}_{1950.366}$	$1950.540^{1950.676}_{1950.376}$	$1950.518^{1950.674}_{1950.370}$
e	$0.202^{0.210}_{0.195}$	$0.202^{0.210}_{0.195}$	$0.203^{0.211}_{0.195}$	$0.203^{0.210}_{0.195}$	$0.202^{0.210}_{0.195}$	$0.202^{0.210}_{0.195}$
$a\ ('')$	$0.031^{0.036}_{0.025}$	$0.031^{0.032}_{0.030}$	$0.031^{0.034}_{0.028}$	$0.032^{0.033}_{0.030}$	$0.032^{0.036}_{0.027}$	$0.032^{0.033}_{0.029}$
$\omega\ (°)$	$198.678^{201.382}_{196.574}$	$199.033^{201.309}_{196.532}$	$199.001^{201.194}_{196.558}$	$199.141^{201.244}_{196.539}$	$199.217^{201.293}_{196.608}$	$198.653^{201.341}_{196.672}$
$\Omega\ (°)$	$305.271^{317.330}_{294.422}$	$304.496^{313.434}_{298.794}$	$304.182^{313.971}_{294.971}$	$305.464^{307.996}_{302.494}$	$305.521^{311.866}_{299.980}$	$307.406^{313.424}_{295.709}$
$i\ (°)$	$110.204^{132.995}_{79.770}$	$108.993^{127.468}_{97.224}$	$111.385^{121.306}_{102.014}$	$106.681^{131.474}_{83.162}$	$109.769^{114.951}_{107.044}$	$106.953^{126.627}_{98.631}$
$V_0\ (km/s)$	$-6.390^{-6.284}_{-6.470}$	$-6.389^{-6.291}_{-6.474}$	$-6.377^{-6.278}_{-6.467}$	$-6.355^{-6.286}_{-6.470}$	$-6.381^{-6.288}_{-6.472}$	$-6.383^{-6.289}_{-6.477}$
$f/\varpi\ (pc)$	$161.221^{199.983}_{142.424}$	$161.508^{192.703}_{147.767}$	$162.614^{193.141}_{143.082}$	$158.391^{196.190}_{147.878}$	$160.074^{196.435}_{140.010}$	$158.118^{194.824}_{147.137}$

Ford, E. B. 2004, 《太平洋天文学会刊》, 116, 1083 —. 2008, 《天文学杂志》, 135, 1008 Geman, S., & Geman,D. 1984, IEEE模式分析与机器智能汇刊, 721 Goodman, J., & Weare, J. 2010, 《应用数学与计算科学通讯》, 5, 65 Gregory, P. C. 2005, 《天体物理学杂志》, 631, 1198, doi: 10.1086/432594—. 2011, 《皇家天文学会月报》, 410, 94, doi: 10.1111/j.1365-2966.2010.17428.xHarris, C. R., Millman,K. J.,van der Walt, S. J., 等 2020, 《自然》, 585, 357 Hastings, W. K. 1970 Hees, A., Dehghanfar, A.,Do, T., 等 2019, 《天体物理学杂志》, 880, 87 Hoffman, M. D., Gelman,A., 等 2014, 《机器学习研究杂志》, 15, 1593 Hou, F., Goodman, J., Hogg, D. W., Weare, J., & Schwab, C. 2012, 《天体物理学杂志》, 745, 198, doi: 10.1088/0004-637X/745/2/198Hunter, J. D. 2007, 《科学与工程计算》, 9, 90 James, O. B. 1985, 施普林格统计学系列, ISBN-10: 0-387-96098-8 和 -13, 978 Kozachenko, L., & Leonenko, N. N. 1987, 《信息传输问题》, 23, 9 Kumar,R., Carroll, C., Hartikainen, A., & Martin, O.2019, 《开放源代码软件期刊》, 4, 1143, doi: 10.21105/joss.01143Lindley, D. V. 1956, 《数理统计年刊》, 27, 986 —。1972, 贝叶斯统计：综述（工业与应用数学学会） Liu, D. C., & Nocedal, J. 1989, 《数学规划》, 45, 503

表3. 将虚拟观测（(,)）纳入后，层级三体系统 LHS 1070 $d_{\tau_i} \sim \mathcal{N}\ f_{\rm kep}\ \theta_{map}\ \tau_i\ \ \sigma^2\ \ i \in$ 的 $\{1,2,3,4,5\}$ σ^2 $\mathcal{D}$ 轨道根数和导出物理参数的边缘后验分布的最大后验估计和95%最高后验密度区间（HDPIs），其中固定为可用系统观测的最小方差。

θ	AS_{triple}	$AS_{triple}+d_{\tau_1}$	$AS_{triple}+d_{\tau_2}$	$AS_{triple}+d_{\tau_3}$	$AS_{triple}+d_{\tau_4}$	$AS_{triple}+d_{\tau_5}$
P_{BaBb} (yr)	$17.282^{17.290}_{17.261}$	$17.280^{17.289}_{17.261}$	$17.270^{17.289}_{17.261}$	$17.274^{17.290}_{17.262}$	$17.274^{17.290}_{17.261}$	$17.276^{17.288}_{17.263}$
T_{BaBb} (yr)	$2005.081^{2005.752}_{2004.754}$	$2005.180^{2005.778}_{2004.749}$	$2005.174^{2005.754}_{2004.725}$	$2005.271^{2005.777}_{2004.731}$	$2005.130^{2005.754}_{2004.738}$	$2005.085^{2005.753}_{2004.750}$
e_{BaBb}	$0.014^{0.016}_{0.013}$	$0.015^{0.016}_{0.013}$	$0.014^{0.016}_{0.013}$	$0.015^{0.016}_{0.013}$	$0.015^{0.016}_{0.013}$	$0.015^{0.016}_{0.013}$
a_{BaBb} $('')$	$0.463^{0.464}_{0.462}$	$0.463^{0.464}_{0.462}$	$0.463^{0.464}_{0.462}$	$0.463^{0.464}_{0.462}$	$0.463^{0.464}_{0.462}$	$0.463^{0.464}_{0.462}$
ω_{BaBb} (°)	$174.313^{188.218}_{167.394}$	$176.356^{188.905}_{167.359}$	$176.142^{188.365}_{166.909}$	$178.222^{188.781}_{166.951}$	$175.308^{188.416}_{167.223}$	$174.283^{188.082}_{167.142}$
Ω_{BaBb} (°)	$14.608^{14.842}_{14.404}$	$14.612^{14.823}_{14.393}$	$14.695^{14.841}_{14.389}$	$14.626^{14.832}_{14.390}$	$14.572^{14.830}_{14.382}$	$14.666^{14.834}_{14.394}$
i_{BaBb} (°)	$62.962^{63.199}_{62.844}$	$62.975^{63.178}_{62.820}$	$62.964^{63.200}_{62.840}$	$63.024^{63.196}_{62.832}$	$63.063^{63.189}_{62.828}$	$63.001^{63.196}_{62.837}$
q_{BaBb}	$0.988^{1.000}_{0.979}$	$0.991^{1.000}_{0.979}$	$0.989^{1.000}_{0.979}$	$0.993^{1.000}_{0.978}$	$0.990^{1.000}_{0.979}$	$0.994^{1.000}_{0.979}$
P_{AB} (yr)	$105.414^{124.480}_{88.241}$	$110.701^{118.228}_{101.827}$	$109.692^{112.803}_{106.627}$	$108.935^{109.430}_{108.388}$	$108.699^{108.907}_{108.560}$	$108.693^{108.790}_{108.592}$
T_{AB} (yr)	$2001.345^{2003.964}_{1996.313}$	$2002.556^{2003.864}_{2000.305}$	$2002.231^{2003.109}_{2001.257}$	$2001.952^{2002.292}_{2001.636}$	$2001.803^{2002.204}_{2001.574}$	$2002.162^{2003.167}_{2000.923}$
e_{AB}	$0.166^{0.259}_{0.071}$	$0.191^{0.227}_{0.150}$	$0.187^{0.200}_{0.175}$	$0.183^{0.184}_{0.182}$	$0.183^{0.184}_{0.181}$	$0.181^{0.186}_{0.177}$
a_{AB} $('')$	$1.886^{2.117}_{1.665}$	$1.947^{2.036}_{1.843}$	$1.935^{1.968}_{1.907}$	$1.927^{1.931}_{1.923}$	$1.926^{1.932}_{1.919}$	$1.921^{1.934}_{1.913}$
ω_{AB} (°)	$349.818^{359.905}_{328.230}$	$354.614^{359.512}_{345.724}$	$353.399^{356.914}_{349.646}$	$352.330^{353.721}_{350.987}$	$351.754^{353.264}_{350.899}$	$353.111^{357.096}_{348.703}$
Ω_{AB} (°)	$13.263^{14.166}_{12.381}$	$13.699^{14.249}_{12.789}$	$13.535^{13.921}_{13.131}$	$13.409^{13.561}_{13.281}$	$13.334^{13.572}_{13.180}$	$13.530^{14.017}_{12.868}$
i_{AB} (°)	$65.577^{66.793}_{63.908}$	$65.800^{66.258}_{65.277}$	$65.753^{65.949}_{65.615}$	$65.745^{65.909}_{65.586}$	$65.753^{65.898}_{65.605}$	$65.652^{65.979}_{65.442}$
Φ (°)	$2.882^{3.864}_{1.646}$	$2.943^{3.325}_{2.599}$	$2.979^{3.200}_{2.684}$	$2.933^{3.178}_{2.694}$	$2.913^{3.186}_{2.737}$	$2.842^{3.332}_{2.446}$

Loredo, T. J., Berger, J. O.,Chernoff, D. F., Clyde, M. A., & Liu, B. 2012，统计方法学，9，101 Lucy, L. 2014，天文学与天体物理学，563，A126 Mendez, R. A., Claveria, R. M., Orchard, M. E., & Silva, J. F. 2017，《天文学杂志》，154，187，doi：10.3847/1538-3881/aa8d6fMuller, P., & Parmigiani, G.1996，贝叶斯统计与计量经济学：向 A. Zellner 致敬的论文集，397 Myung, J. I., & Pitt, M.A. 2009，《心理学评论》，116，499 Neal, R. M., et al. 2011，《马尔可夫链蒙特卡罗手册》，2，2 Nelson, B., Ford, E. B.,& Payne, M. J. 2014，《天体物理学杂志增刊》，210，11，doi：10.1088/0067-0049/210/1/11Parzen, E. 1962，《数理统计年刊》，33，1065 Patat, F., & Hussain, G. 2012，《信使》，150，17 Pukelsheim, F. 2006，《实验的最优设计》（工业与应用数学学会）

Sebastiani, P., & Wynn, H. P. 2000，《皇家统计学会会刊：B辑（统计方法学）》，62，145 Shewry, M. C., & Wynn, H. P. 1987，《应用统计学杂志》，14，165 Sozzetti, A., Casertano, S., Brown, R., & Lattanzi, M. 2002，《太平洋天文学会刊》，114，1173 Videla, M., Mendez,R. A., Claver´ıa, R. M.,Silva, J. F., & Orchard, M. E. 2022年，《天文学杂志》，163，220，doi：10.3847/1538-3881/ac5ab4Villegas, C., Mendez,R. A., Silva, J. F., & Orchard, M. E. 2021年，《太平洋天文学会刊》，133，074501，doi：10.1088/1538-3873/ac0239 Waskom, M. L. 2021年，《开放源代码软件期刊》，6，3021，doi：10.21105/joss.03021