Improving Orbit Estimates for Incomplete Orbits with a New Approach to Priors: with Applications from Black Holes to Planets

K. Kosmo O'Neil[1] , G. D. Martinez[1], A. Hees[1,2] , A. M. Ghez[1] , T. Do[1] , G. Witzel[1] , Q. Konopacky[3] , E. E. Becklin[1], D. S. Chu[1] , J. R. Lu[4] , K. Matthews[5], and S. Sakai[1]

[1] Department of Physics and Astronomy, University of California, Los Angeles, Los Angeles, CA 90095, USA
[2] SYRTE, Observatoire de Paris, Université PSL, CNRS, Sorbonne Université, LNE, 61 avenue de l'Observatoire F-75014 Paris, France
[3] Center for Astrophysics and Space Sciences, University of California, San Diego, La Jolla, CA 92093, USA
[4] Department of Astronomy, University of California, Berkeley, Berkeley, CA 94720, USA
[5] Division of Physics, Mathematics, and Astronomy, California Institute of Technology, Pasadena, CA 91125, USA

Abstract

We propose a new approach to Bayesian prior probability distributions (priors) that can improve orbital solutions for low-phase-coverage orbits, where data cover less than ∼40% of an orbit. In instances of low phase coverage—such as with stellar orbits in the Galactic center or with directly imaged exoplanets—data have low constraining power and thus priors can bias parameter estimates and produce underestimated confidence intervals. Uniform priors, which are commonly assumed in orbit fitting, are notorious for this. We propose a new observable-based prior paradigm that is based on uniformity in observables. We compare performance of this observable-based prior and of commonly assumed uniform priors using Galactic center and directly imaged exoplanet (HR 8799) data. The observable-based prior can reduce biases in model parameters by a factor of two and helps avoid underestimation of confidence intervals for simulations with less than ∼40% phase coverage. Above this threshold, orbital solutions for objects with sufficient phase coverage—such as S0-2, a short-period star at the Galactic center with full phase coverage—are consistent with previously published results. Below this threshold, the observable-based prior limits prior influence in regions of prior dominance and increases data influence. Using the observable-based prior, HR 8799 orbital analyses favor low-eccentricity orbits and provide stronger evidence that the four planets have a consistent inclination of ∼30° to within 1σ. This analysis also allows for the possibility of coplanarity. We present metrics to quantify improvements in orbital estimates with different priors so that observable-based prior frameworks can be tested and implemented for other low-phase-coverage orbits.

Key words: Galaxy: center – Galaxy: fundamental parameters – methods: statistical – planets and satellites: fundamental parameters

Supporting material: tar.gz file

1. Introduction

The advent of adaptive optics (AO) has enabled orbital parameter estimation for individual objects in several new and exciting astrophysical systems, such as the Galactic nuclear star cluster (NSC) and directly imaged exoplanet systems. At the center of the Milky Way, the orbit of S0-2 (period = 16 yr) has provided the best evidence to date for the existence of supermassive black holes, and has now begun to constrain alternative theories of gravity (Ghez et al. 2000, 2003, 2005, 2008; Eckart et al. 2002; Schödel et al. 2002, 2003; Eisenhauer et al. 2003; Gillessen et al. 2009, 2017; Genzel et al. 2010; Meyer et al. 2012; Boehle et al. 2016; Grould et al. 2017; Hees et al. 2017; Parsa et al. 2017). Over time, it has become possible to measure stellar orbits at larger Galacto-centric radii and thereby study the dynamical structure of the Galactic NSC—the only such system for which orbital studies of central black holes and their host galaxies are possible (e.g., Ghez et al. 2005; Gillessen et al. 2017). Similarly, AO has opened up a new window in the field of exoplanets. Orbital motions of exoplanets imaged with AO have enabled the first studies of the 3D dynamical structures of exoplanet systems, which can give insights into the formation and evolution of giant planets (e.g., Marois et al. 2008, 2010; Currie et al. 2011, 2012; Chauvin et al. 2012; Esposito et al. 2013; Maire et al. 2015; Pueyo et al. 2015; Konopacky et al. 2016;

Rameau et al. 2016; Zurlo et al. 2016; Wertz et al. 2017). However, the majority of orbital measurements at the Galactic center (GC) and of directly imaged exoplanets have incomplete orbital phase coverage.

It is difficult to infer accurate orbital estimates for objects with low orbital phase coverage. In these cases, Bayesian prior probability distributions (priors) can easily dominate parameter estimates. In orbit analyses, uniform priors in model parameters are commonly assumed (e.g., Ghez et al. 2008; Boehle et al. 2016; Gillessen et al. 2017); however, uniform priors are subjective, as they depend on a choice of the parameterization of the model. Other prior forms that still depend on model parameters have also been used, including isotropic orientations (as suggested by Ford 2006) and uniformity in the Thiele-Innes elements (Lucy 2014). However, Lucy (2014) showed that, when data are not rigorously constraining (e.g., when observations sweep out an angle that is less than ∼40% of the projected 360° orbit), such subjective model-based priors can lead to biases in parameter estimates and produce inaccurate confidence intervals.

We aim to develop a prior that is less subjective (and thus more objective) than what has previously been used in orbit modeling. While the concept of using an objective prior has not been deeply explored in orbit modeling, there is a rich literature in statistics in which objective-prior frameworks have been developed. Although priors cannot be truly objective, because any probability distribution contains *some* information, objective-prior frameworks aim to

minimize prior influence and thus maximize data influence on parameter inference. The Jeffreys prior (Jeffreys 1946), based on the Fisher information content, is one such paradigm that is designed to be invariant under parameterization. However, in multiple dimensions, the Jeffreys prior has been shown to be inconsistent with other objective prior frameworks, such as reference analysis or location-scale invariance (Bernardo 2005). For a prior to be considered objective, it must inherently sample relevant regions of parameter space (Jeffreys 1946; Neyman & Scott 1948; Hartigan 1964; Jaynes 1968; Bernardo 2005; Berger et al. 2009; Pierini et al. 2011). In other words, it requires consistent sampling of regions of parameter space favored by the data, which ideally would ensure unbiased parameter estimates and accurate confidence intervals. Reference analysis (Berger et al. 2009) accomplishes this by maximizing the relative expected information between posterior and prior distributions in the asymptotic limit of large data sets. However, most real-life experiments only produce a subset of this asymptotically large data set, and thus can have properties that differ from those of the larger data set. Consequently, priors based on this asymptotic assumption (e.g., those from reference analysis) can lead to unintended statistical consequences, such as misleading inferences or biases (Gelman et al. 2017). In addition, such analyses are difficult to calculate, except in simple examples (Pierini et al. 2011).

In this paper, we propose a new approach to priors that can improve orbital estimates for low-phase-coverage orbits. We refer to these new priors as *observable-based priors* because they are based on uniformity in the observables rather than in model parameters. This prior requires all measurements to be equally likely before observation, and thus it promotes consistent sampling, limits prior influence in regions of prior dominance, and increases data influence.

We present the general definition of our observable-based prior in Section 2.1, and describe the specific form of the prior for orbit-modeling applications in Section 2.2. We describe the observational data used for this analysis in Section 3, and use simulations based on these data to test the performance of the observable-based prior compared to that of standard uniform priors in Section 4. We then present updated orbital analyses of S0-2 and the HR 8799 planets in Section 5, and discuss the scientific impact in Section 6.

2. Observable-based Priors

2.1. General Form

We construct a prior that is as objective as possible—in order to promote unbiased parameter estimates and accurate confidence intervals—but that is not based on a hypothetically large data set like in reference analysis (Berger et al. 2009). Defining priors with respect to quantities that can be physically measured ensures a priori that the underlying assumptions (likelihood and priors) make physical sense and that possible observations are not biased before observing data. In other words, a prior should neither de-emphasize areas of parameter space that could be observed, nor emphasize areas of parameter space that cannot be observed. There should theoretically be an equal probability of obtaining observations in all regions of parameter space that are possible to observe, which motivates the construction of a prior based on uniformity in the observables rather than on model parameters.

Observable-based priors assume a prior in observable space, $\mathcal{P}(\mathcal{O})$, which can be transformed into a prior in model-parameter space, $\mathcal{P}(\mathcal{M})$, by inverting the integral

$$\mathcal{P}(\mathcal{O}) = \int d\mathcal{M}\,\delta(\mathcal{O} - f_{\mathcal{O}}(\mathcal{M}))\mathcal{P}(\mathcal{M}) \tag{1}$$

to solve for $\mathcal{P}(\mathcal{M})$. Here, $\mathcal{O} = f_{\mathcal{O}}(\mathcal{M})$ is the definition of the set of observables and $\mathcal{M}$ is the set of model parameters. Observable-based priors require that external prior knowledge and knowledge of the experimental design are encapsulated in $\mathcal{P}(\mathcal{O})$ rather than in $\mathcal{P}(\mathcal{M})$. In this paper, we assume a one-to-one transformation for simplicity. For problems where $\mathcal{O}$ and $\mathcal{M}$ have the same dimensionality, inverting Equation (1) gives:

$$\mathcal{P}(\mathcal{M}) = \mathcal{P}(\mathcal{O}) \left| \frac{\partial(\mathcal{O})}{\partial(\mathcal{M})} \right|. \tag{2}$$

Here, $\partial(\mathcal{O})/\partial(\mathcal{M})$ is the Jacobian defined by the transformation from observable space to model parameter space. If the dimensionality of $\mathcal{M}$ is larger than that of $\mathcal{O}$, then extra constraints must be introduced to produce an unique prior. Alternatively, if the dimensionality of $\mathcal{O}$ is larger, then the observables themselves are not independent of each other. In general, measurements are not independent and identically distributed for orbit analyses, and thus the number of observables increases with the number of measurements. Each observable depends on the epoch of observation, making the distributions $\mathcal{P}(\mathcal{O})$ and $\mathcal{P}(\mathcal{M})$ dependent on the epoch t (e.g., $\mathcal{P}(\mathcal{O}|t)$ and $\mathcal{P}(\mathcal{M}|t)$). This dependence makes it impossible to uniquely specify $\mathcal{P}(\mathcal{O}|t)$ and $\mathcal{P}(\mathcal{M}|t)$ at every observed epoch. Instead, we marginalize over a PDF of the observing schedule, $\mathscr{P}(t)$, to specify a marginal distribution $\mathcal{P}(\mathcal{M})$. This marginal distribution can be approximated as a sum over the epochs of observation:

$$\mathcal{P}(\mathcal{M}) = \int \mathcal{P}(\mathcal{M}|t) P(t) dt \tag{3}$$

$$\approx \sum_j \mathcal{P}(\mathcal{M}|t_j). \tag{4}$$

We define our prior such that the probability of a set of observables $\mathcal{O}$ at time t, $\mathcal{P}(\mathcal{O}|t)$, is uniform within a range that is proportional to the measurement error. Because the likelihood is invariant under coordinate translation, a flat prior distribution in observable space ensures that the posterior is also invariant under the same assumptions (Jaynes 1968). Observable-based priors of other forms in observable coordinates can also be assumed, depending on the form of the likelihood. Assuming uniformity in observable space, Equation (2) conditioned on time t becomes:

$$\mathcal{P}(\mathcal{M}|t) = \mathcal{P}(\mathcal{O}|t) \left| \frac{\partial(\mathcal{O}(t))}{\partial(\mathcal{M})} \right| \tag{5}$$

$$\propto \frac{1}{\prod_i \sigma_i(t)} \left| \frac{\partial(\mathcal{O}(t))}{\partial(\mathcal{M})} \right|, \tag{6}$$

where $\prod_i \sigma_i(t)$ is the product of the measurement uncertainties for the set of observables at time t. That the specified form of the prior is inversely proportional to the measurement uncertainties implies that, in practice, the prior assigns different weight to information on observables carried by each measurement. In other words,

more emphasis is placed on data points with smaller measurement uncertainties. Note that the above prior is identical to the Jeffreys prior (Jeffreys 1946) for a single epoch, though this similarity ends for data sets that span multiple epochs.[6] Under the above assumptions, it follows from Equations (4) and (6) that the observable-based prior in model-parameter space becomes:

$$\mathcal{P}(\mathcal{M}) \approx \sum_j \mathcal{P}(\mathcal{M}|t_j) \tag{8}$$

$$\propto \sum_j \frac{1}{\prod_i \sigma_i(t_j)} \left| \frac{\partial(\mathcal{O}(t_j))}{\partial(\mathcal{M})} \right|. \tag{9}$$

This observable-based prior is based on experimental design, though *not* on actual observed data. This idea has precedence in Bayesian statistics, as the Jeffreys prior is similarly based on a form of the likelihood. While our prior relates to the Jeffreys prior for a single epoch, it has different philosophical motivation. Rather than being based on the abstract concept of information content, our observable-based prior is based on the practical idea that there should be an equal probability of obtaining observations in regions of parameter-space that are possible to observe. In addition, there should be no probability of obtaining observations in regions of parameter space that are not possible to observe, thereby avoiding the pitfalls of asymptotic-based priors (Gelman et al. 2017). Thus, the observable-based prior is less subjective than commonly assumed model-based priors, because it does not directly impose constraints on model parameters. However, it is weakly informative rather than truly objective, in the sense that it imposes constraints on possible observed data.

2.2. Observable-based Priors in Orbit Modeling

We construct an observable-based prior, of the general form defined in Section 2.1, specific to a Keplerian point-potential orbit model in which the central mass, M_{cent}, dominates. In future work, we will extend the model to include post-Newtonian parameters in General Relativity, such as the relativistic redshift. In the Keplerian model described here, seven global parameters (that describe the gravitational potential) and six Keplerian orbital parameters comprise the model $\mathcal{M}$. The seven global parameters are the central mass (M_{cent}), the line-of-sight distance to the primary (R_o), its position on the sky (x_o, y_o), and its three-dimensional velocity ($v_{x,o}, v_{y,o}, v_{z,o}$). The six Keplerian orbital elements, further detailed by Ghez et al. (2005), are the angle to the ascending node (Ω), argument of periapse (ω), inclination (i), orbital period (P), time of closest approach (T_o), and eccentricity (e). The model $\mathcal{M}$ is thus defined as:

$$\mathcal{M} = \{M_{cent}, R_o, x_o, y_o, v_{x,o}, v_{y,o}, v_{z,o}, \Omega, \omega, i, P, T_o, e\}. \tag{10}$$

Observational data $\mathcal{D}$ constrain the set of model parameters $\mathcal{M}$. For data with only astrometric observations (e.g., directly

imaged planets), the set of measured observables $\mathcal{D}$ at time t is:

$$\mathcal{D} = \{x(t), y(t)\}, \tag{11}$$

where x and y describe the orbiting body's position on the sky with respect to the position of the primary, x_o and y_o. For data with both astrometric and RV observations (e.g., GC data),

$$\mathcal{D} = \{x(t), y(t), v_z(t)\}, \tag{12}$$

where v_z is the orbiting body's line-of-sight velocity. Note that x, y, and v_z refer to the position and velocity of the orbiting body, while x_o, y_o, $v_{x,o}$, $v_{y,o}$, and $v_{z,o}$ describe the position and velocity of the primary (in the GC case, the SMBH). Relative positions are characterized by the angular separation between the primary and the orbiting body, projected onto the plane of the sky. While the position of the primary can correspond to a particular R.A. and decl., relative R.A. and decl. coordinates would depend on an object's location on the curved coordinate system, requiring an additional transformation. In the small-angle approximation, there is effectively no curvature in the coordinate system—and thus there is approximately a linear relationship between (Δx, Δy) and (ΔR.A., Δdecl.). In such cases, R.A. and decl. or any other linearly related coordinate system could be used. As long as the transformation between coordinate systems is linear, a uniform distribution in one implies a uniform distribution in the other. In this work, relative astrometric positions, defined as $\mathcal{D} = \{x(t), y(t)\}$, are used because a linear transformation converts measurements from pixel space to positions on the sky (defined with respect to the position of the primary). Relative astrometry is commonly measured in this way (e.g., Konopacky et al. 2016; Gillessen et al. 2017). Further, a functional form exists for $x(t)$ and $y(t)$ such that a prior can easily be formulated.

The *measured* observables $\mathcal{D}$ are related linearly to a set of *orbital* observables $X(E)$, $Y(E)$, $V_X(E)$, and $V_Y(E)$ that describe the position and velocity in the orbital plane by

$$x(t) = \frac{B}{R_o}X(E) + \frac{G}{R_o}Y(E) + v_{x,o}t + x_o \tag{13}$$

$$y(t) = \frac{A}{R_o}X(E) + \frac{F}{R_o}Y(E) + v_{y,o}t + y_o \tag{14}$$

$$v_z(t) = CV_X(E) + HV_Y(E) + v_{z,o}, \tag{15}$$

where the orbital observables $X(E)$, $Y(E)$, $V_X(E)$, and $V_Y(E)$ (e.g., Ghez et al. 2003) are defined as

$$X = a(\cos E - e), \tag{16}$$

$$Y = a(\sqrt{1 - e^2} \sin E), \tag{17}$$

$$V_X = -\frac{\sin E}{1 - e \cos E}\sqrt{\frac{GM}{a}}, \tag{18}$$

$$V_Y = \frac{\sqrt{1 - e^2} \cos E}{1 - e \cos E}\sqrt{\frac{GM}{a}}, \tag{19}$$

and A, B, C, F, G, and H are the Thiele-Innes constants (e.g., Hartkopf et al. 1989; Wright & Howard 2009), defined by

$$A = +\cos\Omega\cos\omega - \sin\Omega\sin\omega\cos i \tag{20}$$

$$B = +\sin\Omega\cos\omega + \cos\Omega\sin\omega\cos i \tag{21}$$

$$C = \sin\omega\sin i \tag{22}$$

$$F = -\cos\Omega\sin\omega - \sin\Omega\cos\omega\cos i \tag{23}$$

---

[6] The Jeffreys prior for multiple epochs is:

$$\mathcal{P}^{\mathcal{J}} = \sqrt{\left| \sum_i \frac{\partial(\mathcal{O}(t_i))}{\partial(\mathcal{M})} \cdot \Sigma_i^{-1} \cdot \frac{\partial(\mathcal{O}(t_i))}{\partial(\mathcal{M})} \right|}, \tag{7}$$

where Σ is the covariance matrix ($\Sigma_{i,j} = \sigma_{i,j}^2$).

$$G = -\sin\Omega\sin\omega + \cos\Omega\cos\omega\cos i \qquad (24)$$

$$H = \cos\omega\sin i. \qquad (25)$$

The factor of R_o in Equations (13) and (14) is included to convert from angular distance to physical distance. Here, M is the mass of the system (approximated as $M_{\rm cent}$), G is the gravitational constant, a is the semimajor axis of the orbit, defined as

$$a = \left(\frac{GMP^2}{4\pi^2}\right)^{1/3}, \qquad (26)$$

and E is the eccentric anomaly, defined in terms of the mean anomaly m as

$$E - e\sin E = \frac{2\pi}{P}(t - T_o) = m. \qquad (27)$$

To obtain a form of the observable-based prior in model-parameter space, as defined in Section 2.1, we must transform from observable space to model-parameter space. Because the measured and orbital observables are linearly related, a uniform distribution in one parameter set implies a uniform distribution in the other—and thus we can use the orbital observables for the transformation. In this paper, we explore two observable-based priors: (a) a prior based solely on astrometric observables $\mathcal{O} = \{X, Y\}$, and (b) a prior based on both astrometric and RV observables $\mathcal{O} = \{V_X, V_Y\}$.

To perform a one-to-one transformation, we apply the transformation only to a subset of the model parameters, $\mathcal{M} = \{e, P\}$. Applying the transformation to e and P is the natural choice, for the reasons described below, though in theory there are innumerable forms that an observable-based prior could take. This one-to-one transformation causes the observable-based prior to be conditioned on the parameters that appear in the Jacobian transformation ($M_{\rm cent}$ and T_o; see Equations (33) and (34)).

For multiple star or planet fits, the orbital parameters are unique for each orbiting body, while the central potential parameters are shared. Thus, to generalize our prior to multiple star or planet fits, the priors on the global parameters should be specified separately, and we base our transformation only on the orbital parameters.

For each of the global parameters that act as translation parameters (x_o, y_o, $v_{x,o}$, $v_{y,o}$, and $v_{z,o}$), we assume a uniform prior—the standard objective prior for parameters that only shift the probability distribution (Bernardo 2005):

$$\mathcal{P}(x_o, y_o, v_{x,o}, v_{y,o}, v_{z,o}) \propto {\rm constant}. \qquad (28)$$

For the prior based solely on astrometric observables, we assume a log-uniform prior on R_o to ensure scale invariance because R_o acts as scale parameter. However, we assume a uniform prior on R_o for the prior based on astrometric and RV observables because of the R_o dependence that appears in Equation 36(a) below. A log-uniform prior is also assumed on $M_{\rm cent}$. For (a) the prior based solely on astrometric observables and (b) the prior based on both astrometric and RV observables, these assumptions on R_o and $M_{\rm cent}$ yield

$$\mathcal{P}(R_o, M_{\rm cent}) \propto \frac{1}{M_{\rm cent}}\frac{1}{R_o} \qquad (29a)$$

$$\mathcal{P}(R_o, M_{\rm cent}) \propto \frac{1}{M_{\rm cent}}. \qquad (29b)$$

Alternatively, without RV measurements, R_o and $v_{z,o}$ cannot be independently constrained, and should be fixed if these values are known. In this paper, for directly imaged planets where we have astrometric but not RV data, we assume that the mass of the primary is well-defined and that positional measurements are defined relative to the host star. We therefore fix all global parameters that describe the central potential in this case.

Of the remaining six orbital parameters, four can easily be specified separately. Because T_o is a translation parameter with respect to time, we again assume a standard objective prior of uniformity:

$$\mathcal{P}(T_o) \propto {\rm constant}. \qquad (30)$$

A uniform distribution in X and Y guarantees uniformity in $x(t)$ and $y(t)$ as described above, so the ω, Ω, and i prior components can be specified separately because X and Y do not depend on these angular parameters (Equations (16)–(19)), reducing the choice of model parameters to e and P. For the angular parameters, we assume a prior whose spatial orientation is uniform in direction (e.g., uniform in cosine of the inclination), as suggested by Ford (2006):

$$\mathcal{P}(i) \propto \sin(i), \qquad (31)$$

and

$$\mathcal{P}(\omega, \Omega) \propto {\rm constant}. \qquad (32)$$

For the prior based on astrometric observables, we set $\mathcal{O} = \{X, Y\}$ and $\mathcal{M} = \{e, P\}$ in Equation (9). The transformation from X and Y to P and e produces a Jacobian:

$$J_{\rm astro} \equiv \frac{\partial(X, Y)}{\partial(e, P)} = -\left(\frac{(GM)^2 P}{2\pi^4}\right)^{1/3}$$
$$\times \frac{[2(e^2 - 2)\sin E + e(3\,m + \sin 2E) + 3\,m\cos E]}{6\sqrt{1 - e^2}}. \qquad (33)$$

Similarly, if given RV data, the RV observables ($\mathcal{O} = \{V_X, V_Y\}$) produce a Jacobian of:

$$J_{\rm RV} \equiv \frac{\partial(V_X, V_Y)}{\partial(e, P)} = \frac{(GM\pi)^{2/3}}{3(2)^{1/3}P^{5/3}}$$
$$\times \frac{[\sin E(e^2\cos 2E + 3e^2 - 2) + \cos E(6\,m - 2e\sin E)]}{\sqrt{1 - e^2}\,(1 - e\cos E)^3}. \qquad (34)$$

Finally, the $\prod_i \sigma_i(t_j)$ term in Equation (9) must be defined within the orbital observable space in units of the measured uncertainties of x, y, and v_z. Because $\prod_i \sigma_i(t_j)$ represents some volume in parameter space, equivalent volumes between measured and orbital observable space can be related by the Jacobian that defines the transformation between these observable spaces. Because the orbital and measured observables are related linearly via Equations (13) and (14), equivalent volumes in parameter space can be transformed by

$$\sigma_X\sigma_Y = \frac{R_o^2}{\begin{vmatrix} B & G \\ A & F \end{vmatrix}}\sigma_x\sigma_y = f(\Omega, \omega, i)R_o^2\sigma_x\sigma_y, \qquad (35)$$

Table 1
Summary of Observational Data

Object	$\delta\phi_{\mathrm{astro}}$ [a]	$\delta\phi_{\mathrm{RV}}$ [b]	Number of Astrometry Points	Number of RV Points	Data Source
S0-2	∼102%	∼96%	38	47	Boehle et al. (2016)
HR 8799b	∼3%	⋯	13	⋯	Konopacky et al. (2016)
HR 8799c	∼5%	⋯	13	⋯	Konopacky et al. (2016)
HR 8799d	∼7%	⋯	12	⋯	Konopacky et al. (2016)
HR 8799e	∼11%	⋯	9	⋯	Konopacky et al. (2016)

Notes.
[a] Angular phase coverage: percentage of the 3D orbit covered by astrometric observations, based on true anomaly.
[b] RV phase coverage: percentage of the 3D orbit covered by RV observations, based on true anomaly.

which reduces to $\sigma_X\sigma_Y \propto R_o^2\sigma_x\sigma_y$ because the Jacobian is simply some function f of the angular parameters. Analogously, $\sigma_{V_X}\sigma_{V_Y} \propto \sigma_{v_z}^2$, where the factor of R_o^2 does not appear because the observable v_z is measured in units of physical distance per unit time, and thus there is no conversion from angular to physical distance. Plugging these relations into Equation (9), we obtain a form of the two observable-based priors used in this paper:

$$\mathcal{P}(P, e) \propto \sum_i \frac{1}{R_o^2\,\sigma_x(t_i)\,\sigma_y(t_i)}|J_{\mathrm{astro}}(t_i)| \qquad (36a)$$

$$\mathcal{P}(P, e) \propto \sum_i \frac{1}{R_o^2\,\sigma_x(t_i)\,\sigma_y(t_i)}|J_{\mathrm{astro}}(t_i)| \\ + \sum_j \frac{1}{\sigma_{v_z}^2(t_j)}|J_{\mathrm{RV}}(t_j)|. \qquad (36b)$$

3. Methodology and Observational Data

We implement the new observable-based prior in *Efit5* (Meyer et al. 2012), an orbit fitting code developed by the *Galactic Center Orbits Initiative*. *Efit5* performs a Bayesian analysis using MULTINEST (Feroz & Hobson 2008; Feroz et al. 2009), a multimodal nested sampling algorithm.

The following analysis uses astrometric and RV data published by Boehle et al. (2016) for the short-period star S0-2, and astrometric data published by Konopacky et al. (2016) for the HR 8799 directly imaged exoplanet system. We use only Keck HR 8799 data, rather than data from multiple cameras and reduction pipelines, to isolate the effects of phase coverage by eliminating the additional systematics originating from the use of multiple data sets (Konopacky et al. 2016). The model for the HR 8799 analysis assumes that the origin of the system is fixed and that the mass of and the line-of-sight distance to the host star are known. The mass of the star was set to $1.516\,M_\odot$ (Baines et al. 2012) with an uncertainty of $\pm 0.15\,M_\odot$ (Konopacky et al. 2016), and the distance to the star was set to 39.4 ± 1.1 pc (van Leeuwen 2007). A summary of the S0-2 and HR 8799 observations is provided in Table 1, and the data are plotted in Figure 1.

In the following section, we test the performance of the observable-based prior compared to that of commonly assumed uniform priors, using simulations based on these data. We then apply the prior directly to these S0-2 and HR 8799 data in Section 5, to evaluate the scientific impact.

4. Testing the Observable-based Prior

4.1. Simulated Data Sets

In this section, we describe the simulated data sets that are used to evaluate the effects of the observable-based prior on fitted model parameters. The simulated GC data are based on the short-period star S0-2, and exoplanet data are based on HR 8799d. The input parameter values from which mock data were generated are provided in Table 2. One hundred simulations were run for each test case. Mock data were drawn from a normal Gaussian distribution with a mean equal to the value predicted by the model at each epoch and a standard deviation equal to the assumed error, as described below.

For the GC test cases, we simulated S0-2 data with varying angular phase coverage—defined here as the percentage of the orbit that is swept by observations, with respect to the true anomaly. The simulations are based on a single star in order to keep variables such as angular orientation constant while only varying phase coverage, though future work will probe the effects of differences in intrinsic orbital parameters. To date, S0-2 is the brightest star in the central cluster with full orbital phase coverage. As such, S0-2's relatively well-defined orbit provides a realistic data set as a basis for comparison. A model orbit was created from a set of assumed parameter values published by Boehle et al. (2016). All simulated data points assume an astrometric and RV error equal to S0-2's average observational error.

Exoplanet mock data were generated in the same manner as were the S0-2 simulations, with sampling dates and average errors based on HR 8799d astrometry published by Konopacky et al. (2016). Simulations were only run for one of the HR 8799 planets, as an example, because the statistical measures used to evaluate prior performance with these simulations are not as robust to multimodality as that which is used to evaluate prior performance with the true data for these astrometry-only cases (see Sections 4.2 and 5.2.1).

Table 3 summarizes the simulated data sets. Test Case 1, which serves as a full-phase-coverage example, samples S0-2 mock data on the same dates for which astrometric and RV data were reported by Boehle et al. (2016). For all subsequent GC test cases, evenly spaced observations of S0-2 were simulated every 6° on the plane of the sky ($\pm 0°.5$ to simulate multiple observations per year). In Test Case 2, data points are centered in time on the periapse. The simulated observations in Test Case 2 cover ∼86% of the orbit based on the true anomaly, but only ∼50% of the orbit on the plane of the sky. All variations of Test Case 3 are centered on the apoapse, with varying degrees of phase coverage but the same observing cadence.
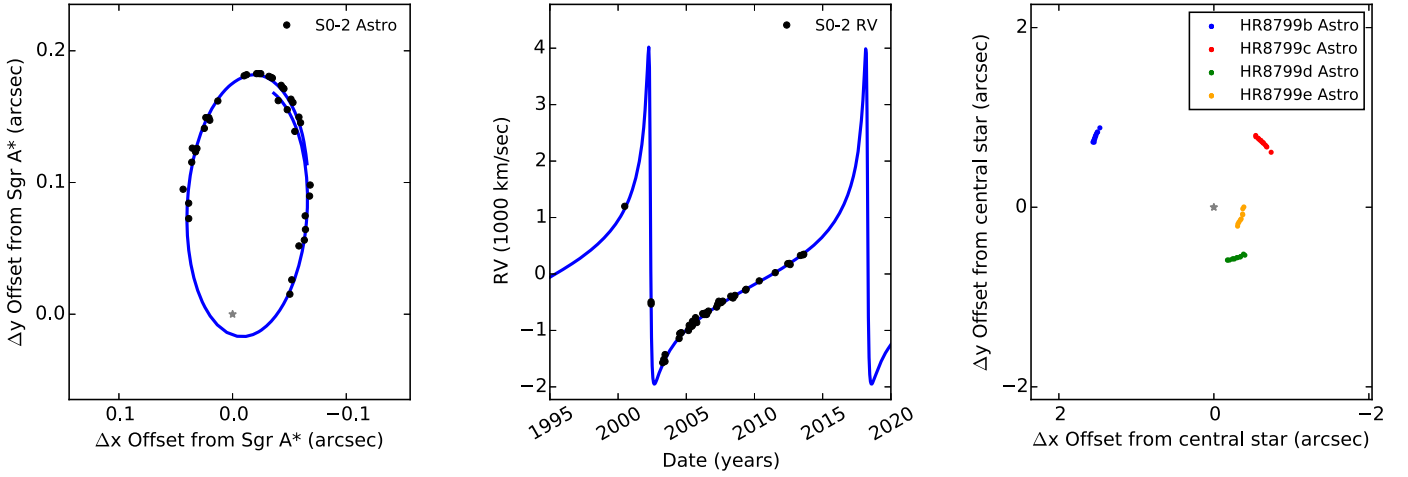
Figure 1. Left: S0-2 Astrometric data published by Boehle et al. (2016), with the best-fit orbit shown in blue. The gray star indicates the position of the central SMBH. The time baseline of observations extends from 1995 to 2013, so although 2002 data (taken during the previous closest approach) are not plotted (to avoid possible effects of confusion between S0-2 and the SMBH), there is full orbital phase coverage. Middle: S0-2 RV data from the same source, with the best-fit RV curve shown in blue. Right: HR 8799 astrometric data published by Konopacky et al. (2016). The best-fit orbits are not plotted here because, with such low phase coverage, parameter posteriors are multimodal. The gray star indicates the position of the central star.

<div style="text-align:center">

Table 2
Simulation Input Parameter Values

</div>

	[a]S0-2 Simulations	[b]HR 8799d Simulations
Global Parameters[c]		
M_{cent} (Solar Masses)	4.02×10^6	1.516
R_o (pc)	7860	39.4
x_o (mas)	2.74	0
y_o (mas)	-5.06	0
$v_{x,o}$ (mas yr^{-1})	-0.04	0
$v_{y,o}$ (mas yr^{-1})	0.51	0
$v_{z,o}$ (km s^{-1})	-15.84	0
Orbital Parameters[c]		
Ω (deg)	228.0	59
ω (deg)	66.8	92.4
i (deg)	134.2	29
P (yr)	15.92	112.7
T_o (yr)	2002.347	1995.4
e	0.892	0.02

Notes.
[a] Input values for S0-2 simulations, derived from a simultaneous orbital fit of S0-2 and S0-38 with a jackknife bias term added for the reference frame (Boehle et al. 2016).
[b] Input values for HR 8799d simulations, based on values used for simulations from Konopacky et al. (2016).
[c] See Section 2.2 for description of parameters.

<div style="text-align:center">

Table 3
Summary of Simulated Test Cases

</div>

Test Case[a]	$\delta\phi_{astro}$[b]	$\delta\phi_{RV}$[c]	Number of Astrometry Points	Number of RV Points	Description
1	101.9%	95.8%	38	47	S0-2 True Sampling
2	86.0%	86.0%	18	18	S0-2 Periapse Centered
3.1	71.9%	71.9%	56	56	S0-2 Apoapse Centered
3.2	66.5%	66.5%	55	55	
3.3	41.4%	41.4%	53	53	
3.4	31.7%	31.7%	52	52	
3.5	27.2%	27.2%	51	51	
3.6	25.9%	25.9%	51	51	
3.7	21.4%	21.4%	50	50	
3.8	18.3%	18.3%	48	48	
3.9	16.2%	16.2%	45	45	
3.10	15.3%	15.3%	46	46	
3.11	13.9%	13.9%	44	44	
3.12	13.7%	13.7%	42	42	
3.13	11.1%	11.1%	40	40	
4	7.2%	…	12	…	HR 8799d True Sampling

Notes.
[a] See Section 4.1 for a description of the simulated data sets.
[b] Angular phase coverage: percentage of the 3D orbit covered by astrometric observations, based on true anomaly.
[c] RV phase coverage: percentage of the 3D orbit covered by RV observations, based on true anomaly.

This is achieved by systematically removing data points closest to the periapse for each subsequent simulation, such that the total orbital phase coverage (and thus the number of data points) decreases without affecting the cadence of remaining measurements. Test Case 4 samples HR 8799d mock data on the same dates for which astrometric data were reported by Konopacky et al. (2016).

4.2. Simulation Results and Statistical Measures

For the simulations described in Section 4.1, two statistical measures are used to evaluate the general performance of our observable-based prior compared to that of uniform priors. We

define a new statistic, the Bias Factor, which we use along with a classical statistic, the statistical efficiency, to quantify how well regions near the true value are sampled. Consistently sampling regions of parameter space that are favored by the data is one property that must be satisfied for posteriors to be truly objective (Bernardo 2005). Although observable-based priors are weakly informative rather than truly objective, we
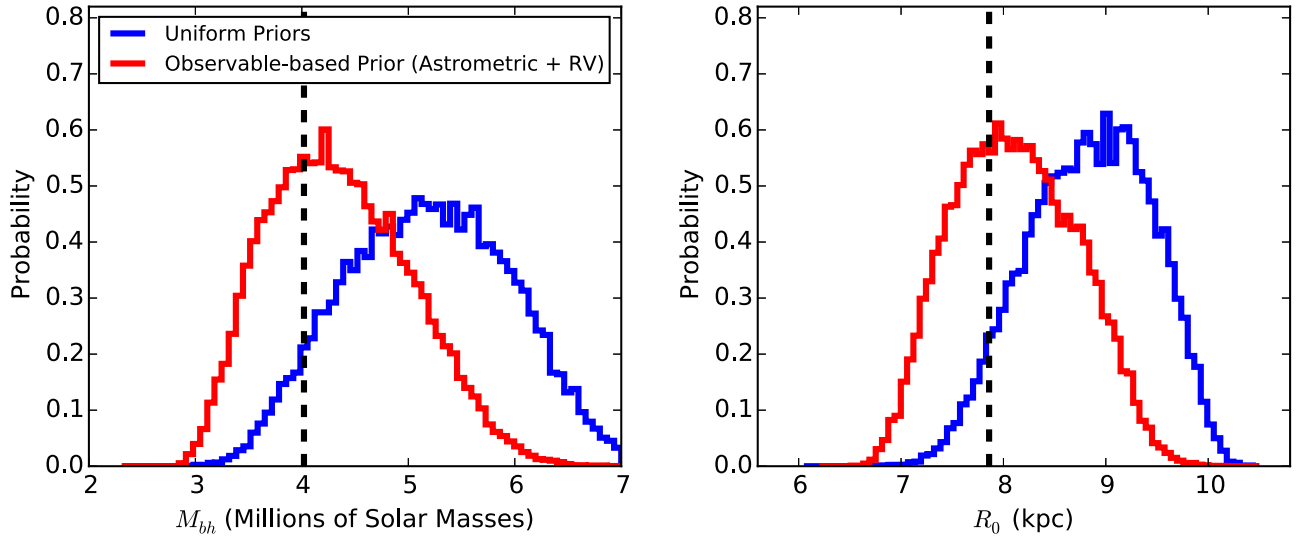
Figure 2. Marginalized 1D Posteriors for the black hole mass (left) and the line-of-sight distance to the GC (right) from a single orbital fit for Test Case 3.9, ~16% phase coverage centered on the apoapse, assuming uniform priors (blue) and a new prior based on uniformity in the astrometric and RV observables (red). All curves are normalized such that the area under the curve is equal to one. For reference, the assumed true value from which the mock data are generated is indicated by the dashed vertical line. The observable-based prior, as compared to standard uniform priors, allows the posteriors to shift closer to the assumed true value, indicating a less biased result (smaller specific bias value).

show that the less-subjective nature of these priors (compared to that of commonly assumed uniform priors) indeed promotes consistent sampling due to improvements in the Bias Factor and statistical efficiency.

4.2.1. Bias Factor

Given a hypothetical large set of N orbital fits resulting from N mock data sets, we define the Bias Factor β_F as

$$\beta_F = \underset{i \in N}{\text{median}}\{\beta_i\}, \qquad (37)$$

where β_i is defined to be the Specific Bias in parameter θ for the ith orbital fit. Here, θ denotes any parameter in the set of model parameters, $\mathcal{M}$. The Specific Bias β_i is given by the difference between the median parameter value ($\hat{\theta}_i$) and the assumed true value (θ_{true}), normalized by half the 68% central credible interval for that parameter (σ_i, see Section 4.2.2):

$$\beta_i = \frac{\hat{\theta}_i - \theta_{\text{true}}}{\sigma_i}. \qquad (38)$$

The Specific Bias effectively measures the deviation from the assumed true parameter value, in units of σ. Ideally, for each orbital fit, $\hat{\theta}$ has a 50% probability of being greater than θ_{true}, and a 50% probability of being less than θ_{true}. If no bias were present, the probability distribution of β_i values would be normally distributed about zero. For a large number of data sets, the Bias Factor should thus, in theory, be consistent with zero. A Bias Factor greater (less) than zero indicates that the parameter is statistically biased high (low).

We first illustrate the Specific Bias by looking at the effects of the prior on a single simulation. Figure 2 shows marginalized 1D M_{bh} and R_o posteriors for a randomly chosen fit in Test Case 3.9 (Table 3), which serves as a low-phase-coverage example. The resulting posteriors shift closer to the assumed true value when using an observable-based prior rather than uniform priors, indicating a Specific Bias value closer to zero (less biased). Figure 3 shows the set of β_i values for all 100 orbit fits for the same test case to confirm that the improvement

detected in Figure 2 is significant. Here, the distribution of β_i values shifts closer to zero with the observable-based prior.

We quantify this shift toward a less-biased result by evaluating the Bias Factor, β_F. Figure 4 shows the Bias Factor for M_{bh} and R_o for all GC test cases as a function of phase coverage. As phase coverage decreases, the prior has a more profound impact on parameter estimates. When more than half of the orbit (based on true anomaly) has been observed, the Bias Factor is low and remains roughly consistent between the uniform prior and the observable-based prior. Below ~40% phase coverage, the Bias Factor increases more rapidly with uniform priors than it does with the observable-based prior. This cutoff is consistent with the onset of bias seen by Lucy (2014). With the observable-based prior, the Bias Factor improves by a factor of two at low phase coverage. For example, the Bias Factor for R_o rises to greater than 1.2 with uniform priors, but remains ~0.6 with the observable-based prior. Similarly, the Bias Factor for M_{bh} rises to greater than 1.0 with uniform priors, but remains less than 0.5 with the observable-based prior. If the Bias Factor for a given parameter is 1.0, then the average output value of that parameter is 1σ greater than the true value. A Bias Factor of 1.0 would indicate a consistent shift from the "true" value, which can have profound effects on the accuracy of resulting inferences.

Although we only highlight the Bias Factor for M_{bh} and R_o in Figure 4 for clarity, the Bias Factor is improved with the observable-based prior for all global and orbital parameters in cases of low phase coverage. Figure 5 shows this improvement in the Bias Factor for all parameters for Test Case 3.9, which again serves as an example of a low-phase-coverage test. In this case, by assuming an observable-based prior rather than uniform priors, the Bias Factor decreases from 0.98 ± 0.06 to 0.35 ± 0.06 for M_{bh} (a 65% improvement), and from 1.02 ± 0.07 to 0.46 ± 0.06 for R_o (a 54% improvement). There are similar improvements in all parameters in this case— a 55% improvement in x_o, 40% in y_o, 64% in $v_{x,o}$, 61% in $v_{y,o}$, 57% in $v_{z,o}$, 57% in Ω, 75% in ω, 62% in i, 80% in P, 36% in T_o, and 31% in e.
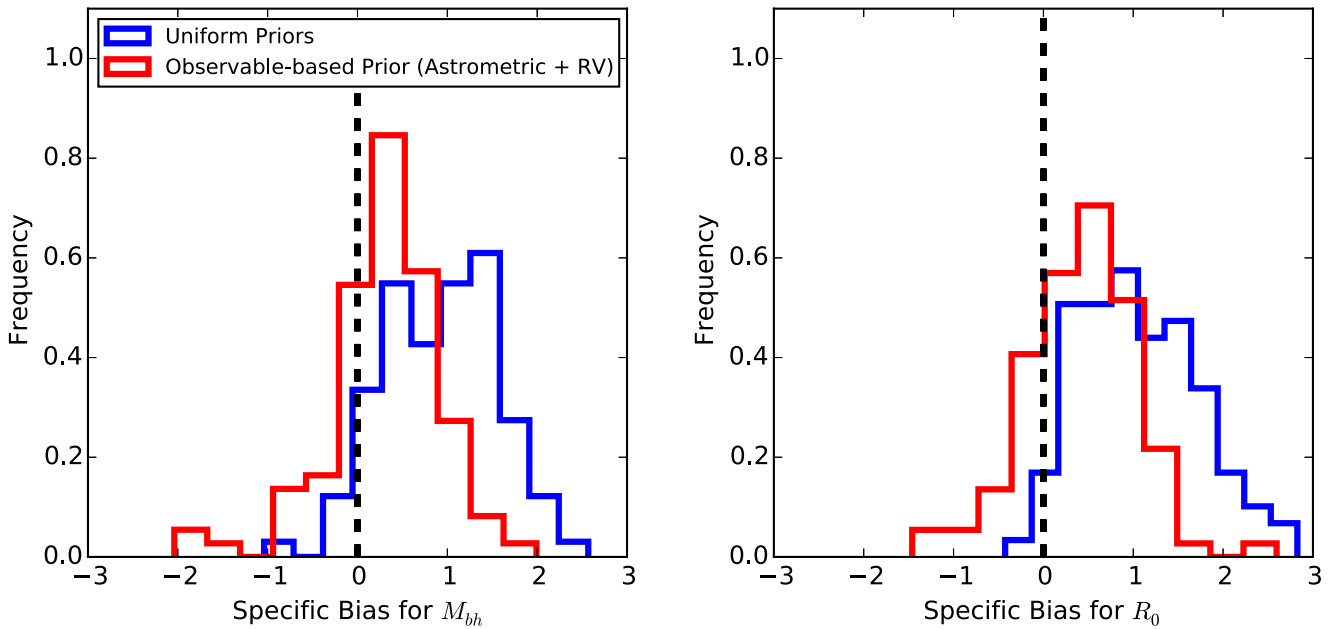
Figure 3. Probability distribution of Specific Bias values (Equation (38)) for the set of $N = 100$ orbit fits in Test Case 3.9, $\sim 16\%$ phase coverage centered on the apoapse. The distributions are plotted for the black hole mass (left) and the line-of-sight distance to the GC (right) assuming uniform priors (blue) and a new prior based on uniformity in the astrometric and RV observables (red). All curves are normalized such that the area under the curve is equal to one. The dashed line at zero indicates an unbiased result. The distribution shifts closer to zero with the observable-based prior as compared to the uniform prior, indicating that the reduction in bias seen in Figure 2 is statistical. The Bias Factor is then defined as the median of each of the above distributions.

Because credible intervals are difficult to define for multimodal distributions, data must have some minimum constraining power in order for the Bias Factor to be a robust performance metric. With greater than $\sim 10\%$ phase coverage and with both astrometric and RV data, resultant posteriors are monomodal and we can define a central credible interval for the GC cases presented. We do not present the Bias Factor for GC test cases with less than $\sim 10\%$ phase coverage, because the errors become so large that nearly the entire parameter range is covered and the utility of the Bias Factor as a performance metric breaks down. Similarly, in astrometry-only cases such as HR 8799, posteriors are often multimodal without the additional constraints provided by RV data, causing central credible intervals not to be well-defined. We therefore do not run simulations for all HR 8799 planets, because the applicable analyses depend on central credible interval construction. However, it is interesting to evaluate the statistical properties of eccentricity estimates nonetheless, because low phase coverage is known to bias eccentricity estimates toward artificially high values (e.g., Konopacky et al. 2016). As an example, we run simulations for HR 8799d and evaluate the eccentricity bias using the 68% error on the mode rather than central credible intervals. Eccentricity estimates indeed are biased high when uniform priors are assumed, but are consistent with the input value with our new observable-based prior. For HR 8799d, the Bias Factor on e is reduced from 3.17 ± 0.21 with uniform priors to an unbiased value of 0.07 ± 0.26 with the observable-based prior. This result indicates that using the observable-based prior with low-phase-coverage data can help mitigate the known risk of biasing eccentricity estimates toward artificially high values. We discuss the reason for this improvement in Section 5.2.

4.2.2. Statistical Efficiency

We use the statistical efficiency to investigate how the observable-based prior affects the reliability of calculated

confidence intervals. Bayesian confidence intervals (credible intervals) and classical confidence intervals are fundamentally different in both construction and interpretation (Host et al. 2007). Unlike credible intervals, which are defined simply as regions containing a prescribed posterior probability, classical confidence intervals have a more intricate definition. In the classical definition, for a sufficiently large number of experiments, the confidence interval inferred from each experiment will contain, or cover, the universally "true" value a prescribed fraction of the time (confidence level $\times 100\%$) (Neyman 1937). By this definition, a 68% confidence interval requires that 68 out of 100 possible observed (or randomly drawn) data sets produce a confidence interval that covers the true value. Confidence intervals that cover the true value more than the prescribed frequency are said to over-cover, whereas confidence intervals that cover the true value less than the prescribed frequency are said to under-cover. Unfortunately, most algorithms used to calculate confidence intervals do not guarantee exact coverage, and thus can either under- or over-cover. Under- or over-covering is common in prior-dominated regimes where data are not rigorously constraining. In such cases, prior information can have a profound impact on the resulting confidence intervals (Lucy 2014).

Statistical efficiency is a powerful performance diagnostic that is used to investigate the accuracy of calculated confidence or credible intervals. Statistical efficiency is defined as the ratio of effective coverage (the experimentally determined percentage of data sets in which the inferred confidence or credible interval covers the true value) to stated coverage (68% for a 1σ confidence interval). While based on the classical definition of confidence intervals, effective coverage can be used to evaluate the effectiveness of both credible and confidence intervals (Cameron 2011). An efficiency of 1 indicates exact coverage, whereas an efficiency less (greater) than 1 indicates that the credible interval under-(over-)covers.
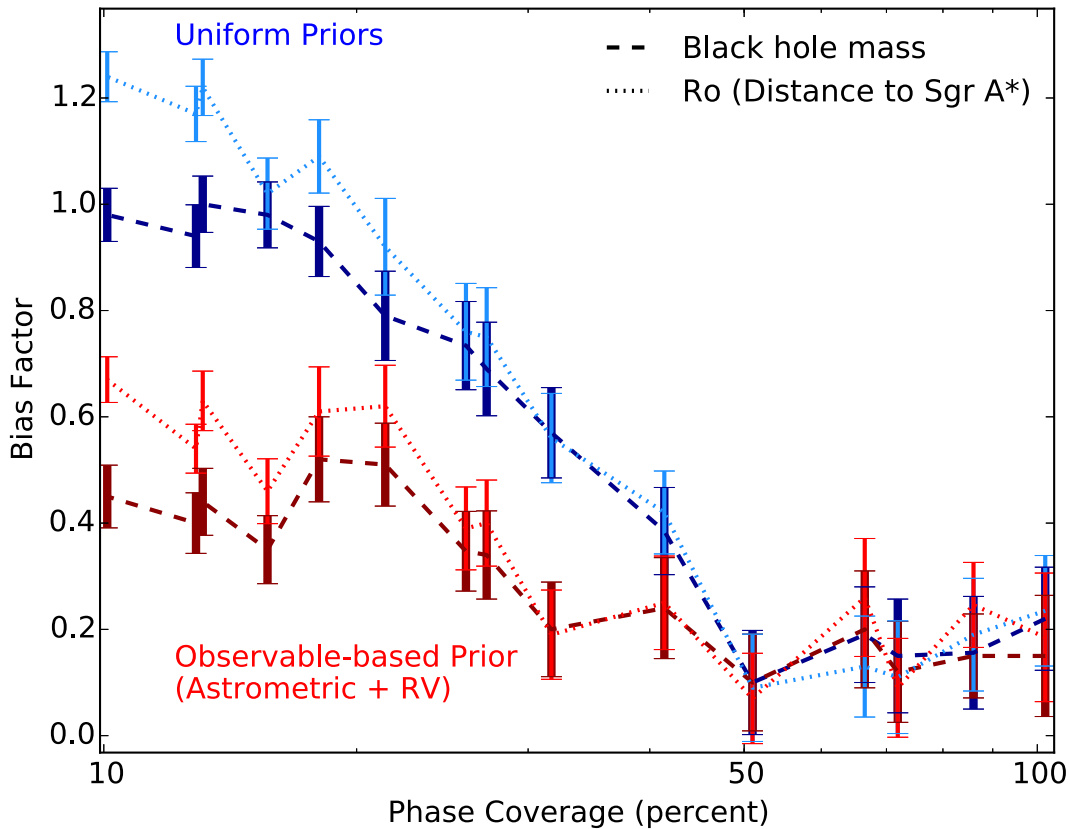
Figure 4. Bias Factor (Section 4.2.1) as a function of phase coverage for the black hole mass and the line-of-sight distance to the GC, assuming uniform priors (blue) and a new prior based on uniformity in the astrometric and RV observables (red). Error bars indicate 1σ standard errors on the Bias Factor. The onset of bias occurs at less than ~40% phase coverage, and is less dramatic with the observable-based prior than with uniform priors. At low phase coverage, there is a factor of two improvement in the Bias Factor with the observable-based prior over standard uniform priors.

Figure 6 presents the statistical efficiency as a function of phase coverage for all simulated GC test cases, again focusing on the black hole mass (M_{bh}) and the line-of-sight distance to the GC (R_o), for clarity. With phase coverage greater than ~40%, credible intervals are well-defined because the statistical efficiency is consistent with one for both M_{bh} and R_o, indicating nearly exact coverage. Below ~40% phase coverage, the statistical efficiency decreases to less than one with uniform priors, indicating that the derived credible intervals under-cover and errors are underestimated. This drop in statistical efficiency is consistent with the onset of bias around 40% phase coverage (Section 4.2.1). The risk of underestimating errors below ~40% phase coverage is mitigated by assuming an observable-based prior. The statistical efficiency remains greater than or equal to one for all S0-2 test cases with the observable-based prior. With statistical efficiencies greater than one, the credible intervals over-cover and error estimates are overly conservative. While exact coverage is ultimately desired, it is better to be too conservative than to underestimate the errors.

In light of the predicted and detected eccentricity biases for HR 8799d, we also investigate the HR 8799d eccentricity efficiency. As with the Bias Factor, statistical efficiency depends on credible interval construction, and thus must be evaluated cautiously for cases such as HR 8799 where posteriors are multimodal and central credible intervals are not always well-defined. As such, we again use the 68% error on the mode of the posterior distribution rather than using central credible intervals. For HR 8799d, statistical efficiency

for eccentricity increases from 0.6 ± 0.05 (under-covering) to greater than one (sufficiently defined credible intervals) with the observable-based prior, indicating that the errors on the mode are no longer underestimated with the new prior.

5. Application of the Observable-based Prior

We now apply the observable-based prior to real S0-2 and HR 8799 data (described in Section 3), to evaluate the scientific impact.

5.1. S0-2 Results

Table 4 shows that, with full orbital phase coverage, the orbital solution for S0-2 derived with the observable-based prior is consistent with that published by Boehle et al. (2016). S0-2 astrometric and RV data have high constraining power, and thus the orbital solution is not influenced heavily by prior effects. This result is supported by simulation results for the high-phase-coverage limit in Figures 4 and 6. There are many stars with low phase coverage in the GC whose orbital solution could be improved by the observable-based prior; as such, future work will include using the new prior to fit orbits to these stars in order to probe the effects of the prior on stars with different eccentricity distributions and angular orientations.

5.2. HR 8799 Results and Statistical Analysis

Here, we investigate how the inferred orbital plane configurations of the HR 8799 planets change based on prior choice. Prior choice affects whether the planets are hypothesized to have
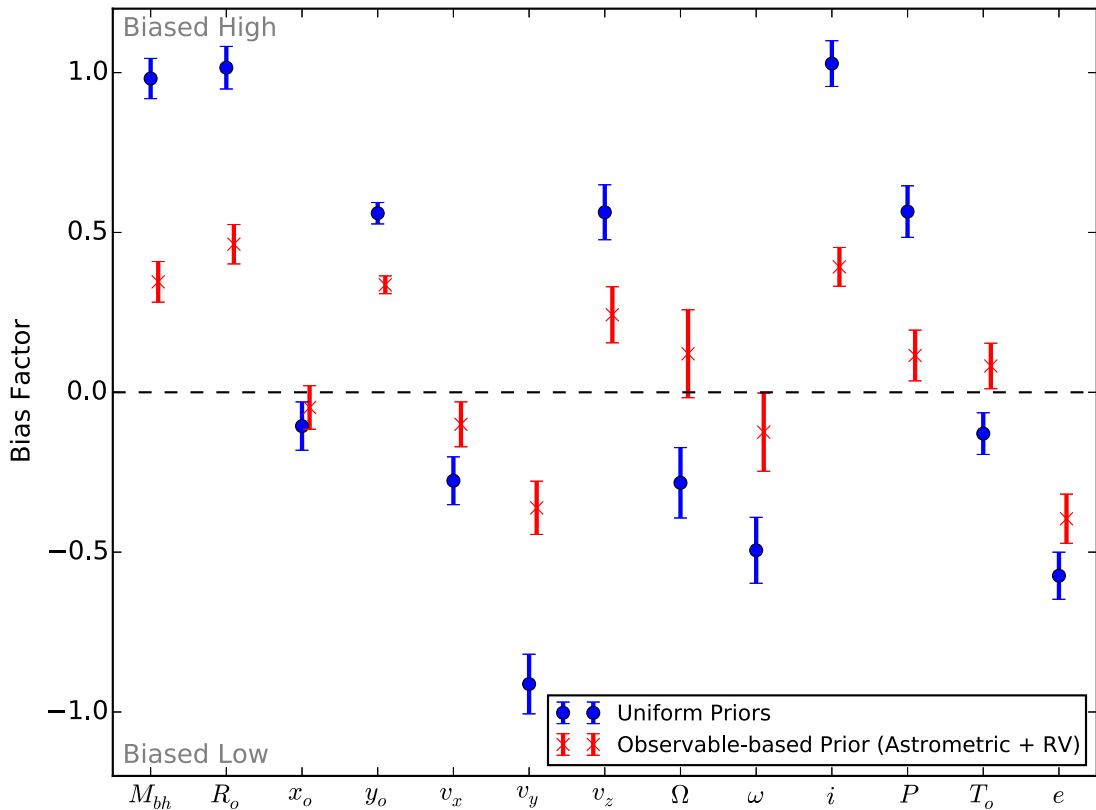
Figure 5. Bias Factor (Section 4.2.1) for all parameters for Test Case 3.9, which serves as an example of low angular phase coverage (∼16%), assuming uniform priors (blue) and a new prior based on uniformity in the astrometric and RV observables (red). Error bars indicate 1σ standard errors on the Bias Factor. The Bias Factor is reduced for all parameters by assuming an observable-based prior rather than standard uniform priors.

consistent inclinations, and thus prior influences must be taken into consideration in this low-phase-coverage regime. The observable-based prior provides stronger evidence that the four planets have a consistent inclination of ∼30° to within 1σ (see inclination posteriors in Figure 7). Figure 8 shows 1σ contours for the joint probability distribution functions between eccentricity and inclination, as well as between angle of ascending node (Ω) and inclination, for each of the four HR 8799 planets. While we cannot claim coplanarity between the host star and the planets, because the Ω posteriors remain largely unconstrained, the possibility is allowed by this analysis (bottom panel of Figure 8). Table 5 lists the median values of the inclination posteriors derived with both flat priors and the observable-based prior, for each of the four planets.

Further, the observable-based prior favors lower eccentricity estimates for all four planets than those inferred with uniform priors, providing more stringent eccentricity upper limits than those inferred with uniform priors and allowing for the possibility of nearly circular orbits (top panel of Figure 8). This result is consistent with the HR 8799d simulation reported in Section 4.2.1, which shows that HR 8799d eccentricity estimates are biased high with uniform priors but are unbiased with the observable-based prior. Previous works have noted that such artificially high eccentricities can result from T_o estimates being biased toward the epochs of observation (e.g., Konopacky et al. 2016). The reason for this is that higher eccentricity modes that correspond to this biased region of T_o parameter space are accentuated with flat priors (see e, T_o joint posterior in left panel of Figure 9). The observable-based prior accounts for this bias by suppressing these regions (right panel of Figure 9). For completeness, 1D marginalized posteriors for

all orbital parameters for HR 8799 b, c, d, and e are presented in the Appendix (Figures 11–14), and the Monte Carlo chains are available in the .tar.gz package.

5.2.1. Expected Information Gained

Biases introduced in the previous analysis affect the posterior information content. Ideally, a prior should be chosen such that the data contribute maximally to the posteriors (e.g., the prior, as compared to the likelihood, adds the least amount of information possible to the posterior). For a given data set $\mathscr{D}$, this is equivalent to maximizing the relative entropy (Kullback & Leibler 1951) between the posterior and prior of some model parameter set $\mathscr{M}$,

$$\kappa \equiv \int d\mathscr{M}\, \mathcal{P}(\mathscr{M}|\mathscr{D}) \log\left[\frac{\mathcal{P}(\mathscr{M}|\mathscr{D})}{\mathcal{P}(\mathscr{M})}\right]. \tag{39}$$

Equation (39) is not an ideal measure of information gained, as it assumes only one possible data set. Instead, the average relative entropy between the posterior and the prior, or the average information gained in the posterior over the prior (defined as "expected information" by Lindley (1956)), can be used for this purpose:

$$\mathcal{I} \equiv \int d\mathscr{D}\, \mathcal{P}(\mathscr{D}) \int d\mathscr{M}\, \mathcal{P}(\mathscr{M}|\mathscr{D}) \log\left[\frac{\mathcal{P}(\mathscr{M}|\mathscr{D})}{\mathcal{P}(\mathscr{M})}\right] \tag{40}$$

$$= \int\int d\mathscr{D}\, d\mathscr{M}\, \mathcal{P}(\mathscr{M}, \mathscr{D}) \log\left[\frac{\mathcal{P}(\mathscr{M}, \mathscr{D})}{\mathcal{P}(\mathscr{D})\mathcal{P}(\mathscr{M})}\right] \tag{41}$$
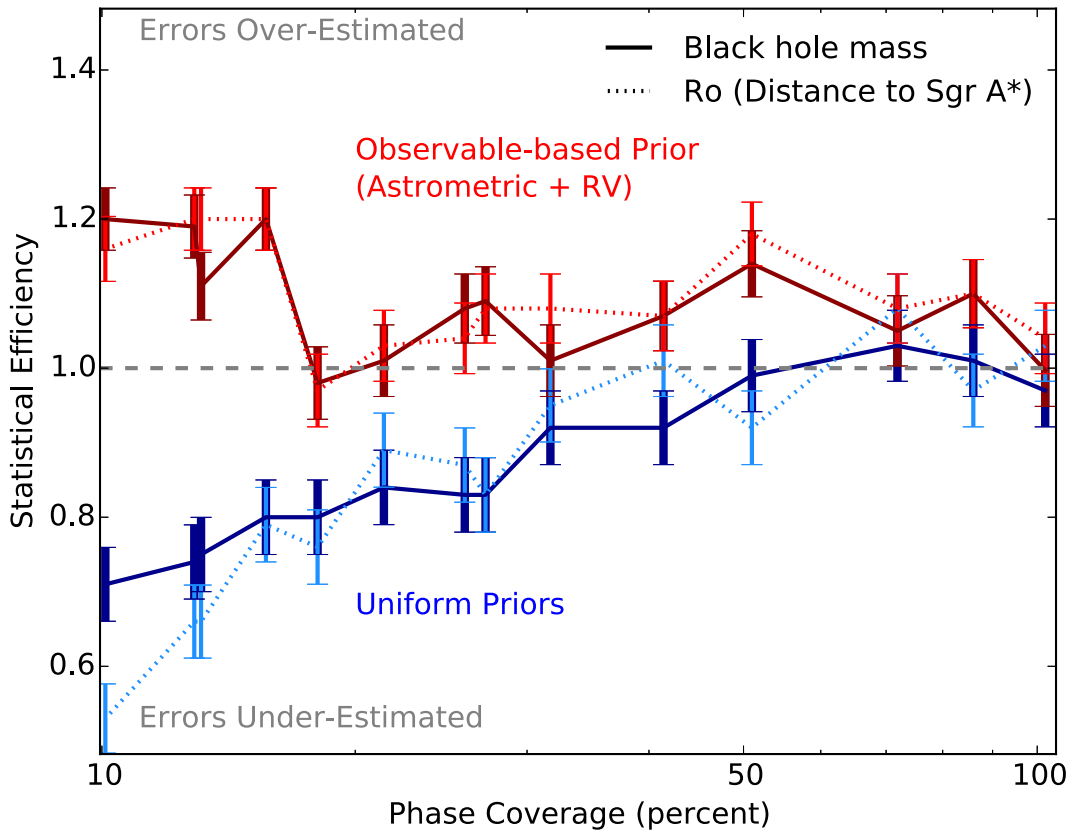
Figure 6. Statistical efficiency (Section 4.2.2) as a function of phase coverage for the black hole mass and the line-of-sight distance to the GC, assuming a uniform prior (blue) and a new prior based on uniformity in the astrometric and RV observables (red). Errors on the statistical efficiency are derived from a binomial distribution (with $N = 100$ trials). The dashed horizontal line at one indicates exact coverage. Below $\sim$40% phase coverage, the statistical efficiency decreases to less than one when assuming a uniform prior, indicating that credible intervals under-cover (errors are underestimated). Credible intervals remain well-defined when assuming an observable-based prior, as the statistical efficiency remains at or greater than one. In this case, the observable-based prior eliminates the problem of underestimated errors.

$$= \int d\mathcal{M}\, \mathcal{P}(\mathcal{M}) \int d\mathcal{D}\, \mathcal{P}(\mathcal{D}|\mathcal{M}) \log\left[\frac{\mathcal{P}(\mathcal{D}|\mathcal{M})}{\mathcal{P}(\mathcal{D})} \right]. \qquad (42)$$

Other objective-prior frameworks (e.g., Berger et al. 2009) derive priors by maximizing the expected information. The sampling algorithm that we use to estimate the expected information is described in Appendix A (Algorithm 1).

Table 6 lists the expected information—gained in the posteriors over the prior—for orbital fits of each of the HR 8799 planets. When assuming an observable-based prior rather than a uniform prior, the expected information increases by $\sim$29% for HR 8799b, $\sim$26% for HR 8799c, $\sim$35% for HR 8799d, and $\sim$33% for HR 8799e. This increase in expected information indicates that, with respect to the prior, the information contained in the data contributes 25–35% more to the posterior for each of the HR 8799 planets when assuming an observable-based prior.

6. Discussion

Many stars in the Galactic Center and many directly imaged exoplanets have low orbital phase coverage, causing data to have low constraining power. Prior assumptions dominate parameter estimates in these low-phase-coverage regimes, potentially introducing biases in fitted parameters and producing inaccurate confidence intervals. Uniform priors, commonly assumed in orbit fitting, exacerbate these issues in regions of prior-dominance. In this paper, we propose a new

prior that is based on uniformity in observable space, rather than in model parameter space, to limit the impact of subjective model selection. Statistical tests applied to both simulated and real GC and HR 8799 data indicate that observable-based priors perform better than uniform priors in prior-dominated regimes.

6.1. Galactic Center Orbits

Orbits of the short-period stars within the central arcsecond (the S-stars) can be used to probe the dynamics of the GC. There are currently $\sim$40 S-stars with measured orbits, 17 of which have been used in a multistar fit to constrain the central potential (Gillessen et al. 2017). Although this multistar fit reduces the uncertainty on fundamental parameters such as the mass of and distance to the central SMBH, it is essential to ensure that accuracy is not jeopardized for this precision. As such, we must ensure that using low-phase-coverage stars in a multistar fit does not introduce biases due to the statistical effects of prior dominance. In the multistar fit from Gillessen et al. (2017), each star is weighted according to the number of data points it contributes to account for biases; however, simulated S0-2 test cases indicate that bias is correlated with orbital phase coverage, though not *necessarily* with the number of observations. This is evidenced by the fact that Test Case 1 and Test Case 2 have fewer data points than do all other S0-2 test cases (Table 3), yet have higher phase coverage and smaller bias values. However, the number of data points or
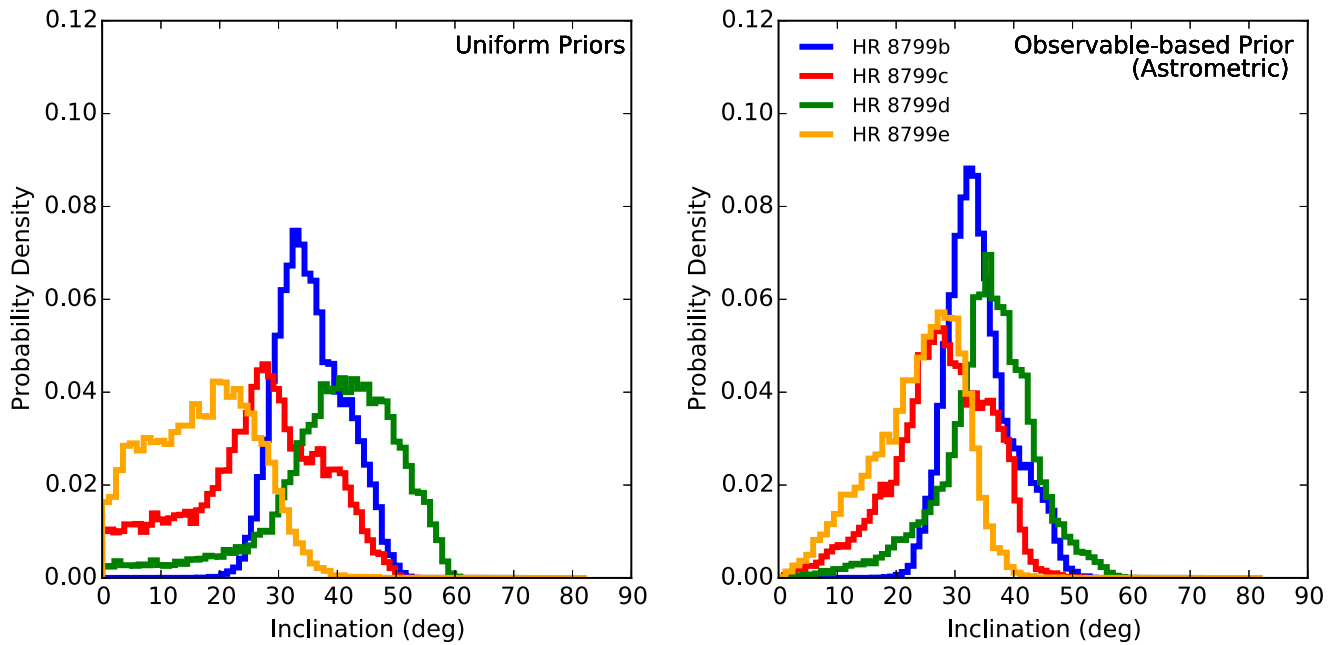
Figure 7. Inclination posteriors for HR 8799b, c, d, and e, assuming uniform priors (left) and a new prior based on uniformity in the astrometric observables (right). The observable-based prior provides stronger evidence that the inclinations of the four planets are consistent around 30° to within 1σ.

Table 4
Orbital Parameter Estimates for S0-2 Derived with the Observable-based Prior, Compared to Previously Published Values

	[a]Prior Ranges	[b]Observable-based Prior	[c] Boehle et al. (2016)
Global Parameters[d]			
M_{bh} (10^6 Solar Masses)	[2.3, 7.0]	4.01 ± 0.31	4.03 ± 0.31
R_o (kpc)	[5.90, 10.5]	7.98 ± 0.36	8.01 ± 0.36
x_o (mas)	[−6, 8]	2.04 ± 0.56	2.02 ± 0.56
y_o (mas)	[−8, 8]	-3.70 ± 1.34	-3.64 ± 1.32
$v_{x,o}$ (mas yr^{-1})	[−0.3, 0.8]	-0.10 ± 0.03	-0.10 ± 0.03
$v_{y,o}$ (mas yr^{-1})	[−0.5, 1.5]	0.71 ± 0.07	0.72 ± 0.07
$v_{z,o}$ (km s^{-1})	[−120., 120.]	-20 ± 10	-19 ± 10
Orbital Parameters[d]			
Ω (deg)	[221, 233]	227.9 ± 0.8	227.9 ± 0.8
ω (deg)	[60, 71]	66.6 ± 0.9	66.5 ± 0.9
i (deg)	[128, 139]	134.6 ± 0.9	134.7 ± 0.9
P (yr)	[15.4, 16.9]	15.90 ± 0.04	15.90 ± 0.04
T_o (yr)	[2002.29, 2002.38]	2002.344 ± 0.008	2002.343 ± 0.008
e	[0.865, 0.915]	0.890 ± 0.005	0.890 ± 0.005

Notes.
[a] Prior ranges for the observable-based prior are still specified in model-parameter space.
[b] Mean posterior estimates for S0-2 derived using same data as in column 4, but with the observable-based prior. All results with the new prior are consistent with previously published values, as expected with full phase coverage.
[c] Mean posterior estimates reported by Boehle et al. (2016) derived from an individual orbital fit of S0-2, without the jackknife bias term added for the reference frame. Though Boehle et al. (2016) also report values derived from a simultaneous fit of S0-2 and S0-38, we compare the values for S0-2 alone here, for simplicity.
[d] See Section 2.2 for description of parameters.

cadence of observations may affect prior performance for orbits with different angular orientations or eccentricities. As such, future work will include testing the performance of the observable-based prior for additional S-stars, and using information on how each star affects global parameter biases to appropriately weight the contribution of each star in a multistar fit.

In addition, orbital estimates of the S-stars are used to test formation hypotheses that attempt to explain the observed abundance of early-type stars in the GC—the so-called "paradox of youth" (e.g., Morris 1993; Ghez et al. 2003). For example, the eccentricity distribution of the S-stars can be compared to distributions expected for different formation scenarios (e.g., Gillessen et al. 2009, 2017; Perets et al. 2009; Chen & Amaro-Seoane 2014; Madigan et al. 2014). One proposed mechanism of S-star formation is tidal capture of a binary star system, which would result in an ejected hypervelocity star and a highly eccentric bound star whose
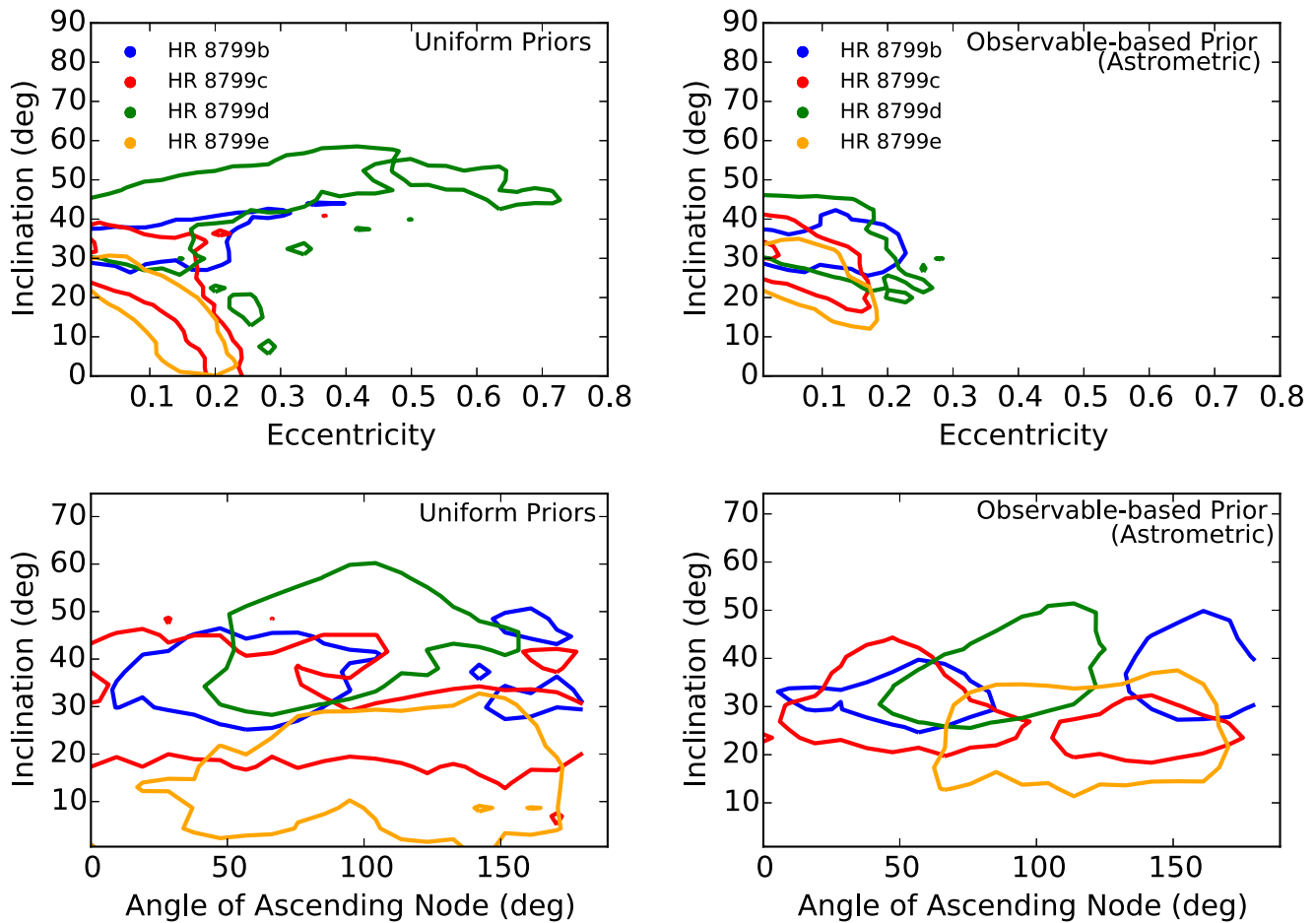
Figure 8. The 1σ contours for the joint probability distribution functions between eccentricity and inclination (top) and between angle of ascending node and inclination (bottom) for HR 8799b, c, d, and e, assuming uniform priors (left) and a new prior based on uniformity in the astrometric observables (right). With the observable-based prior, low-eccentricity orbits are favored for all four planets with an inclination of $\sim30°$ to within 1σ, placing stronger constraints on the orbital plane configuration of the HR 8799 system. The bottom panel shows that coplanar solutions are still allowed, though not necessitated, with the observable-based prior. Using the same data set, prior choice can influence our physical constraints on the system, highlighting the fact that prior effects must not be ignored.

Table 5
Inclination Estimates for the HR 8799 Planets, Derived with Uniform Priors and with the Observable-based Prior

Planet	[a]Prior Ranges (deg)	[b,c]Observable-based Prior (deg)	[d,c]Uniform Priors (deg)
HR 8799b	[0, 180]	$33.4^{+6.6}_{-4.3}$	$35.1^{+7.0}_{-5.1}$
HR 8799c	[0, 180]	$26.7^{+3.4}_{-11.2}$	$26.9^{+2.7}_{-11.8}$
HR 8799d	[0, 180]	$35.5^{+5.6}_{-7.8}$	$38.1^{+9.9}_{-13.5}$
HR 8799e	[0, 180]	$25.0^{+6.2}_{-9.5}$	$17.5^{+9.0}_{-10.7}$

Notes.
[a] Prior ranges for the observable-based prior are still specified in model-parameter space.
[b] Inclination posterior median with the associated 68% credible interval, derived using the observable-based prior.
[c] Monte Carlo chains are made available so that posterior distributions can be evaluated independently, because MAP estimates differ slightly from the median.
[d] Inclination posterior median with the associated 68% credible interval, derived using uniform priors.

orbit would circularize over a relaxation timescale (e.g., Hills 1988; Brown 2015). Whereas a relaxed stellar system expects a thermal distribution, this binary capture scenario expects the eccentricity distribution to peak toward higher

eccentricities because the two-body relaxation time ($\sim10^9$ yr; Perets et al. 2007) is an order of magnitude longer than the maximum lifetime of a B star ($\sim10^8$ yr). On the other hand, a lower-than-thermal distribution may indicate a disk-migration scenario. Gillessen et al. (2009) found that the eccentricity distribution is consistent with a thermal distribution to within 3σ, though they highlight a slight peak toward higher eccentricities. With a larger sample of stars, Gillessen et al. (2017) subsequently found that the eccentricity distribution indeed is consistent with a thermal distribution, leaving inconclusive evidence of the formation history. We can also look at the orbital plane orientations of the S-stars to see if they are compatible with the clockwise stellar disk. The inclination and angle of ascending node distributions reported by Gillessen et al. (2009) and confirmed by Gillessen et al. (2017) indicate that a majority of the S-stars have randomly distributed orientations and do not appear to be associated with the clockwise disk of stars located outside the central arcsecond. To confirm these findings or perform more robust tests of these formation scenarios, we must ensure that the measured orbits of the early-type stars with low phase coverage are not biased. Future work will include using observable-based priors to test the effects of the prior on stars with different eccentricity distributions and angular orientations.
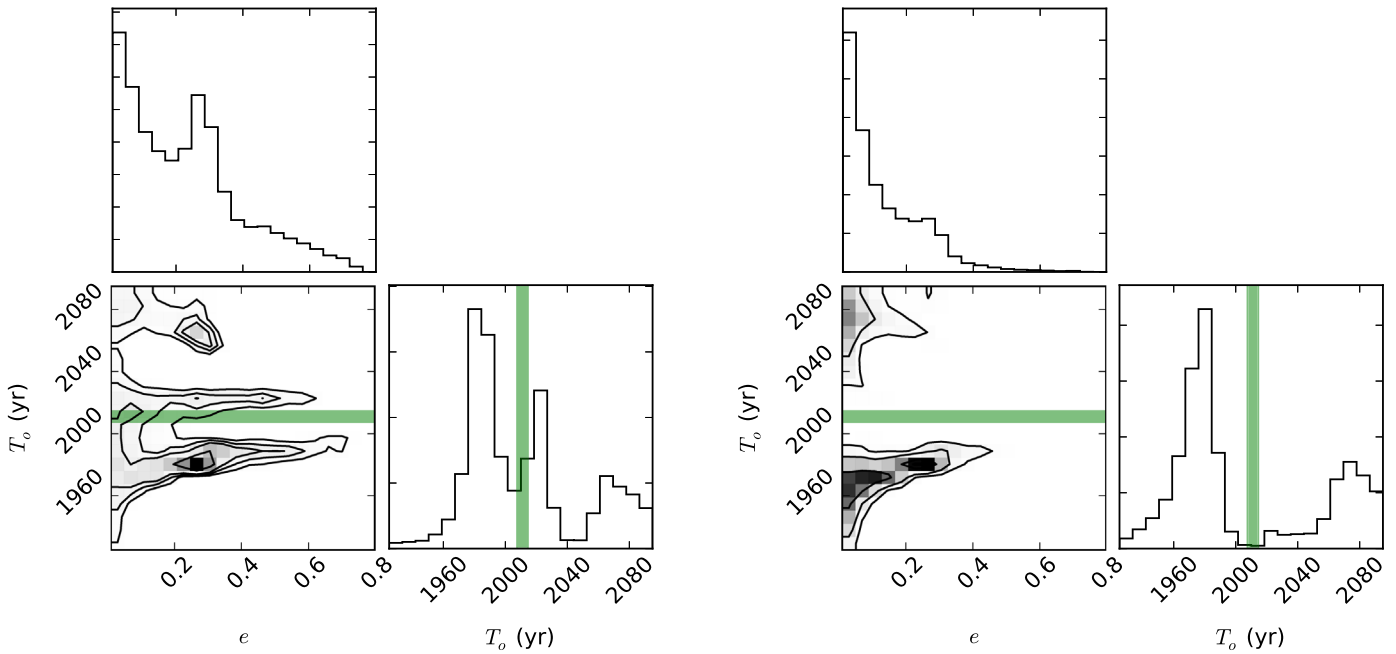
Figure 9. Joint probability distribution functions between eccentricity and T_o, assuming uniform priors (left) and an observable-based prior (right) for HR 8799d. With uniform priors, an artificial T_o mode emerges near the epochs during which observations were taken, indicated by the green shaded region. The left panel indicates that, with uniform priors, this biased region of T_o parameter space corresponds with an artificially high eccentricity mode. The observable-based prior mitigates this bias by suppressing the artificial T_o mode and consequently the artificially high eccentricity mode.
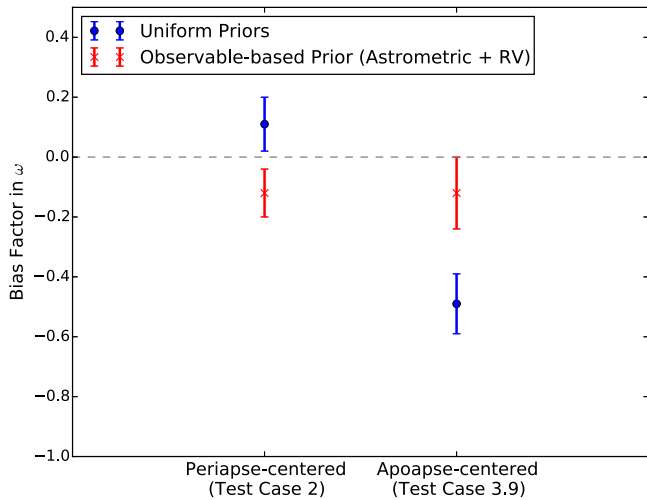
Figure 10. Bias Factor (Section 4.2.1) for the argument of periapse ω derived from simulations with uniform priors (blue) and observable-based priors (red). Values are plotted for Test Cases 2 and 3.9 (Table 3), which are centered on S0-2's periapse and apoapse with ~86% and ~16% angular phase coverage, respectively. These two simulations both cover ~50% of the orbit on the plane of the sky, making for an interesting comparison. Differences in ω between two such fits could, in theory, hint at a detection of the precession of the periapse (Parsa et al. 2017); however, the fact that Bias Factor values with the observable-based prior are consistent between the two tests, while values with uniform priors differ by over 0.5σ, indicates that differences in ω could in fact be due to statistical biases if prior effects are not considered.

Table 6
Expected Information Gained in the Posterior Over the Prior for the HR 8799 Planets

Planet	[a]Uniform Prior (Commonly Assumed)	[b]Observable-based Prior	[c] Percent Increase
HR 8799b	16.4 ± 0.1	21.2 ± 0.1	$29.3\% \pm 0.1\%$
HR 8799c	16.7 ± 0.1	21.0 ± 0.1	$25.7\% \pm 0.1\%$
HR 8799d	16.1 ± 0.2	21.7 ± 0.3	$34.8\% \pm 0.4\%$
HR 8799e	14.0 ± 0.1	18.6 ± 0.4	$32.9\% \pm 0.4\%$

Notes.
[a] Average relative entropy between posteriors and uniform priors as a measure of information gained (Section 5.2.1).
[b] Average relative entropy between posteriors and observable-based priors as a measure of information gained (Section 5.2.1).
[c] Observable-based priors allow information from data to contribute 25%–35% more to resulting inferences than it could with uniform priors.

Accurate orbital estimates in the GC are also critical, as S0-2 has recently gone through its closest approach to the SMBH in 2018, and small deviations from a Keplerian orbit are under investigation. Because S0-2 has full phase coverage, Keplerian orbital estimates with the observable-based prior are consistent with previously published results. However, for a given data set, prior choice becomes more important because the size of

the effect under consideration decreases and the complexity of the likelihood increases (Gelman et al. 2017), indicating that prior choice may play a larger role in deciphering small post-Newtonian effects. Future work will include extending this bias analysis to post-Newtonian parameters in General Relativity. Without considering prior influences on posteriors, statistical aberrations may be confused with actual physical processes. In prior-dominated regimes, statistical effects can obscure physical effects such as those from General Relativity (e.g., Grould et al. 2017; Hees et al. 2017; Parsa et al. 2017). For example, biases induced by prior dominance can cause inferred model parameters to differ when we fit an orbit to the upper and lower portions of S0-2's trajectory independently. Note that, because of the angular orientation of S0-2's orbit, the on-sky projection of the orbit in the left panel of Figure 1 does not properly convey how eccentric the true 3D orbit is. As such, angular phase coverage differs greatly from the apparent coverage on

the plane of the sky. For example, simulated data in Test Case 3.9 (apoapse-centered) and Test Case 2 (periapse-centered) both cover ∼50% of the orbit on the plane of the sky, though they differ greatly in angular phase coverage (∼16% and ∼86% of the 3D orbit based on true anomaly, respectively). Because of this difference in angular phase coverage and consequent difference in information content, the Bias Factor on the argument of periapse ω differs between these two simulations by over 0.5σ with uniform priors (Figure 10). With the observable-based prior, however, the Bias Factor on ω remains consistent between the two test cases (Figure 10), indicating that differences in ω between these apoapse- and periapse-centered fits could be due to statistical effects of prior dominance rather than a hint of the precession of the periapse.

6.2. Exoplanet Orbits

Similarly, physical models for planet formation depend heavily on orbital estimates. The four directly observed young giant planets that comprise the HR 8799 system provide an unparalleled opportunity to study the formation and evolution of giant planets (e.g., Marois et al. 2010). Accurate constraints on the planets' orbital plane parameters are essential to understanding the system's dynamical history. Consequently, our physical interpretation can be obscured by the statistical effects of low-phase-coverage data that are not rigorously constraining.

We show that prior choice affects the inference of orbital plane parameters, especially eccentricity and inclination—key parameters that provide constraints on dynamical models. Konopacky et al. (2016) note that low phase coverage can cause eccentricity estimates to be biased high. By calculating the statistical efficiency and Bias Factor on eccentricity for HR 8799d simulations (based on the 68% error on the mode of the posterior distribution rather than on central credible intervals, due to the multimodality of resulting posteriors), we confirm that eccentricity estimates indeed are biased high when uniform priors are assumed, but are unbiased with the new observable-based prior. Additionally, the errors on the mode of the posterior distribution are no longer underestimated with the new prior. This improvement in statistical efficiency and reduction in bias allows greater confidence to be placed in eccentricity estimates inferred with an observable-based prior than in those inferred with uniform priors. In fits to the HR 8799 astrometry, lower eccentricity estimates are favored with the observable-based prior than are estimated with uniform priors, allowing for the possibility of nearly circular orbits—particularly for planets d and e (Figures 11–14). In contrast to this result, analyses presented by Wertz et al. (2017) do not support assumptions of circular orbits, as they estimate an eccentricity of approximately 0.35 for HR 8799d—similar to the value we infer with uniform priors. Although Wertz et al. (2017) argue against the circular orbit hypothesis, they note that astrometric biases or underestimation of astrometric errors should not be neglected. Building off of this caveat, we suggest that biases and underestimation of errors in posterior parameter estimates—not just in initial astrometric measurements—also cannot be ignored. We therefore suggest that prior considerations be taken into account before ruling out any dynamical models.

The relative inclination of the four planets is another open question whose investigation could provide constraints on dynamical models. Evaluating whether the planets have consistent inclinations requires obtaining accurate estimates of the orbital plane parameters. As such, this question has been widely disputed due to the low constraining power of the data. In this paper, we show that the evidence for consistent inclinations of the four HR 8799 planets is stronger when assuming an observable-based prior than when assuming uniform priors (Figure 8). Early works suggest that the HR 8799 planets do not have similar inclinations (e.g., Currie et al. 2012; Pueyo et al. 2015); however, using a self-consistent data set, Konopacky et al. (2016) found evidence that the orbital planes of the four planets are consistent within 2σ. Building off of this work, using the same self-consistent data set, we provide stronger evidence that the four planets have a consistent inclination of ∼30° to 1σ. Although the Ω estimates are disjointed, there remains a large overlap in Ω parameter space with the observable-based prior, indicating that coplanar solutions could still be consistent with the data. While Wertz et al. (2017) similarly show that the HR 8799 planets have consistent inclinations between ∼20° and 38° with respect to the plane of the sky, they use 3D dynamical modeling techniques to suggest that the system might not be coplanar, at 2σ significance, due to the disjointed Ω estimates. Future work will include combining an assessment of the three-dimensional orientations of the orbits (Wertz et al. 2017) and of the system dynamics and stability with this prior analysis to further assess the possibility of coplanarity. Here, we simply demonstrate that the prior has a profound effect on parameter estimates and consequently on our physical understanding of the system itself. Prior considerations must be taken into account before confirming or denying physical models when fitting data with such low phase coverage.

6.3. Statistical Context

Basing an objective prior on experimental design is not an unprecedented idea in Bayesian statistical inference. Reference priors and the Jeffreys prior are objective priors that are defined by the structure of the likelihood. Like reference priors, observable-based priors are dependent on a form of the likelihood, with the goal that the resultant inference is maximally dominated by data (Berger et al. 2009). Observable-based priors differ from these paradigms in that they are not based on the asymptotic nature of an experiment. This protects against the statistical consequences of asymptotic priors, but also implies that our observable-based prior is weakly informative rather than truly objective (Gelman et al. 2017). There is some subjectivity added by the choice of the prior in observable space and the wide choice of conversions from higher-dimensional parameter space to lower-dimensional observable space (Equation (1)). For example, we transform from the observables to e and P for the reasons stated in Section 2, though we could in theory have transformed to e and T_o instead. In short, while including relevant information that can influence possible observations, we seek to limit prior influence in regions of prior dominance and thus maximize areas of data dominance. Because we base our prior analysis on possible data that can be observed (not in the asymptotic limit), we do not achieve the same objectivity as standard objective priors. However, because the prior's parameterization is determined by the observables, it is less subjective than are uniform priors. Thus, our prior analysis lies between that of

truly objective priors and uniform priors, which are commonly assumed in orbit fitting.

7. Conclusion

Data sets with low orbital phase coverage have low constraining power, and thus prior assumptions can bias parameter estimates, produce inaccurate confidence intervals, and profoundly impact inferred posteriors. To improve orbital estimates for objects with low phase coverage—in particular, stars in the Galaxy's central stellar cluster or directly imaged exoplanets—we develop a prior framework that is based on uniformity in observable space rather than in model parameter space. This observable-based prior limits prior influence and allows the data to contribute more heavily to resultant posteriors.

Compared to uniform priors, which are commonly assumed in orbit fitting, the observable-based prior reduces biases in model parameters by up to a factor of two and ensures that credible intervals are not underestimated for simulated Galactic center data with less than $\sim$40% phase coverage. Applying the new prior to simulated HR 8799d data shows that the observable-based prior can mitigate the known issue that eccentricity estimates are biased high when data are not rigorously constraining.

While S0-2 astrometric and RV data have full phase coverage and thus high constraining power, HR 8799 astrometric data have low phase coverage without the additional constraints of RV data. Thus, Section 5.1 shows that S0-2's orbital solution derived with the new prior is consistent with that published by Boehle et al. (2016), while Section 5.2 shows that orbital solutions for the HR 8799 planets are impacted by the observable-based prior, thus influencing our physical interpretation of the system. By limiting prior influence in prior-dominated regions and allowing data to have stronger influence over inferred posteriors, we see stronger evidence for lower eccentricity orbits and for consistent inclinations of the four HR 8799 planets at $\sim$30° to within 1σ, and do not exclude the possibility of coplanarity.

There are innumerable forms that an observable-based prior can take, though we have only specified two in this work. The objective choice of the prior form can vary and thus should be tested for other models. This framework of prior creation and evaluation can and should be extended to different models and data sets to more accurately estimate orbits of objects with low phase coverage. Such applications include (but are not limited to) Galactic Center orbits, directly imaged planetary systems, and visual binaries.

Appendix A
Expected Information Sampling

Calculation of the expected information comes directly from Equations (41) and (42), which imply that the expected information is the expectation of $\log[\mathcal{P}(\mathcal{D}|\mathcal{M})/\mathcal{P}(\mathcal{D})]$ over the joint probability distribution $\mathcal{P}(\mathcal{D}, \mathcal{M})$, e.g.,

$$\mathcal{I} \approx \frac{1}{N} \sum_{i=0}^{N} \log\left[\frac{\mathcal{P}(\mathcal{D}_i|\mathcal{M}_i)}{\mathcal{P}(\mathcal{D}_i)}\right], \quad \mathcal{D}, \mathcal{M} \sim \mathcal{P}(\mathcal{D}, \mathcal{M}) \quad (43)$$

over some large N. Equation (43) implies a simple algorithm to calculate $\mathcal{I}$: iteratively sample $\mathcal{D}$ and $\mathcal{M}$ from $\mathcal{P}(\mathcal{D}, \mathcal{M})$ and calculate the average $\log[\mathcal{P}(\mathcal{D}|\mathcal{M})/\mathcal{P}(\mathcal{D})]$. Here, $\mathcal{P}(\mathcal{D}, \mathcal{M})$ is sampled by first sampling $\mathcal{M}$ from the prior, $\mathcal{P}(\mathcal{M})$, and then drawing a mock data set, $\mathcal{D}$, from the likelihood, $\mathcal{P}(\mathcal{D}|\mathcal{M})$. This is summarized in Algorithm 1.

Algorithm 1: Expected information sampling algorithm.
Draw samples, $\{\mathcal{M}\}$, from $\mathcal{P}(\mathcal{M})$
$n \longleftarrow 0$
while $n < N$ **do**
  Draw $\mathcal{M}_{\text{mock}}$ from $\{\mathcal{M}\}$
  Draw $\mathcal{D}_{\text{mock}}$ from $\mathcal{P}(\mathcal{D}|\mathcal{M}_{\text{mock}})$
  Find evidence, $\mathcal{P}(\mathcal{D}_{\text{mock}})$, of $\mathcal{P}(\mathcal{D}_{\text{mock}}|\mathcal{M})\mathcal{P}(\mathcal{M})$
  $v_n \longleftarrow \log\left[\frac{\mathcal{P}(\mathcal{D}_{\text{mock}} | \mathcal{M}_{\text{mock}})}{\mathcal{P}(\mathcal{D}_{\text{mock}})}\right]$
  $n \longleftarrow n + 1$
end
$\mathcal{I} \longleftarrow \frac{1}{N}\sum_i v_i$

Appendix B
HR 8799 Orbital Parameter Posteriors

Figures 11–14 show 1D marginalized posteriors and the prior for all orbital parameters for HR 8799b, c, d, and e, respectively. Prior choice affects the resulting posteriors, and thus must be taken into consideration. All probability densities, including the prior probability densities, are derived from Monte Carlo simulations. The Monte Carlo chains are available in the .tar.gz package.

Note that the form of the observable-based prior defined in Section 2.2 is approximated by summing the Jacobian over all epochs. For each individual epoch, the prior distribution in T_o should be uniform (modulo boundary effects) according to Equation (30). However, when summing over a finite number of epochs, the sum is not guaranteed to be flat, as is evident in Figures 11 and 13.
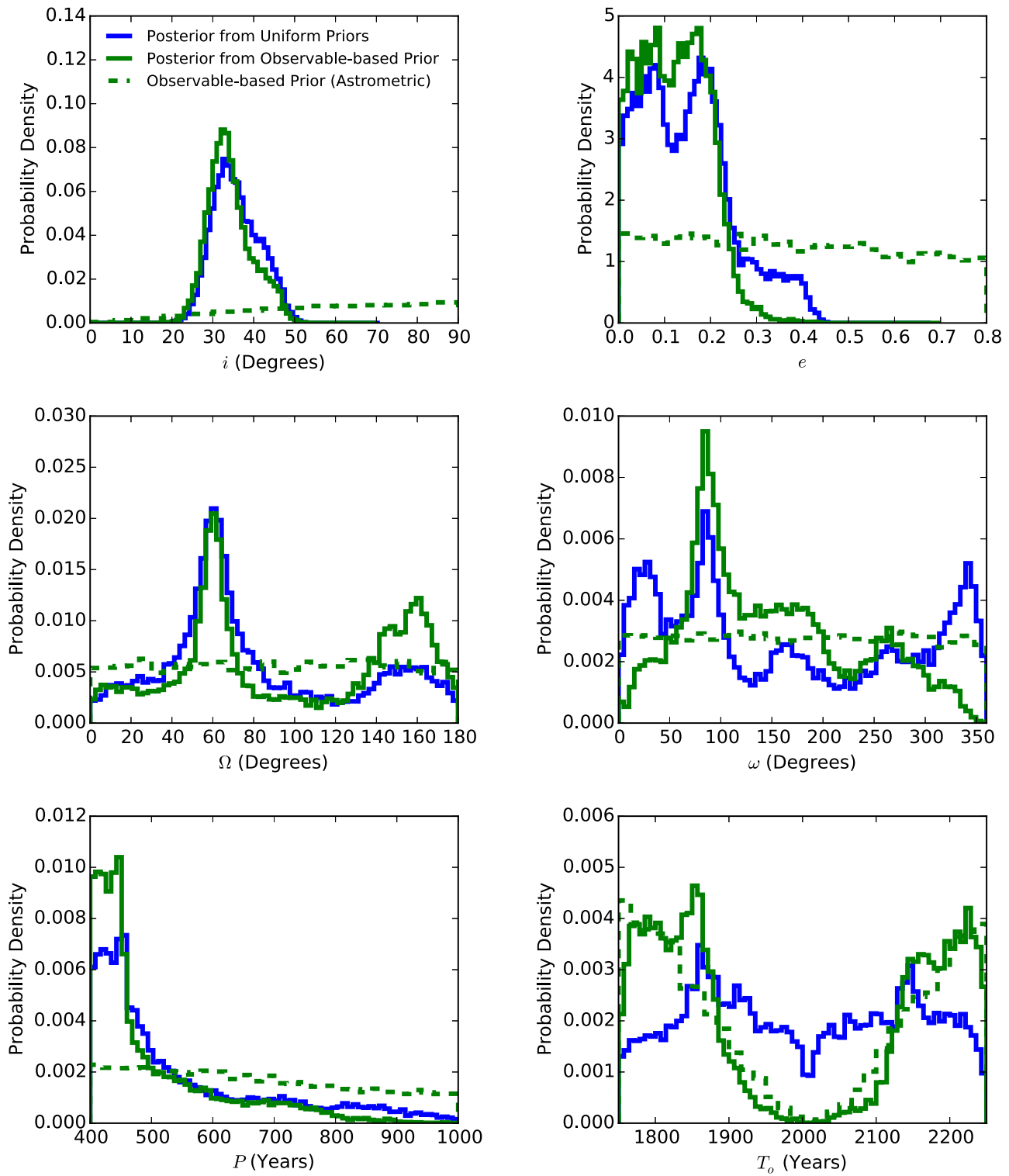
Figure 11. Marginalized 1D posteriors, assuming a uniform prior (blue) and a new prior based on uniformity in the astrometric observables (green) for all orbital parameters resulting from a fit to HR 8799b astrometric data from Konopacky et al. (2016). The observable-based prior itself is also shown with the dashed green line.
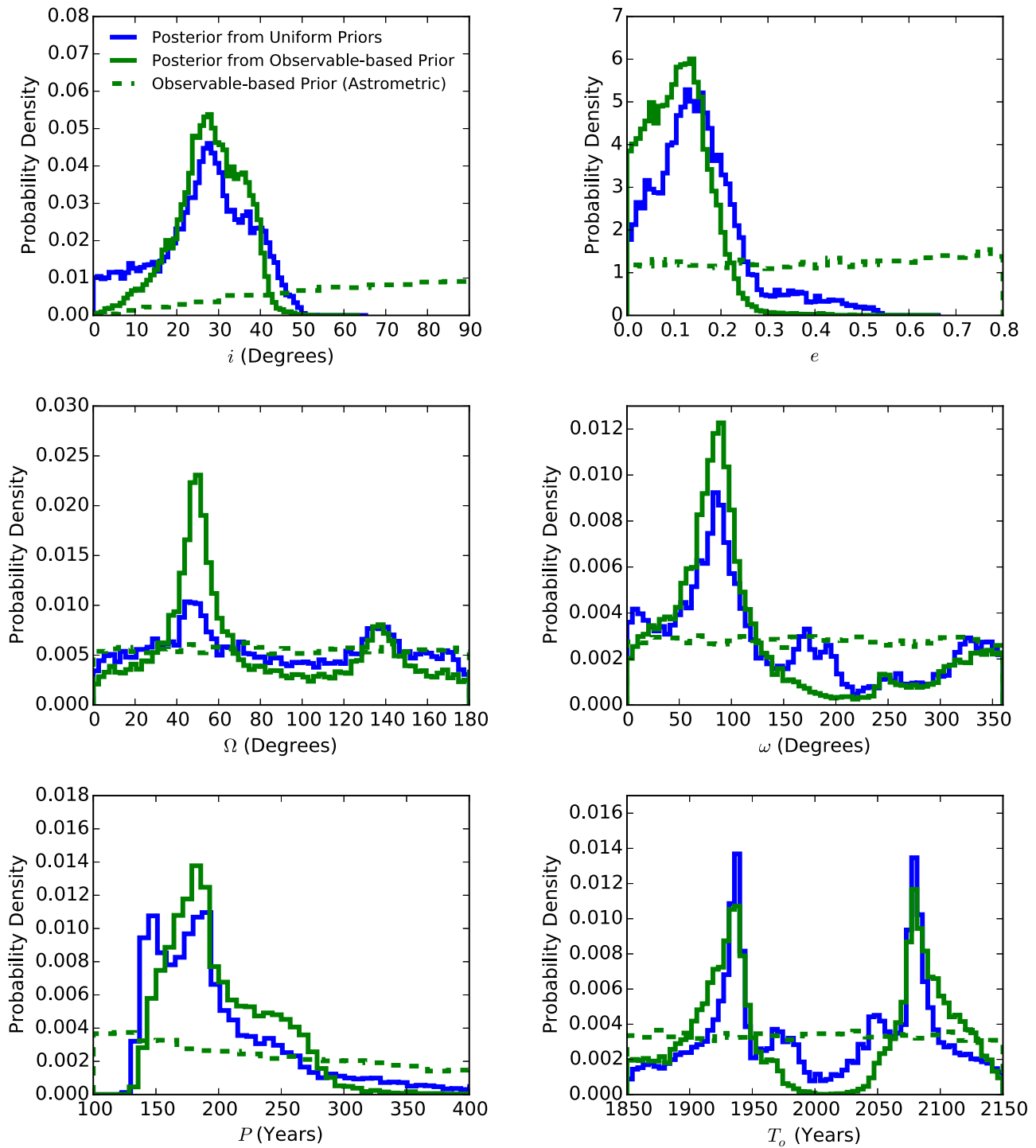
Figure 12. Marginalized 1D posteriors, assuming a uniform prior (blue) and a new prior based on uniformity in the astrometric observables (green) for all orbital parameters resulting from a fit to HR 8799c astrometric data from Konopacky et al. (2016). The observable-based prior itself is also shown with the dashed green line.
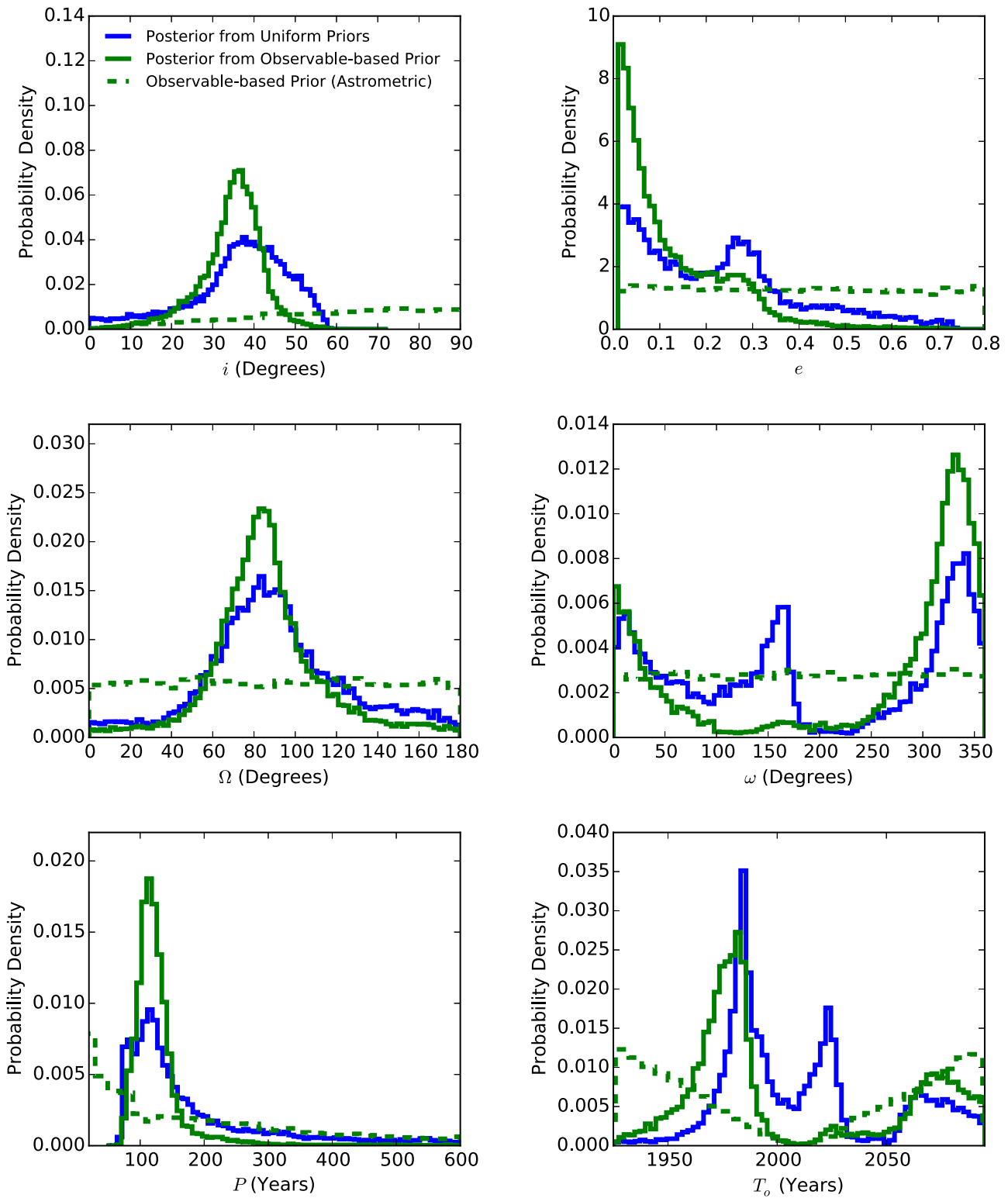
Figure 13. Marginalized 1D posteriors, assuming a uniform prior (blue) and a new prior based on uniformity in the astrometric observables (green) for all orbital parameters resulting from a fit to HR 8799d astrometric data from Konopacky et al. (2016). The observable-based prior itself is also shown with the dashed green line.
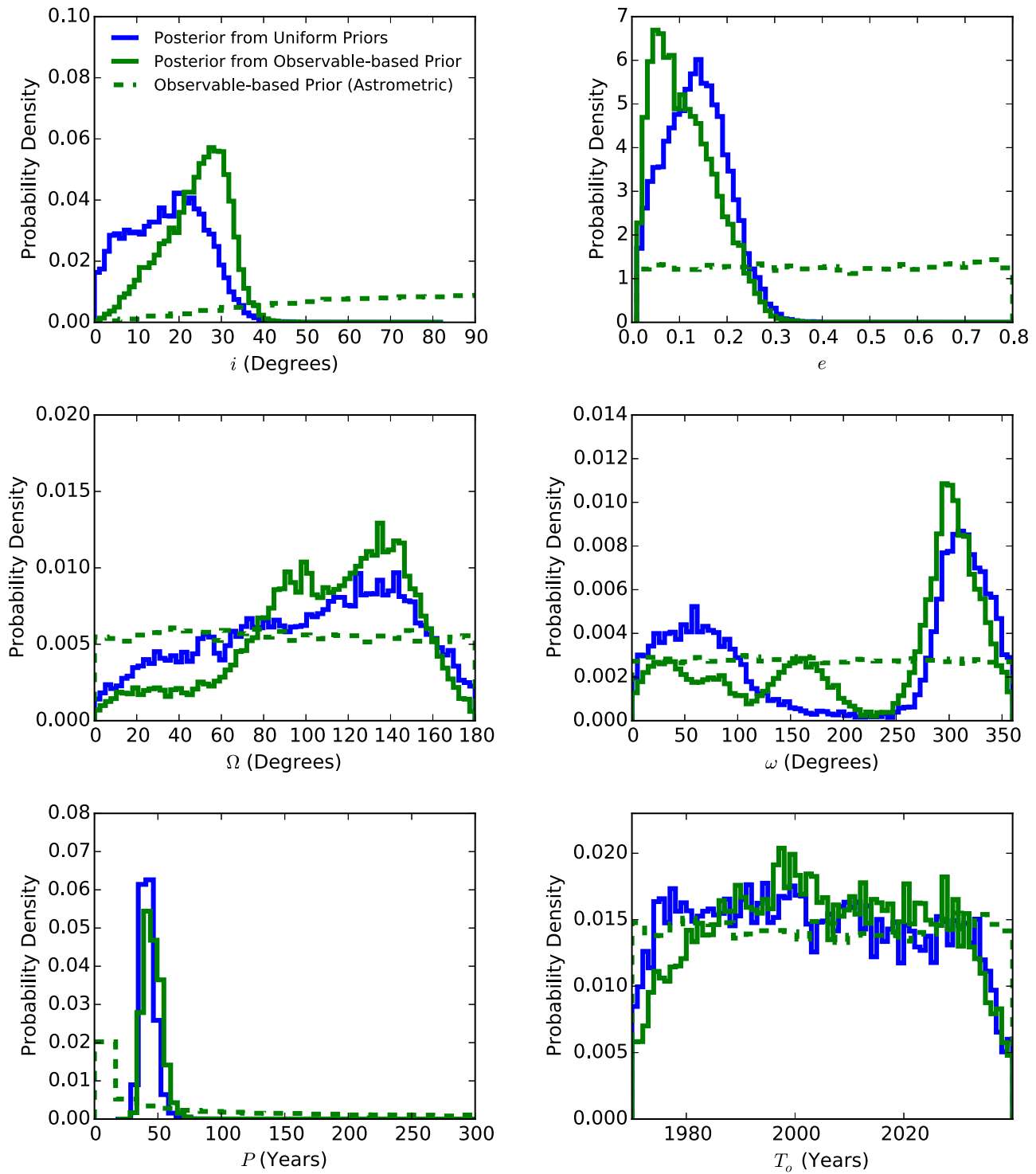
Figure 14. Marginalized 1D posteriors, assuming a uniform prior (blue) and a new prior based on uniformity in the astrometric observables (green) for all orbital parameters resulting from a fit to HR 8799e astrometric data from Konopacky et al. (2016). The observable-based prior itself is also shown with the dashed green line.

ORCID iDs

K. Kosmo O'Neil ⓘ https://orcid.org/0000-0003-2400-7322
A. Hees ⓘ https://orcid.org/0000-0002-2186-644X
A. M. Ghez ⓘ https://orcid.org/0000-0003-3230-5055
T. Do ⓘ https://orcid.org/0000-0001-9554-6062
G. Witzel ⓘ https://orcid.org/0000-0003-2618-797X
Q. Konopacky ⓘ https://orcid.org/0000-0002-9936-6285
D. S. Chu ⓘ https://orcid.org/0000-0003-3765-8001

J. R. Lu ⓘ https://orcid.org/0000-0001-9611-0009
S. Sakai ⓘ https://orcid.org/0000-0001-5972-663X

References

Baines, E. K., White, R. J., Huber, D., et al. 2012, ApJ, 761, 57
Berger, J. O., Bernardo, J. M., & Sun, D. 2009, Ann. Statist., 37, 905
Bernardo, J. M. 2005, in Handbook of Statistics, ed. D. K. Dey & C. R. Rao (Amsterdam: Elsevier), 17

Boehle, A., Ghez, A. M., Schödel, R., et al. 2016, ApJ, 830, 17
Brown, W. R. 2015, ARA&A, 53, 15
Cameron, E. 2011, PASA, 28, 128
Chauvin, G., Lagrange, A.-M., Beust, H., et al. 2012, A&A, 542, A41
Chen, X., & Amaro-Seoane, P. 2014, ApJL, 786, L14
Currie, T., Burrows, A., Itoh, Y., et al. 2011, ApJ, 729, 128
Currie, T., Fukagawa, M., Thalmann, C., Matsumura, S., & Plavchan, P. 2012, ApJL, 755, L34
Eckart, A., Genzel, R., Ott, T., & Schödel, R. 2002, MNRAS, 331, 917
Eisenhauer, F., Schödel, R., Genzel, R., et al. 2003, ApJL, 597, L121
Esposito, S., Mesa, D., Skemer, A., et al. 2013, A&A, 549, A52
Feroz, F., & Hobson, M. P. 2008, MNRAS, 384, 449
Feroz, F., Hobson, M. P., & Bridges, M. 2009, MNRAS, 398, 1601
Ford, E. B. 2006, ApJ, 642, 505
Gelman, A., Simpson, D., & Betancourt, M. 2017, Entrp, 19, 555
Genzel, R., Eisenhauer, F., & Gillessen, S. 2010, RvMP, 82, 3121
Ghez, A. M., Becklin, E., Duchjne, G., et al. 2003, ANS, 324, 527
Ghez, A. M., Morris, M., Becklin, E. E., Tanner, A., & Kremenek, T. 2000, Natur, 407, 349
Ghez, A. M., Salim, S., Hornstein, S. D., et al. 2005, ApJ, 620, 744
Ghez, A. M., Salim, S., Weinberg, N. N., et al. 2008, ApJ, 689, 1044
Gillessen, S., Eisenhauer, F., Trippe, S., et al. 2009, ApJ, 692, 1075
Gillessen, S., Plewa, P. M., Eisenhauer, F., et al. 2017, ApJ, 837, 30
Grould, M., Vincent, F. H., Paumard, T., & Perrin, G. 2017, A&A, 608, A60
Hartigan, J. 1964, Ann. Math. Statist., 35, 836
Hartkopf, W. I., McAlister, H. A., & Franz, O. G. 1989, AJ, 98, 1014
Hees, A., Do, T., Ghez, A. M., et al. 2017, PhRvL, 118, 211101
Hills, J. G. 1988, Natur, 331, 687
Host, O., Lahav, O., Abdalla, F. B., & Eitel, K. 2007, PhRvD, 76, 113005

Jaynes, E. T. 1968, IEEE Trans. Systems, Science and Cybernetics, 4, 227
Jeffreys, H. 1946, RSPSA, 186, 453
Konopacky, Q. M., Marois, C., Macintosh, B. A., et al. 2016, AJ, 152, 28
Kullback, S., & Leibler, R. A. 1951, Ann. Math. Statist., 22, 79
Lindley, D. V. 1956, Ann. Math. Statist., 27, 986
Lucy, L. B. 2014, A&A, 563, A126
Madigan, A.-M., Pfuhl, O., Levin, Y., et al. 2014, ApJ, 784, 23
Maire, A.-L., Skemer, A. J., Hinz, P. M., et al. 2015, A&A, 579, C2
Marois, C., Macintosh, B., Barman, T., et al. 2008, Sci, 322, 1348
Marois, C., Zuckerman, B., Konopacky, Q. M., Macintosh, B., & Barman, T. 2010, Natur, 468, 1080
Meyer, L., Ghez, A. M., Schödel, R., et al. 2012, Sci, 338, 84
Morris, M. 1993, ApJ, 408, 496
Neyman, J. 1937, RSPTA, 236, 333
Neyman, J., & Scott, E. L. 1948, Econometrica, 16, 1
Parsa, M., Eckart, A., Shahzamanian, B., et al. 2017, ApJ, 845, 22
Perets, H. B., Gualandris, A., Kupi, G., Merritt, D., & Alexander, T. 2009, ApJ, 702, 884
Perets, H. B., Hopman, C., & Alexander, T. 2007, ApJ, 656, 709
Pierini, M., Prosper, H. B., Sekmen, S., & Spiropulu, M. 2011, arXiv:1107.2877
Pueyo, L., Soummer, R., Hoffmann, J., et al. 2015, ApJ, 803, 31
Rameau, J., Nielsen, E. L., De Rosa, R. J., et al. 2016, ApJL, 822, L29
Schödel, R., Ott, T., Genzel, R., et al. 2002, Natur, 419, 694
Schödel, R., Ott, T., Genzel, R., et al. 2003, ApJ, 596, 1015
van Leeuwen, F. 2007, A&A, 474, 653
Wertz, O., Absil, O., Gómez González, C. A., et al. 2017, A&A, 598, A83
Wright, J. T., & Howard, A. W. 2009, ApJS, 182, 205
Zurlo, A., Vigan, A., Galicher, R., et al. 2016, A&A, 587, A57