The Astronomical Journal, 157:188 (9pp), 2019 May https://doi.org/10.3847/1538-3881/ab0f2e


# Three Direct Imaging Epochs Could Constrain the Orbit of Earth 2.0 inside the Habitable Zone

Claire Marie Guimond[1] and Nicolas B. Cowan[1,2]
[1] Department of Earth & Planetary Sciences, McGill University, 3450 rue University, Montréal, QC, H3A 0E8, Canada; claire.guimond2@mail.mcgill.ca
[2] Department of Physics, McGill University, 3600 rue University, Montréal, QC, H3A 2T8, Canada
*Received 2018 December 24; revised 2019 March 10; accepted 2019 March 11; published 2019 April 22*

## Abstract

Space-based direct imaging missions (HabEx, LUVOIR) would observe reflected light from exoplanets in the habitable zones of Sun-like stars. The ultimate—but not sole—goal of these concept missions is to characterize such planets. Knowing an exoplanet's orbit would help twofold: (i) its semimajor axis informs whether the planet might harbor surface liquid water, making it a priority target; and (ii) predicting the planet's future location would tell us where and when to look. The science yields of HabEx and LUVOIR depend on the number, cadence, and precision of observations required to establish a planet's orbit. We produce mock observations using realistic distributions for the six Keplerian orbital parameters, experimenting with both beta and uniform eccentricity distributions, and accounting for imperfect astrometry ($\sigma = 3.5$ mas) and obscuration due to the inner working angle of a high-contrast imaging system (inner working angle = 31 mas). Using Markov chain Monte Carlo methods, we fit the orbital parameters, and retrieve their average precisions and accuracies as functions of cadence, number of epochs, and distance to the target. Given the time at which it was acquired, each image provides two data: the $x$ and $y$ position of the planet with respect to its star. Parameter retrieval based on one or two images is formally underconstrained, yet the semimajor axis posterior can be obtained semi-analytically. For a planet at 1 au around a star at a distance of 10 pc, three epochs constrain the semimajor axis to within $\lesssim$5%, if each image is taken at least 90 days apart.

*Key words:* planets and satellites: detection – planets and satellites: terrestrial planets – telescopes

## 1. Introduction

The first generation of space-based direct imaging missions would observe planets in the habitable zones (HZs) of Sun-like stars. As recommended by the National Academies of Sciences, Engineering, and Medicine's recent Exoplanet Science Strategy, "NASA should lead a large strategic direct imaging mission capable of measuring the reflected-light spectra of temperate terrestrial planets orbiting Sun-like stars."[3] To acquire these spectra, we must first plan the missions cost effectively. Yet identifying contenders for Earth-like planets is not straightforward with direct imaging alone. A single direct imaging observation would distinguish Earths from similarly bright sub-Neptunes with a one-in-five success rate, due to degeneracies between radius, albedo, and phase. If the planet's semimajor axis is known, on the other hand, the success rate is one in two (Guimond & Cowan 2018).

In this paper we help prepare for future direct imaging missions such as the Habitable Exoplanet Observatory (HabEx) and the Large UV/Optical/IR Surveyor (LUVOIR) by determining how many visits are required to each newfound planet before we can confidently know its semimajor axis. In doing so, we hope to quantify a key input parameter in more comprehensive mission simulations (see Stark et al. 2018).

### 1.1. Images to Orbits

A single epoch of direct imaging provides the two-dimensional position of the planet in the sky plane relative to its host star: $x$ and $y$, where the star is at the origin. Six parameters are needed to uniquely describe the three-dimensional Keplerian orbit of a planet (listed in Table 1). Given the time at which it was taken, each image constrains $x$ and $y$, so at most three measurements should be required to fit all six parameters.

The fitting methods employed in previous direct imaging orbit-retrieval efforts come from either Markov chain Monte Carlo (MCMC) or Bayesian rejection sampling (De Rosa et al. 2015; Pueyo et al. 2015; Rameau et al. 2016; Wang et al. 2016, 2018; Blunt et al. 2017; Kosmo O'Neil et al. 2018). The latter type are carefully demonstrated in Blunt et al. (2017) for closely spaced observations over tiny orbital coverage (~3%).

Mede & Brandt (2017) present a software package to simultaneously fit direct imaging observations and radial velocity observations. Such an approach may be fruitful for eventual Earth twins if high-precision radial velocity can be sensitive to the 10 cm s$^{-1}$ signals (Fischer et al. 2016).

As for future direct imaging, previous design reference missions for HabEx and LUVOIR have not always implemented an optimized number or cadence of observations. Stark et al. (2014) look only at single-visit yields; Stark et al. (2015, 2016) do not consider that a mission would revisit candidates for the strict purpose of establishing their orbits, assuming that most stars would be revisited regardless to increase the total yield.

In this paper, we anticipate space-based observations from HabEx and LUVOIR. The work presented here is fundamentally distinct from these earlier efforts in that we explicitly quantify the number, cadence, and precision of observations required to establish, for targets at any distance, the orbit of a planet within its star's HZ.

### 1.2. A Note on Habitable Zones

The circumstellar HZ is a theoretical shell around a star within which planets can harbor liquid water at their surfaces.

[3] https://www.nap.edu/read/25187/

**Table 1**
The Keplerian Orbital Parameters and Their Priors

| Name | Symbol | Prior Distribution | Prior Range | Input |
|---|---|---|---|---|
| Semimajor axis | $a$ | Uniform in natural log | [0.01, 50] au | [0.95, 1.70] au |
| Eccentricity | $e$ | Beta ($\sigma = 0.081$) or uniform | [0, 1) | [0, 1) |
| Inclination | $i$ | Uniform in cosine | [0, $\pi$] | [0, $\pi/2$] |
| Argument of periapsis | $\omega_p$ | Uniform | [0, $2\pi$) | [0, $2\pi$) |
| Longitude of the ascending node | $\Omega$ | Uniform | [0, $2\pi$) | [0, $2\pi$) |
| Mean anomaly at first epoch | $M_0$ | Uniform | [0, $2\pi$) | [0, $2\pi$) |

**Note.** The prior range refers to the allowable parameter space explored in the Markov chain, while truth values are drawn from the input range.

The inner and outer edges of the HZ have been variously modeled based on assumptions about climatic systems. Early HZ estimates (Hart 1979) were physically driven by the destabilizing ice-albedo and water vapor feedbacks and were narrower than in current conceptions.

The HZ estimates we use come from physics gleaned from the Earth. Assuming the surface of a planet has some exposed silicate rock, $CaSiO_3$ minerals in the rock will react to consume atmospheric $CO_2$. The rate of this reaction increases exponentially with temperature, so the planetary effect of warming the atmosphere is to sequester more carbon and weaken the greenhouse effect (Walker et al. 1981; Kasting 1988; Kasting & Toon 1989). This is the well-known silicate weathering feedback, and is the basis of the HZs proposed in Kasting et al. (1993) and updated in Kopparapu et al. (2013).

This conception of the HZ has yet to be empirically validated (but it could through the statistical study of exoplanet properties; Bean et al. 2017). If a planet does not have exposed surface rock, for example, then the HZs given by a model based on the silicate weathering feedback might not apply (e.g., Abbot et al. 2012). We therefore base our analysis on a variety of hypothetical HZ widths, defined as 0.01, 0.05, 0.1, 0.25, and 0.5 au. For completeness, we also consider the inner and outer HZ limits proposed by Kopparapu et al. (2013). The semimajor axis precision that would satisfy us depends on the sizes of these HZs relative to our semimajor axis estimate.

## 2. The Orbit-retrieval Model

This work quantifies, statistically, the orbit-retrieval accuracy and precision for many simulated HZ planets. We simulate planets with each orbital element randomized according to the distributions in Table 1, and assuming one planet per star. For each epoch of each planet, we calculate the planet's ($x$, $y$) position relative to its star, given a fixed cadence in days and with added Gaussian noise—here we adopt $\sigma_\theta = 3.5$ mas as the baseline astrometric precision (HabEx Team 2018; LUVOIR Team 2018). Then we retrieve the posteriors on orbital elements from our synthetic data set using an MCMC. Figure 1 depicts an example retrieval. We repeat this numerical experiment under varying assumptions, as described below.

*Eccentricity distribution*. Since the eccentricity distribution of HZ terrestrial exoplanets is unconstrained, we repeat our experiment assuming different eccentricity distributions—a beta distribution with parameters $a = 0.867$, $b = 3.03$ (Nielsen et al. 2008; Kipping 2013), or a uniform distribution—for both the underlying true distribution and the prior distribution used in the retrieval.

*Cadence*. The retrieval experiments are repeated, using the same synthetic orbits, for cadences of 30, 90, 180, and 270 days. These cadences were chosen as a starting point because they are well-separated round numbers. This part of the study reveals which cadence gives the best fit, for the same number of epochs. Our result is optimized for orbital periods within the G2V HZ; we expect the best cadence to be set by how much the planet has moved, which is some fraction of the period. In principle, a mission simulation would consider data from previous epochs to choose when to revisit a star. We take the simpler approach of a fixed cadence to obtain a broad result, which could conveniently be applied to mission prognoses.

*Distance*. The distance of the target star will impact orbit retrieval through astrometric error and the inner working angle (IWA) of starlight suppression, both of which correspond to greater projected separations at greater distances. We take a baseline distance of 10 pc, and repeat our experiment for distances of 5 and 20 pc.

Each scenario considers 100 planets with orbital parameters $a$, $i$, $e$, $\omega_p$, $\Omega$, and $M_0$, sampled from the underlying distributions (Table 1). These are the prior distributions described in Section 2.4—except $e$, which we allow to have different underlying and prior distributions, as detailed later. Our sample size of 100 planets is justified in that halving this number gives similar results.

To retrieve posterior probability distributions of the orbital parameters, we use `emcee`, an MCMC ensemble sampler (Foreman-Mackey et al. 2013). The goal of any MCMC implementation is to evaluate the posterior Prob (model|observation, $\sigma$) $\propto$ Prob(observation|model, $\sigma$) $\times$ Prob (model), where $\sigma$ is the measurement precision. Given estimates of the right-hand side, we can estimate the left-hand side. Thus setting up `emcee` requires:

1. Estimates of the prior probability distribution, Prob (parameter), for each orbital parameter,
2. a forward model that calculates the planet positions ($x_0$, $y_0$), ($x_1$, $y_1$), ..., ($x_k$, $y_k$) up to the $k$th epoch given orbital parameters and $t_0$, $t_1$ ... $t_k$,
3. a likelihood function, Pr(observation|parameters, $\sigma$), that calculates the probability of the observed ($x$, $y$) position given the forward model.

The MCMC uses 30 walkers, randomly initialized in a Gaussian ball around the best-fit parameters from a quick `scipy.optimize` likelihood maximization. The walkers have a burn-in time of 1000 steps and run for up to $5 \times 10^5$ steps. The rest of the MCMC setup is detailed in Sections 2.1 through 2.4.

### 2.1. Likelihood Function for Detections

The likelihood of the observed position of the planet at some epoch is a normal distribution centered on the true position

with a width, $\sigma_{xy}$. If we have $k$ images in which the planet is detected, then the log likelihood is

$$\ln L = -\frac{\chi^2}{2} - k(\ln \pi + 2 \ln \sigma_{xy}), \quad (1)$$

where

$$\chi^2 = \sum_{i=1}^{k} \frac{(x_{\mathrm{obv},i} - x_{\mathrm{model},i})^2 + (y_{\mathrm{obv},i} - y_{\mathrm{model},i})^2}{\sigma_{xy}^2}. \quad (2)$$

### 2.2. Likelihood Function for Non-detections

The first epoch is—by definition—a detection, as we are concerned with the orbit retrieval of actual planets rather than the odds of spotting a planet in the first place. Subsequent epochs may yield non-detections, however.

A planet may go undetected in an image for two reasons: (i) it is imaged at small projected separation and is occluded along with the starlight, or (ii) it does not reflect enough light. We ignore the latter case as the brightness of the planet depends on the geometric albedo and the scattering phase function, about which we presume nothing. Titan, for example, appears brighter at the crescent phase, due to a distinctly non-Lambertian phase function (García Muñoz et al. 2017).

The relevant criterion in (i) is the IWA of the telescope, defined as the point where photon transmission through the instrument has decreased by 50%. We adopt an IWA of 31 mas. For a planet of given size, albedo, and phase function, and given a contrast floor, there will be a unique IWA at which it becomes undetectable at the gibbous phase, and another where it becomes undetectable at the crescent phase. We ignore brightness information and hence adopt a single hard-cutoff IWA.

A non-detection therefore still provides a constraint; it means that the angular separation of the planet and the star is less than the IWA. In this way, a non-detection is analogous to a measurement centered at the origin with an uncertainty of $\theta_{\mathrm{IWA}}$. We use this information in our retrieval exercises. If the planet is not detected at some epoch, then any set of parameters that would place it outside the IWA are assigned zero likelihood, Prob(observation|model, $\sigma$) = 0. If an orbit-fitting model were incorporating photometry, then a soft-edge IWA would be more appropriate; for example, the non-detection framework in Ruffio et al. (2018). The width of this soft-edge matters when one knows the orbit and is trying to measure the brightness of the planet, which will vary from 0% to 100% of its true value across the soft edge. However, we are interested in the earlier stage, when we do not yet know the orbit, and do not strictly care about the measured brightness of the planet. In other words, there is no need for a soft edge at this point because the probability of the planet being detected outside the IWA is always unity, although this statement requires assumptions discussed further in Section 4.5.

### 2.3. Posterior Probability for Less Than Three Detections

Given only one or two detections, we expect the MCMC to converge slowly, if at all, since the problem is formally underconstrained. The walkers would be exploring a nearly flat plane of probability; we would not believe the width of their retrieved distribution. Thus we eschew Kepler's laws for a naive, semi-analytic approach, as long as the number of detections is less than three.

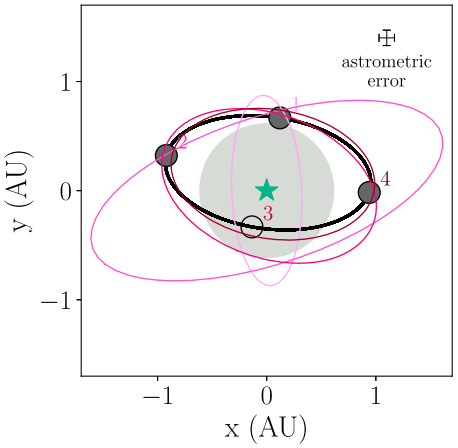


**Figure 1.** Demonstration of orbit retrieval. The true orbit (bold black ellipse) is shown to scale with the region obscured by the inner working angle (gray circle). The true location of the planet at each epoch is shown as a small circle with the radius corresponding to the astrometric error; the hollow circle at the third epoch indicates that the planet is inside the inner working angle and is not detected. The retrieved orbits after each epoch are shown from light (epoch 1) to dark (epoch 4). For this nominal planet $r = 20$ pc, $\sigma_\theta = 3.5$ mas, and the observations are spaced 90 days apart.

That is, for a given detection, the posterior probability distribution of $\ln(a)$ is a strong function of the projected separation, $a_{\mathrm{proj}}$. For circular orbits we can express this analytically as

$$\mathrm{Prob}(a|a_{\mathrm{proj}}, e = 0) = \frac{1}{\beta^2\sqrt{\beta^2 - 1}}, \quad \beta \in [1, \infty), \quad (3)$$

where $\beta = a/a_{\mathrm{proj}}$. Nonzero eccentricities complicate this analytic distribution, but we can model it numerically for given values of $a_{\mathrm{proj}}$ and $a_{\mathrm{IWA}}$ by drawing 1000 random orbits from the priors in Table 1 with $a_{\mathrm{proj}}$ fixed at the observed value. We keep drawing all six Keplerian parameters until we have 1000 orbits with the desired $a_{\mathrm{proj}}$. As Figure 2 shows, the distribution of semimajor axes of these orbits peaks at the true $a$, and is hence a reasonable proxy for the posterior, for a given measurement of $a_{\mathrm{proj}}$. If a planet is not detected at some epoch, we also simulate random orbits but with $a_{\mathrm{proj}} \in (0, a_{\mathrm{IWA}}]$.

For two or more epochs, we construct independent posteriors for each $a_{\mathrm{proj}}$ observation and multiply them together. The resulting joint probability distribution usually has one peak. Of course these measurements would not be truly independent, since they are linked by a Keplerian orbit. Thus our semi-analytic posterior distributions give an upper limit on the true uncertainty. This semi-analytic method is exchanged for a full MCMC once we have three detections, allowing the latter algorithm to converge. Despite the superior speed of the semi-analytic method, it does not use information about the relative timing of the different epochs, and hence does not fully leverage the Keplerian orbit. Figure 3 demonstrates the relatively poor performance of the semi-analytic method compared to the MCMC after three epochs.

### 2.4. Prior Distributions

As summarized in Table 1, parameterizations are chosen such that all priors save for eccentricity are flat. The log-uniform prior on the semimajor axis is roughly consistent with recent literature (Petigura et al. 2013; Foreman-Mackey

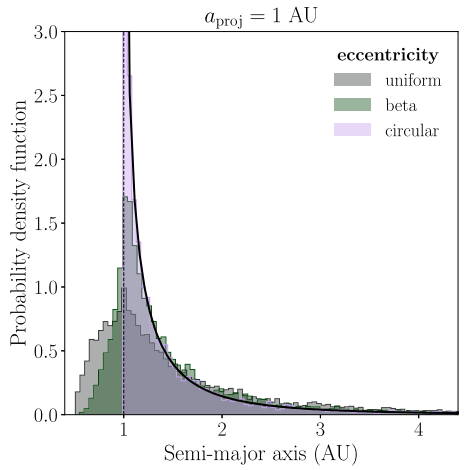


**Figure 2.** Probability of the semimajor axis given a projected separation. Normalized histograms of the semimajor axes of $10^3$ simulated planets with $a_{\rm proj} = 1$ au (dashed black line) and random $a$, $i$, $e$, $\omega_p$, $\Omega$, and $M$. The colors indicate the different underlying eccentricity distributions: lavender for circular, green for beta (Nielsen et al. 2008; Kipping 2013), and gray for uniform. The thick black line traces the analytic derivation for the circular case (Equation (3)).

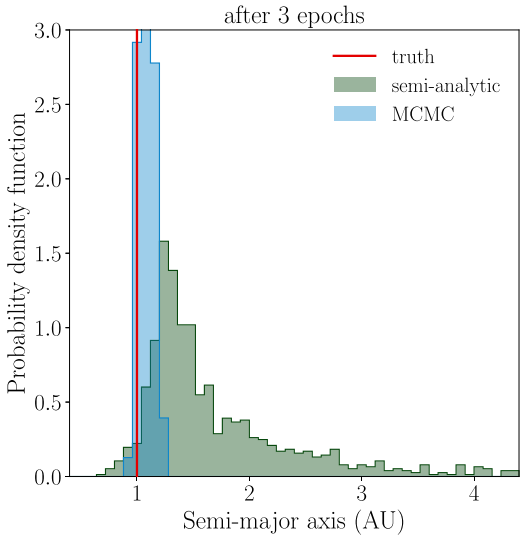


**Figure 3.** Comparison of semi-analytic and MCMC posterior retrievals. Green: normalized joint probability density resulting from multiplying together the three histograms (Figure 2) for three independent observations of one planet's $a_{\rm proj}$, assuming 90 days between each image. Blue: posterior probability density after three simulated observations of the same system using an MCMC. The vertical red line shows the true semimajor axis. The improved performance of the MCMC comes from the added use of Keplerian orbits over merely using geometry.

et al. 2014; Burke et al. 2015; Christiansen et al. 2015; Silburt et al. 2015; Kopparapu et al. 2018). We adopt a wide prior range of $a \in [0.1, 50]$ au.

The mean anomaly at the time of the first epoch is uniformly distributed in $[0, 2\pi)$, as are arguments of periapsis and longitudes of the ascending node. The inclination is uniform in $\cos i \in [-1, 1]$.

Eccentricity is drawn from a beta distribution with parameters $a = 0.867$, $b = 3.03$, as determined for radial velocity planets by Kipping (2013) and in agreement with Nielsen et al. (2008). Because this parameter's underlying distribution is especially hard to constrain, we repeat our experiment with a uniform prior and uniformly distributed true eccentricities, and again with a beta prior and uniformly distributed true eccentricities (where we would be overconfident).

## 3. Results

Figure 1 illustrates the results of our orbit retrieval for five epochs of a planet at 20 pc with a 90-day cadence. This particular planet goes undetected in the second epoch, yet we still see an increased similarity between the retrieved and true orbits. In this example, the best-fit orbit after each epoch is the Markov chain link with the maximum posterior probability.

### 3.1. Semimajor Axis Retrieval under Different Eccentricity Scenarios

The three panels in Figures 4 and 5 show the accuracy and precision—respectively—of semimajor axis retrieval for different prior assumptions and underlying distributions of orbital eccentricity. The distributions are either both beta (realistic), both uniform (pessimistic), or beta-distributed in the prior and uniform in the underlying truths (i.e., a pathological scenario which may underestimate the error).

Figure 4 shows the accuracy of the retrieved $a$ estimates as a function of epochs for a sample of 100 planets. Errors are retrieved from semi-analytic posteriors of $a$ for the under-constrained first two epochs, while the MCMC is used for the third epoch and up. Accuracy is the difference between the median of the posterior and the true semimajor axis.

Figure 5 is complementary to Figure 4 and presents the precisions on the retrieved semimajor axes, along with the sample median of these precisions. Again, an MCMC is used for the third epoch on. Precision is the half-width of the $1\sigma$ confidence interval, and it shrinks with each additional observation. This is as expected by degrees of freedom: the inflection point near the third epoch indicates that here we begin to gain less precision with additional measurements.

Our baseline case is shown in the left panels of Figures 4 and 5. Here, by the second epoch we achieve quite good precision. At $1\sigma$ confidence, we can constrain $a$ to $<25\%$ with two epochs, and to $\lesssim 10\%$ with three epochs. These represent the average results across planets with true semimajor axes between 0.95 and 1.70 au.

The central subplots in Figures 4 and 5 demonstrate a case where our precision is overstated; the prior eccentricity distribution is narrower than the underlying distribution. As shown in Table 2, the standard deviation of the z-scores for this scenario of amiss mixed distributions is much greater than unity, which confirms the overstatement.[4] Therefore, if we were to assume a narrower eccentricity distribution than nature provides, we would not be able to believe our retrieved semimajor axis precisions.

The other endpoint draws eccentricity guesses from a uniform distribution (right panels of Figures 4 and 5). Despite this pessimistic prior, we can still constrain the semimajor axis to 25% by the third epoch, and by the fifth epoch retrieval accuracies are not much different from the best-case scenario. In practice, we would start with a uniform prior until we know better, but the left panel is realistic in the steady state.

[4] The z-score of a fit, defined here as the ratio between our accuracy and precision, has a standard deviation of about unity if the fit precision reliably encompasses the true value and the retrieved values are Gaussian.

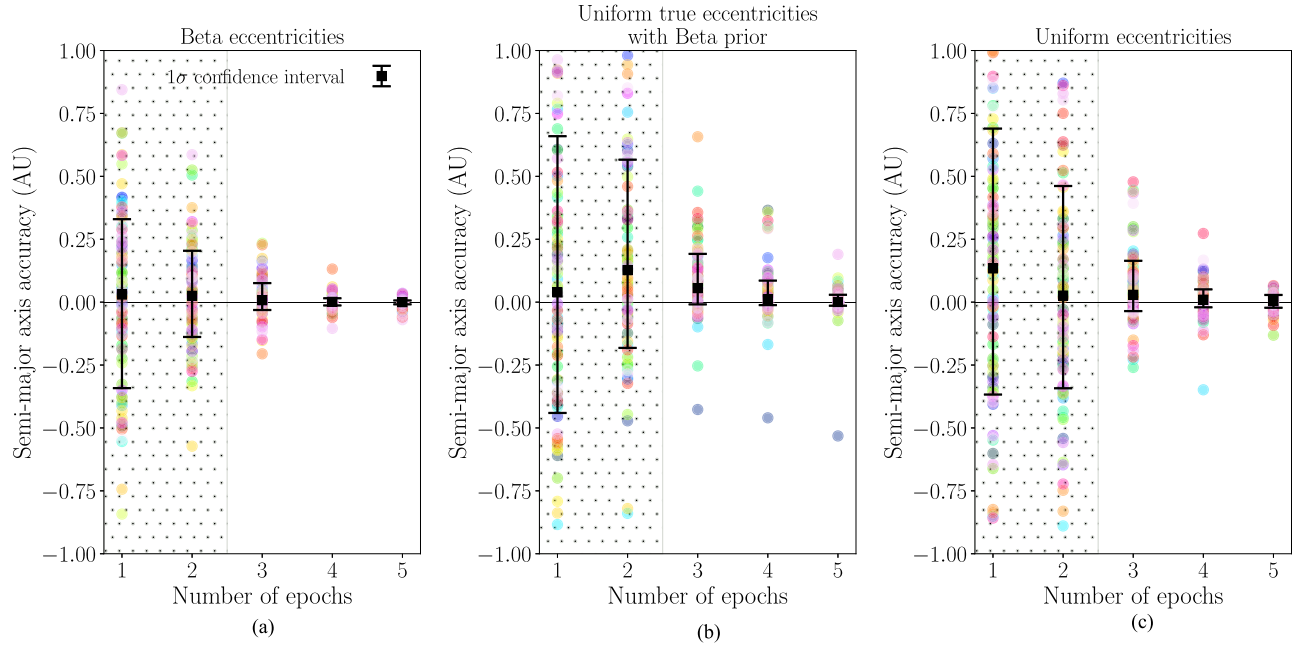


**Figure 4.** Semimajor axis accuracy by epoch. The epoch-dependence of accuracy is defined as ($a_{\rm fitted} - a_{\rm true}$), where $a_{\rm fitted}$ is the median of the posterior semimajor axis distribution from the MCMC. All simulated planets (100 in each scenario) have a distance of 10 pc and are observed every 90 days. Each dot color represents one planet. The black error bars show the 16th and 84th percentiles of the retrieved accuracy (i.e., the standard error). Until a planet is detected three times, orbits are retrieved with a semi-analytic method as opposed to the MCMC and represent a conservative upper limit; the hatched region marks off this distinction. Eccentricities are either (a) drawn from a low-dispersion beta distribution for both the underlying true orbits and the prior, (b) drawn from a uniform distribution for the underlying truths and a beta distribution for the prior, or (c) drawn from a uniform distribution for both the input and prior.

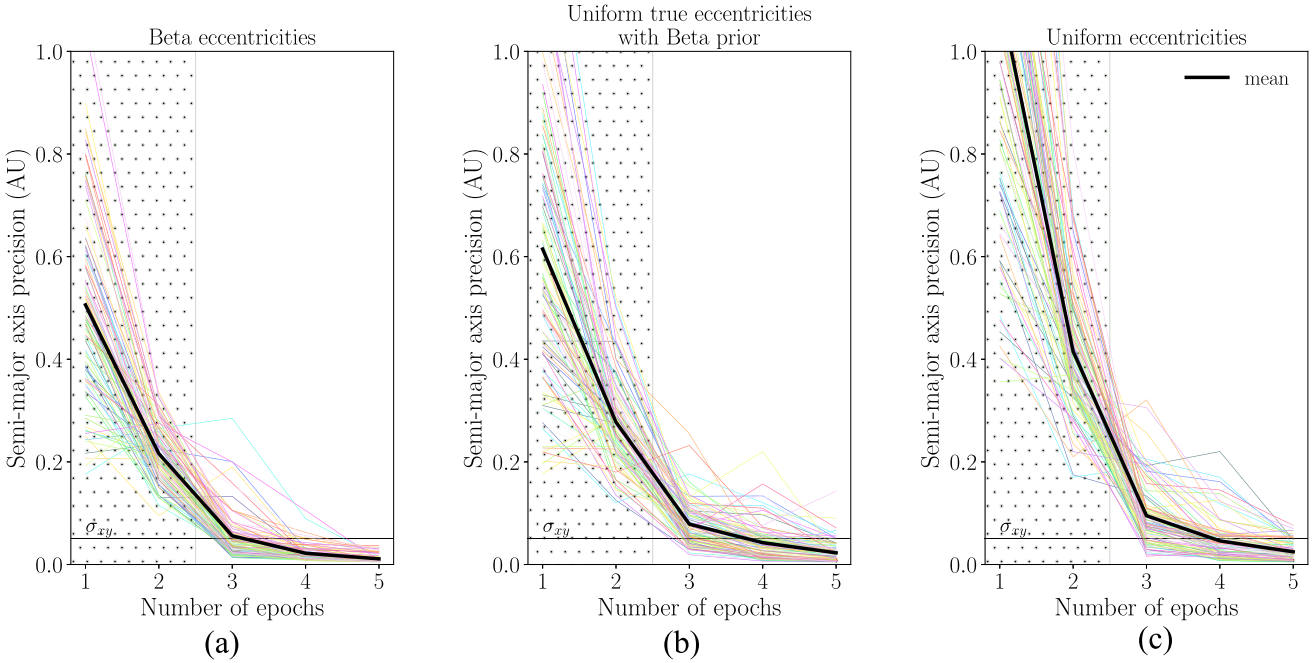


**Figure 5.** 1$\sigma$ precision on the semimajor axis retrieval as a function of the number of epochs. The colored lines show the 100 individual runs using the same scheme as in Figure 4, while the bold black line is the mean precision (the median is slightly better). Precision is defined as ($a_{+1\sigma} - a_{-1\sigma}$)/2, where the subscripts refer to the retrieved upper and lower limits of the 68% confidence interval. The horizontal line indicates the measurement uncertainty for 3.5 mas astrometry at 10 pc. Until a planet is detected three times, fits are done with a semi-analytic method as opposed to the MCMC and represent a conservative upper limit; the hatched region marks off this distinction. In this scenario, planets are observed every 90 days, and eccentricities are either (a) drawn from a low-dispersion beta distribution for both the underlying true orbits and the prior, (b) drawn from a uniform distribution and the prior is a beta distribution, or (c) drawn from a uniform distribution and the prior is uniform. Around the third epoch, the problem becomes constrained, and we see an inflection point in precision.

**Table 2**
Z-score Statistics for Data in Figures 4 and 5

| | (a) Beta | | (b) Mixed | | (c) Uniform | |
|---|---|---|---|---|---|---|
| | Mean | Standard deviation | Mean | Standard deviation | Mean | Standard deviation |
| Epoch 3 | −0.47 | 1.19 | −1.02 | 2.56 | −0.18 | 1.26 |
| Epoch 4 | −0.10 | 0.99 | −0.28 | 2.59 | −0.14 | 1.11 |
| Epoch 5 | 0.15 | 1.12 | 0.23 | 4.11 | −0.13 | 1.10 |

**Note.** Cases (a) and (c) are fine, but (b) is dangerous.

### 3.2. Optimal Cadence

We ran our experiment for different observational cadences (30, 90, 180, and 270 days) to see which retrieves the most accurate orbits in the least number of epochs. Here, the host stars are all at 10 pc and planets have beta-distributed eccentricities. Our optimal cadence pertains to planets with periods in or near the G2V HZ.

The results are shown in Figures 6 and 7—by the third epoch, $a$ can be constrained to within 10% of its true value for 68% of the samples only for cadences of 90 days or greater. In half a period, the planet will move to the furthest possible position from its original location. Revisiting the planet at its antipodal point means that we are less likely to miss the planet inside the IWA (e.g., compared to revisiting a quarter-period later), although this is tied to us ensuring the first epoch is a detection.

Cadences of 180 and 270 days result in the best accuracy on the semimajor axis at epoch 3, but 180-day cadences for near-360-day orbits may not break formal degeneracies between $\omega_p$ and $\Omega$. Hence the optimal cadence is not necessarily the same for $a$ as for other orbital parameters. While the 270-day cadence also performs well, in practice this long wait could risk losing track of the planet, as discussed below.

### 3.3. Effects of Distance and Astrometric Precision

We have adopted a distance of 10 pc and a measurement error of $\sigma_\theta = 3.5$ mas for our numerical experiments thus far. However, the number of epochs needed to constrain an orbit is distance-dependent. More-distant targets impede orbit retrievals in two ways: the corresponding measurement error on the planet position increases (by a factor of four from 5 to 20 pc); and the larger projected separation of the IWA is more likely to obscure targets (nearly 0% non-detections at 5 pc to ∼70% at 20 pc).

To quantify the effect of distance, we repeated the retrieval experiments for the same planets at 5 and 20 pc, assuming 90-day cadences. At our assumed IWA and range of true semimajor axes, HZ planets will be obscured 100% of the time by 55 pc. This experiment is equivalent to increasing the IWA and astrometric uncertainty. Particular sources of astrometric error and their impact on orbit retrievals are discussed in Pueyo et al. (2015).

Figure 8 illustrates the change in the semimajor axis retrieval accuracy with distance. As expected, the width of the $1\sigma$ confidence interval increases for more-distant targets, for all epochs.

The encouraging performance of the fit at 5 pc is attributed to the small astrometric error. For these nearest targets we would be capable of measuring $a$ to within <2% by the fourth epoch.

We are only showing this distance dependence of accuracy for one cadence. The optimal cadence is distance-dependent; with increasing distance, obscuration dominates the measurement error in worsening the accuracy, preferring 180-day (antipodal) cadences.

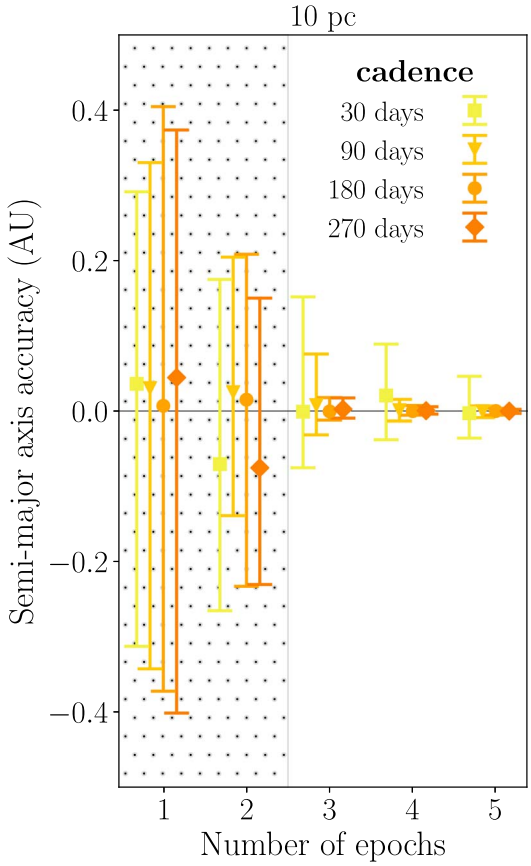


**Figure 6.** $1\sigma$ constraints on MCMC retrieval accuracies of the semimajor axis, as a function of the number of epochs, compared across increasing observational cadences. Accuracy is given as $a_{\text{fitted}} - a_{\text{true}}$, where $a_{\text{fitted}}$ is the median of the posterior semimajor axis distribution from the MCMC. The error bars show the 16th and 84th percentiles of the sample accuracy (100 in each scenario), and symbols mark the medians, with colors representing different cadences.

## 4. Discussion

The minimum number of epochs we need depends on the science question we are asking.

### 4.1. Placing the Planet in the HZ

We first consider hypothetical HZs defined only by how narrow they are. Compare the widths of these hypothetical HZs with the semimajor axis retrievals of planets on 1 au orbits (Figure 9, left panel): at $1\sigma$ confidence, for an optimistic, 25%-wide HZ centered on 1 au, a single epoch of direct imaging

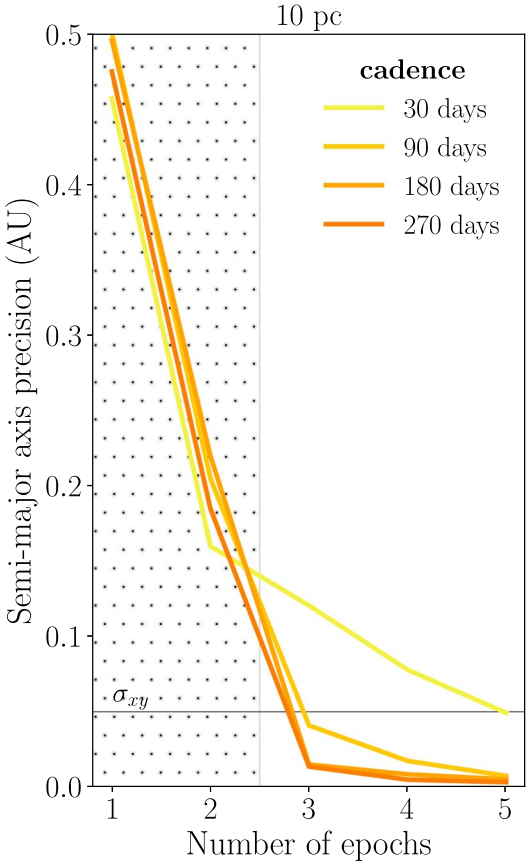


**Figure 7.** Sample mean MCMC retrieval precisions of the semimajor axis, as a function of the number of epochs, compared across observational cadences. Precision is given as half the distance between the 16th and 84th percentiles of the retrieved posteriors. The colored lines show the median sample precision (100 in each scenario) for different cadences. The $\sigma_{xy}$ line marks the measurement precision for 3.5 mas astrometry at 10 pc.

generally suffices to nail the semimajor axis accurately. For a pessimistic 1%-wide HZ, four or five epochs are needed.

To compare these results with theoretical HZs from the literature, the right panel in Figure 9 shows the same semimajor axis constraints to scale with both a LUVOIR-esque IWA and with the theoretical HZ inner limits for solar twins from Kopparapu et al. (2013). By the fourth epoch we are 68% sure a planet is beyond the moist greenhouse inner limit. To constrain an orbit beyond the more optimistic recent Venus inner limit at 0.75 au at the same confidence, about one epoch is required.

We are implicitly assuming that the HZ is a function of the semimajor axis only. While high eccentricities may negatively affect the stability of surface water (Bolmont et al. 2016), the high thermal inertia of extant oceans can buffer against transitions to snowball states at apoastron (Dressing et al. 2010). Either way, the long-term stability of a planetary climate depends, to the first order, on the average incident flux over the entire orbit (Williams & Pollard 2002), which is proportional to $(1-e^2)^{-1/2}$.

### 4.2. Predicting the Future Location of a Planet

In some cases, we would need to constrain all six orbital elements. For example, knowing the entire orbit would help us predict the exoplanetary ephemerides. The best theoretical precision on extrapolating the position of a planet is set by the

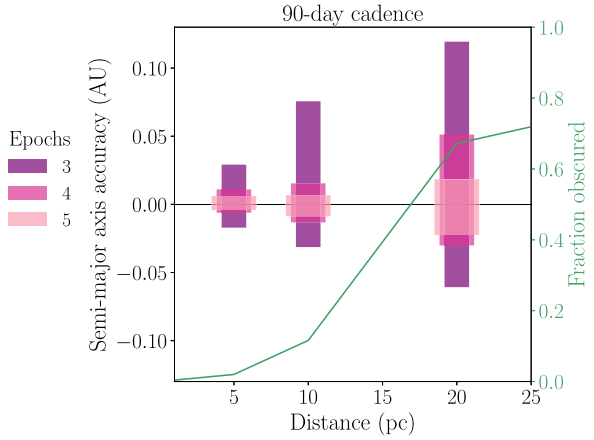


**Figure 8.** Targets at distance. The height of each rectangle represents the $1\sigma$ confidence interval of the semimajor axis retrieval accuracy for 100 planets, with shading according to epoch—darker colors are less visits. The green line indicates the actual sample fraction of observations, where $\theta_{\rm proj} < \theta_{\rm IWA}$. We would expect accuracy to worsen linearly with distance if only astrometric error mattered (a geometric effect); it worsens faster than linearly because non-detections start to dominate after around 17 pc.

detector pixel scale, although in practice, the precision we achieve depends on the post-processing. Using three epochs, we have predicted the position of a planet at the fourth epoch to within a given number of pixels on average. For 180-day cadences, we find the position to be constrained to within one pixel; for 90-day cadences, within two pixels, with five epochs necessary to get down to a one-pixel accuracy. The trade-off is that waiting longer between observations provides better accuracies for the Keplerian parameters, but any errors are amplified by the greater distance the planet will have traveled.

On a related note, we have only considered single-planet systems in our study. However, we expect a substantial fraction of planets to occur in multiplanet systems (Zhu et al. 2018). This presents an interesting tension with respect to the optimal cadence: for multiplanet systems, we may prefer shorter observational cadences because the planet will move less, so there is less risk of confusing it with other planets. Nevertheless, planet disambiguation in multiplanet systems will be made a bit easier by the fact that we can expect these targets to have near-circular orbits (Van Eylen et al. 2019).

### 4.3. Extension to FGK Stars

We have only considered solar-twin stars in this work. However, LUVOIR and HabEx will target a larger range of host stars, whose HZs will scale with stellar mass according to $r/1\ \mathrm{au} \propto (M_*/M_\odot)^2$. This matters because tighter HZs will be obscured more often inside the IWA. In terms of angular separation, then, this scaling is equivalent to varying the distance to the host star. For example, a planet receiving the same stellar flux as Earth and orbiting a K2V star ($0.70\,M_\odot$) will have a separation of 0.49 au, which is essentially isomorphic to doubling the distance to an equivalent G2V system, as in Figure 8 (barring the astrometric precision change). The optimal cadence for such a system would also scale with the orbital period as $P/365\ \mathrm{days} \propto (M_*/M_\odot)^3$, holding to the logic that 90 days $\approx 0.25P$ for solar twins.

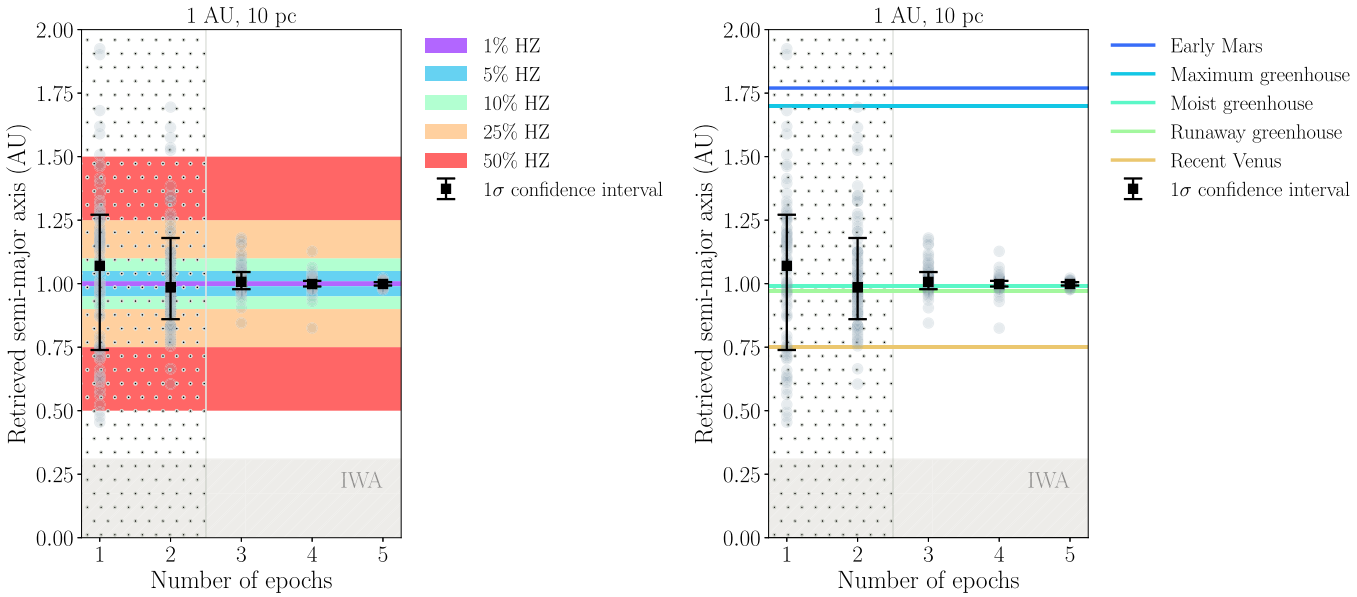


**Figure 9.** Constraints on the semimajor axis compared to HZs. Error bars show the $1\sigma$ sample error of the retrieved semimajor axes for a sample of 100 planets, where all planets are assigned a true semimajor axis of 1 au. The gray hatched area shows the region obscured by the coronagraph inner working angle (31 mas) for a target at 10 pc. Different conceptual HZs are superimposed. Left: colored swaths represent arbitrary HZs relative to the true semimajor axis. For a hypothetical HZ with 5% width, three epochs can constrain an orbit to lie within, at $1\sigma$. Right: colored horizontal lines indicate HZ limits from Kopparapu et al. (2013). At $1\sigma$, constraining a planet to lie within the wider recent Venus inner limit requires about one epoch.

### 4.4. Uncertainties on Stellar Distance

This model has not considered imperfect knowledge of the distance to the host star. The uncertainty on this distance would propagate linearly to an uncertainty on the planet ($x$, $y$) position, which compounds the astrometric measurement error, $\sigma_\theta$. LUVOIR and HabEx target stars would be chosen from the *Hipparcos* and *Gaia* catalogs (HabEx Team 2018; LUVOIR Team 2018), with known trigonometric parallaxes. For a G2V star within 50 pc (the maximum LUVOIR survey distance), the end-of-mission *Gaia* parallax uncertainty will be on the order of 0.1 mas (de Bruijne et al. 2014); i.e., at most $\pm 1$ pc in distance, and an order of magnitude lower than our adopted value of $\sigma_\theta$.

### 4.5. Limitations of Using Astrometry Alone

This work demonstrates a first step in the orbit determination problem, considering only the planet position data. Yet direct imaging would also measure the brightness of the planet relative to its star. Including this photometric information could both help and hurt the conclusions of this paper.

On the one hand, the brightness at a given ($x$, $y$) position represents a constraint on the orbital phase of the planet. Although the phase functions of our target planets are unknown a priori, Lambertian reflection is a fair assumption at phase angles smaller than the crescent phase (Robinson et al. 2010).

On the other hand, we have so far ignored that a planet may go undetected at a given epoch not because it is inside the IWA, but because it is imaged at an orbital phase not bright enough for the contrast floor. This would limit our ability to use non-detections as a constraint on the planet position for nearby targets. For more-distant targets, the IWA will tend to consume crescent-phase planets either way—we will have less leverage on orbital phase constraints from brightness variations, and missing planets because they are too faint will stop mattering.

Further, photometry brings with it other sources of noise beyond astrometric precision, namely, distinguishing real planets from speckles. This work implicitly assumes that a given contrast floor (e.g., $10^{-10}$; LUVOIR Team 2018) already accounts for the removal of speckles in post-processing. Including photometry in our model may increase the number of visits to achieve the same precision, although it would not increase the required number of detections.

The question this work has asked is: how well can one locate a planet in the HZ? Under the assumption of one planet per star, we find that cadences of 180 and 270 days have the best precision and accuracy. In a realistic mission, however, these long cadences would likely be poor at resolving confusion between multiple planets in the same system. A 90-day cadence is a good compromise.

Even if the third image placed a planet in the HZ at 95% confidence, we would still need more epochs to establish its orbit before pursuing expensive spectroscopy. Yet the point is that we could begin prioritizing targets after only one or two epochs. This quantifies a key parameter in design reference missions for future direct imaging concept missions HabEx and LUVOIR. In the LUVOIR study, constraining orbits within the HZ is the third step in identifying a habitable planet after (1) establishing the target star list and (2) performing multicolor point source photometry to rule out background objects (LUVOIR Team 2018). Our work finds that the orbit-constraining step can be done more efficiently than before, reducing the minimum number of observations from six to three. This extra time could be spent characterizing more planets.

This work is supported by the McGill Space Institute, the TEPS training program, and an NSERC Discovery Grant. The authors wish to thank the HabEx AEIWG for their useful conversations, especially Leslie Rogers, Eric Nielsen, and Scott Gaudi. Comments from an anonymous reviewer also greatly

improved the quality of this manuscript. All research was conducted on the territory of the Kanien'kehá:ka, the keepers of the Eastern Door of the Haudenosaunee Confederacy.

## Appendix
## Forward Model

We assume that all planets are on bound Keplerian orbits.

If we hover with the planet's orbital plane below us and our right ears toward the $+x$ reference direction, we will see an ellipse. Its shape can be described parametrically as the first two-dimensions of a three-dimensional matrix,

$$\begin{bmatrix} x_k \\ y_k \\ z_k \end{bmatrix} = \begin{bmatrix} a\cos(M_k) - c \\ b\sin(M_k) \\ 0 \end{bmatrix}, \tag{4}$$

where $a$ is the semimajor axis, $b = a\sqrt{1-e^2}$ is the semiminor axis, $c = \sqrt{a^2 - b^2}$ is an $x$-intercept, $M_k$ is the mean anomaly at the $k$th epoch, and the orbital plane is defined by $\boldsymbol{x}$–$\boldsymbol{y}$.

The ellipse we see in the detector plane, $\boldsymbol{x}'$–$\boldsymbol{y}'$, has been rotated through three angles in the order $\Omega$, $i$, $\omega_p$. There is a Euclidean rotation matrix corresponding to this series of rotations,

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} \cos\gamma & -\sin\gamma & 0 \\ \sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\beta & -\sin\beta \\ 0 & \sin\beta & \cos\beta \end{bmatrix} \cdot \begin{bmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix}, \tag{5}$$

where $\alpha = \Omega$, $\beta = i$, and $\gamma = \omega_p$.

The projected separation in the image plane is the Euclidean distance, $\sqrt{(x')^2 + (y')^2}$.

### *A.1. Planets Move*

Both the cadence, $\delta t$, and the walker's current guesses of $a$ and $M_0$ control the planet's position at the $k$th epoch (i.e., set by the instantaneous mean anomaly $M_k$):

$$M_k = M_0 + \sqrt{\frac{\mu}{a^3}}\, k\delta t\, \frac{86400\ \mathrm{s}}{\mathrm{day}}, \tag{6}$$

where $\sqrt{\mu/a^3}$ is the mean motion of the planet (in radians per second) and $\mu$ is the standard gravitational parameter.

## ORCID iDs

Claire Marie Guimond https://orcid.org/0000-0003-1521-5461
Nicolas B. Cowan https://orcid.org/0000-0001-6129-5699

## References

Abbot, D. S., Cowan, N. B., & Ciesla, F. J. 2012, ApJ, 756, 178
Bean, J. L., Abbot, D. S., & Kempton, E. M.-R. 2017, ApJL, 841, L24
Blunt, S., Nielsen, E. L., De Rosa, R. J., et al. 2017, AJ, 153, 229
Bolmont, E., Libert, A.-S., Leconte, J., & Selsis, F. 2016, A&A, 591, A106
Burke, C. J., Christiansen, J. L., Mullally, F., et al. 2015, ApJ, 809, 8
Christiansen, J. L., Clarke, B. D., Burke, C. J., et al. 2015, ApJ, 810, 95
de Bruijne, J. H. J., Rygl, K. L. J., & Antoja, T. 2014, EAS, 67, 23
De Rosa, R. J., Nielsen, E. L., Blunt, S. C., et al. 2015, ApJL, 814, L3
Dressing, C. D., Spiegel, D. S., Scharf, C. A., Menou, K., & Raymond, S. N. 2010, ApJ, 721, 1295
Fischer, D. A., Anglada-Escude, G., Arriagada, P., et al. 2016, PASP, 128, 066001
Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306
Foreman-Mackey, D., Hogg, D. W., & Morton, T. D. 2014, ApJ, 795, 64
García Muñoz, A., Lavvas, P., & West, R. A. 2017, NatAs, 1, 0114
Guimond, C. M., & Cowan, N. B. 2018, AJ, 155, 230
HabEx Team 2018, arXiv:1809.09674
Hart, M. H. 1979, Icar, 37, 351
Kasting, J. F. 1988, Icar, 74, 472
Kasting, J. F., & Toon, O. B. 1989, in Origin and Evolution of Planetary and Satellite Atmospheres, ed. S. K. Atreya, J. B. Pollack, & M. S. Matthews (Tucson, AZ: Univ. Arizona Press), 423
Kasting, J. F., Whitmire, D. P., & Reynolds, R. T. 1993, Icar, 101, 108
Kipping, D. M. 2013, MNRAS, 434, L51
Kopparapu, R. K., Hébrard, E., Belikov, R., et al. 2018, ApJ, 856, 122
Kopparapu, R. K., Ramirez, R., Kasting, J. F., et al. 2013, ApJ, 765, 131
Kosmo O'Neil, K., Martinez, G. D., Hees, A., et al. 2018, arXiv:1809.05490
LUVOIR Team 2018, arXiv:1809.09668
Mede, K., & Brandt, T. D. 2017, AJ, 153, 135
Nielsen, E. L., Close, L. M., Biller, B. A., Masciadri, E., & Lenzen, R. 2008, ApJ, 674, 466
Petigura, E. A., Howard, A. W., & Marcy, G. W. 2013, PNAS, 110, 19273
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
Robinson, T. D., Meadows, V. S., & Crisp, D. 2010, ApJL, 721, L67
Ruffio, J.-B., Mawet, D., Czekala, I., et al. 2018, AJ, 156, 196
Silburt, A., Gaidos, E., & Wu, Y. 2015, ApJ, 799, 180
Stark, C. C., Bolcar, M., Fogarty, K., et al. 2018, LPICo, 2065, 2061
Stark, C. C., Roberge, A., Mandell, A., et al. 2015, ApJ, 808, 149
Stark, C. C., Roberge, A., Mandell, A., & Robinson, T. D. 2014, ApJ, 795, 122
Stark, C. C., Shaklan, S., Lisman, D., et al. 2016, JATIS, 2, 041204
Van Eylen, V., Albrecht, S., Huang, X., et al. 2019, AJ, 157, 61
Walker, J. C. G., Hays, P. B., & Kasting, J. F. 1981, JGR, 86, 9776
Wang, J. J., Graham, J. R., Dawson, R., et al. 2018, AJ, 156, 192
Wang, J. J., Graham, J. R., Pueyo, L., et al. 2016, AJ, 152, 97
Williams, D. M., & Pollard, D. 2002, IJAsB, 1, 61
Zhu, W., Petrovich, C., Wu, Y., Dong, S., & Xie, J. 2018, ApJ, 860, 101