



Identifying Exo-Earth Candidates in Direct Imaging Data through Bayesian Classification

Alex Bixel^{1,2} and Dániel Apai^{1,2,3}

¹ Department of Astronomy/Steward Observatory, The University of Arizona, 933 N. Cherry Avenue, Tucson, AZ 85721, USA

² Earths in Other Solar Systems Team, NASA Nexus for Exoplanet System Science, USA

³ Lunar and Planetary Laboratory, The University of Arizona, 1640 E. University Boulevard, AZ 85721, USA

Received 2019 June 21; revised 2019 October 4; accepted 2019 October 26; published 2019 December 10

Abstract

Future space telescopes may be able to directly image between approximately 10 and 100 planets with sizes and orbits consistent with habitable surface conditions (“exo-Earth candidates” or EECs), but observers will face difficulty in distinguishing these from the potentially hundreds of nonhabitable “false positives” which will also be detected. To maximize the efficiency of follow-up observations, a prioritization scheme must be developed to determine which planets are most likely to be EECs. In this paper, we present a Bayesian method for estimating the likelihood that any directly imaged extrasolar planet is a true EEC by interpreting the planet’s apparent magnitude and separation in the context of existing exoplanet statistics. As a specific application of this general framework, we use published estimates of the discovery yield of future space-based direct-imaging mission concepts to conduct “mock surveys” in which we compute the likelihood that each detected planet is an EEC. We find that it will be difficult to determine which planets are EECs with >50% confidence using single-band photometry immediately upon their detection. The best way to reduce this ambiguity would be to constrain the orbit of a given planet by revisiting the system multiple times or through a radial velocity precursor survey. Astrometric or radial velocity constraints on the mass of the planet would offer a lesser benefit. Finally, we show that a Bayesian approach to prioritizing targets would improve the follow-up efficiency of a direct imaging survey versus a blind approach using the same data. For example, the prioritized approach could reduce the amount of integration time required for the spectral detection (or rejection) of water absorption in most EECs by a factor of two.

Unified Astronomy Thesaurus concepts: Direct imaging (387); Habitable planets (695); Astronomical techniques (1684); Bayesian statistics (1900); Space telescopes (1547); Astrobiology (74)

1. Introduction

One of the primary science goals driving the development of new space telescopes is the detection and characterization of Earth-like planets around nearby stars. Spectroscopy of the planets’ reflected light spectra would reveal whether they are potentially habitable—i.e., liquid water could exist on their surfaces. The presence of biosignature gases such as O₂, O₃, and CH₄ could be interpreted as evidence for life beyond Earth (e.g., Seager et al. 2016; Fujii et al. 2018), although this interpretation would not be straightforward as many of these gases can be produced abiotically (e.g., Catling et al. 2018; Meadows et al. 2018).

Recently, final study reports have been published for two space telescope design concepts with a primary science goal of directly imaging and spectroscopically characterizing Earth-like planets around nearby stars. These are the Habitable Exoplanet Observatory (HabEx, The HabEx Study Team 2019) and the Large UV/Optical/Infrared Surveyor (LUVOIR, The LUVOIR Study Team 2019, hereafter L19). Following Stark et al. (2014), both reports provide estimates for the expected yield of planets across a range of sizes and insulations. The HabEx report predicts the detection and characterization of 8⁺⁹₋₅ approximately Earth-sized planets in the habitable zone with a 4 m aperture, while the LUVOIR report predicts 54⁺⁶¹₋₃₄ with the 15 m aperture “LUVOIR-A” design. These design concepts have been thoroughly investigated, and it is likely that the design of any future direct imaging space telescope would enable it to detect comparable numbers of potentially habitable planets with moderate signal-to-noise ratio (S/N) photometry.

However, these dozens of “exo-Earth candidates” (EECs) would be detected amidst hundreds of planets with atmospheric compositions or equilibrium temperatures not conducive to Earth-like life, including planets outside the habitable zone, large mini-Neptunes with thick H/He envelopes, or Mars-sized planets which have been stripped of their atmospheres. These nonhabitable planets often demonstrate the same *observable* parameters (e.g., apparent magnitude and separation) as the EECs, but are far more common and may therefore cause a significant number of false-positive detections. We demonstrate this problem in Figure 1. In fact, Guimond & Cowan (2018) show that separation- and contrast-based selection criteria could suffer from a false discovery rate as high as 77% given only the detection data in a single band, or 47% if prior constraints on the orbit are available.

While the “false positive” planets would be interesting to characterize in their own right, spectroscopy across the full wavelength range could take weeks for the faintest targets, diverting time and resources from higher-priority targets. To separate the potentially habitable and nonhabitable planets, a survey could use multi-epoch broadband photometry to characterize their orbits and spectroscopic observations of H₂O absorption features to confirm the presence of water vapor in habitable zone targets. However, even within the habitable zone, there exist many planets that are too small or too large to be habitable, and the identification of water absorption features will require a significant investment of time (Kawashima & Rugheimer 2019). Additionally, warmer planets within the runaway greenhouse limit could exhibit water absorption features. In order to maximize the efficiency of an exo-Earth

False positives for exo-Earth candidate detections

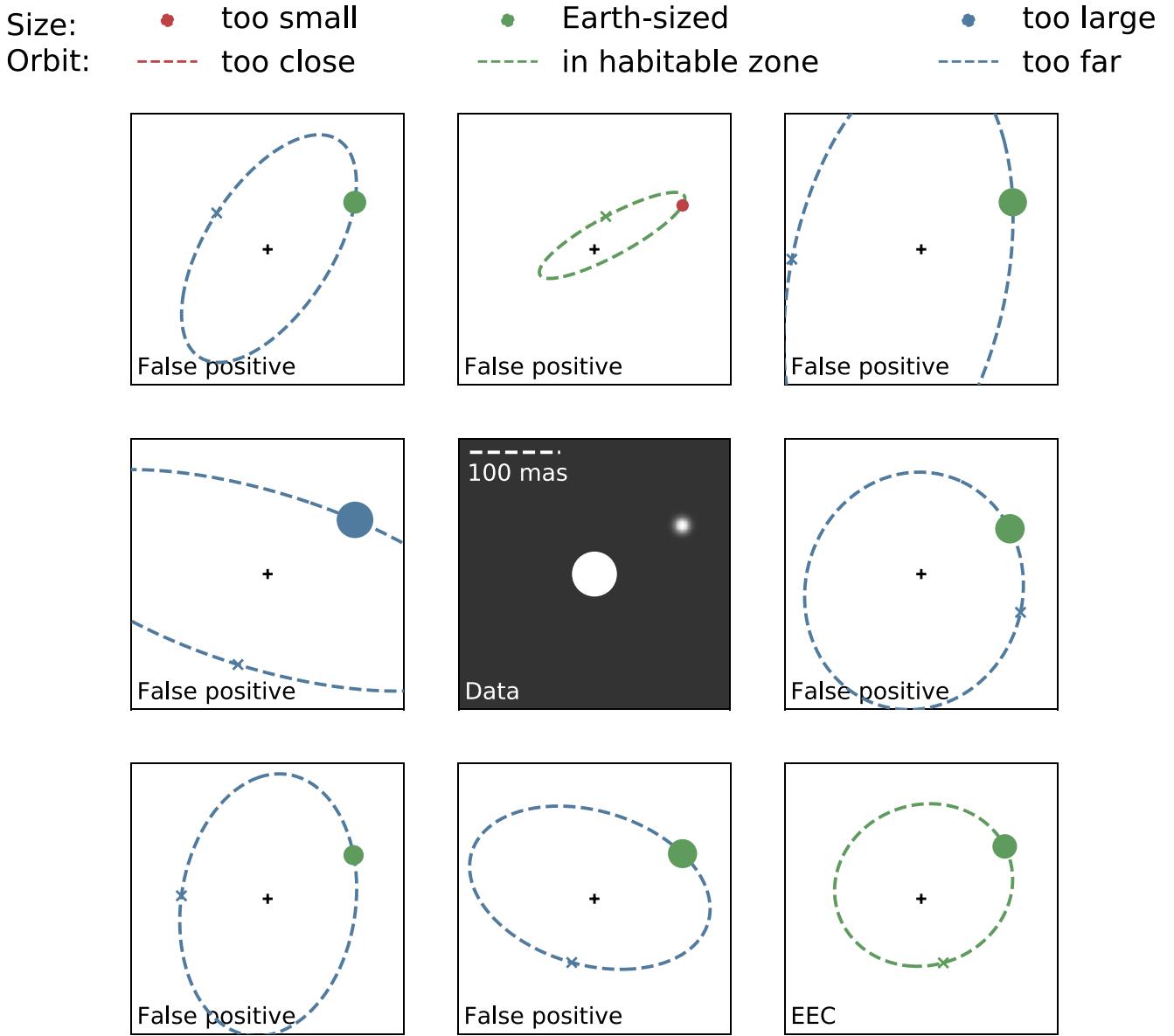


Figure 1. To illustrate the degeneracies which affect the interpretation of direct imaging data, we simulate the detection of a planet orbiting a Solar-type star at 15 pc (center panel), as well as several planets of varying sizes, orbits, and albedos which have a similar projected separation and magnitude (surrounding panels). It is not clear whether this data point represents a true EEC, or one of many potential false positives. The color and size of each circle represents the potential radius of the planet; only green points are approximately Earth-sized ($\sim 0.8\text{--}1.6 R_{\oplus}$). The color of the potential orbit represents its insolation; only green orbits are in the habitable zone. An “x” marks the planet’s closest approach to the observer.

imaging mission, it is necessary to develop a method for identifying those planets that are most likely to be Earth analogs using the limited data which will be available upon detection.

We previously advocated a Bayesian approach to assessing the potential habitability of newly detected exoplanets (Apai et al. 2018). The Bayesian approach allows one to probabilistically constrain the properties of the planet by leveraging knowledge from exoplanet statistics on planet radii, masses, and orbital properties. It also allows one to fold in predictions from theoretical models of planet formation and evolution as prior knowledge. As an example, in Bixel & Apai (2017) we used the Monte Carlo method to infer the likely composition of

Proxima Cen b (Anglada-Escudé et al. 2016) in light of well-established statistical priors and the limited data available about the system. We found that it is $\sim 90\%$ likely that the planet is small and rocky as opposed to a “mini-Neptune.” In this paper, we extend this approach to assess the likelihood that a directly imaged planet has an appropriate composition and orbit to be potentially habitable.

In Section 2 we review our Bayesian framework and give an example of how it can be applied to characterize directly imaged planets. In Section 3 we discuss the prior assumptions upon which our framework is based, and how they might be improved in the coming decade. Using the planet yield estimates provided in LUVOIR interim report (The LUVOIR Team 2018,

Table 1
A List of the Prior Assumptions which we Use to Build our Sample in Case 1, Along with Relevant Literature References

Parameter	Description of Prior	References
Radius and period	<i>Kepler</i> occurrence rates	Mulders et al. (2018)
Class	“sub-terrestrial,” “terrestrial,” or “ice giant” based on radius, following Figure 4	Fulton et al. (2017), Zahnle & Catling (2017)
Mass	Empirical mass–radius relationship with intrinsic spread	Wolfgang et al. (2016)
Habitable zone boundaries	Planet mass-dependent LWHZ models (runaway and maximum greenhouse limits)	Kopparapu et al. (2014)
Albedo	Uniform from 0.2 to 0.7	
Eccentricity	Beta distribution ($\alpha = 0.867$, $\beta = 3.03$)	Kipping (2013)
ω , Ω , M^a	Uniform from 0 to 2π	
$\cos(i)$	Uniform from -1 to 1	
Exo-Earth Candidates (EECs)	“terrestrial” class planets in the LWHZ	

Note. The additional cases in Table 2 may modify these assumptions to reflect new prior knowledge or data.

^a Argument of pericenter; longitude of the ascending node; mean anomaly.

hereafter L18) as a baseline estimate for the yield of a hypothetical direct-imaging mission, in Section 4 we conduct mock surveys where we detect each target and estimate the apparent likelihood that it has a potentially habitable composition and orbit. In Section 5 we discuss the results of these surveys, including the types of nonhabitable planets that would be mistaken for potentially habitable planets, and which additional data could help to resolve this ambiguity. Finally, we show that our approach to target prioritization could greatly enhance the efficiency of follow-up observations after all of the planets have been detected.

2. A Bayesian Framework for Classifying Directly Imaged Planets

2.1. Monte Carlo Method

Here we review the Monte Carlo method for Bayesian inference. This method allows an observer to constrain the *unobservable* properties of a planet based on limited-precision measurements of its observable data values by assuming some understanding of the prior distribution of intrinsic properties and their relationship to the observable data values.

Suppose planets can be described by some set of intrinsic properties θ and some resulting set of observable data values x which can be calculated from θ . Given a prior probability distribution for the values of θ , $P(\theta)$, then the probability distribution for θ given x can be calculated using Bayes’ equation:

$$P(\theta|x) = \frac{P(x|\theta)P(\theta)}{\int_{\theta} P(x|\theta)P(\theta)}.$$

The left term is commonly referred to as the posterior distribution of θ , and $P(x|\theta)$ as the likelihood function.

For most astrophysical applications there is no analytical solution to this equation, so it must be solved numerically. Under the Monte Carlo method, we use the prior probability distribution $P(\theta)$ to simulate a set of properties θ_{sim} , then calculate the simulated data values x_{sim} directly. Next, we accept or reject this simulated planet based on the value of its likelihood function for some observed set of data values x_{obs} . Assuming the data values are drawn from independent, normal probability distributions with standard deviations (measurement uncertainties) σ_{obs} , the likelihood function is that of a

multivariate Gaussian:

$$P(x|\theta) = \prod_{i=1}^m \exp((x_{\text{obs},i} - x_{\text{sim},i})^2 / 2\sigma_{\text{obs},i}^2).$$

If only an upper limit is available for a component of x_{obs} (e.g., magnitude), then the prior sample is first cleaned of simulated members exceeding that limit.

This procedure is repeated in parallel for a large number of planets until a statistically sufficient number are accepted. The result of the likelihood-based selection is a sample of planets whose properties θ_{sim} are distributed according to the posterior distribution. In other words, a histogram of the accepted values of θ_{sim} represents the probability distribution for θ_{obs} , the properties of the observed planet.

2.2. Constructing the Prior Sample

We use the priors in Table 1 to construct the prior sample according to the algorithm visualized in Figure 2, and we discuss the prior assumptions in detail in Section 3. We consider several cases governing the amount of data available to the observer—these are listed in Table 2.

Before we simulate the properties of the directly imaged planet, we first consider the properties of its host star, which the observer will only know with finite precision. We represent this measurement uncertainty by drawing a unique stellar radius and mass for each system from normal distributions with $\sigma = 3\%$ and 7% . We assume nearly exact measurements of the distance and luminosity of the host star, and we discuss these assumptions in Section 3.1.

We generate the radius and period of the planet using *Kepler* occurrence rates as a prior probability distribution, then use the radius to classify the planet as too small (“sub-terrestrial”), too large (“ice giant”), or of the proper size to maintain a habitable atmosphere against stellar irradiation (“terrestrial”). This requires us to extrapolate to planets smaller ($R \lesssim 0.5 R_{\oplus}$) or with longer periods ($P \gtrsim 100$ days) than those readily available in the *Kepler* sample, as we discuss in Section 3.2. We calculate the mass from an empirical mass–radius relationship, where we assume some intrinsic variance due to stochastic planet formation histories and differences between the host stars of the planets on which these relationships are based. We draw eccentricities from a beta distribution and the remaining orbital elements are assumed to be isotropically distributed.

The liquid water habitable zone (hereafter LWHZ) is the range of orbital separations over which a broadly Earth-like

Simulating the prior sample of planets

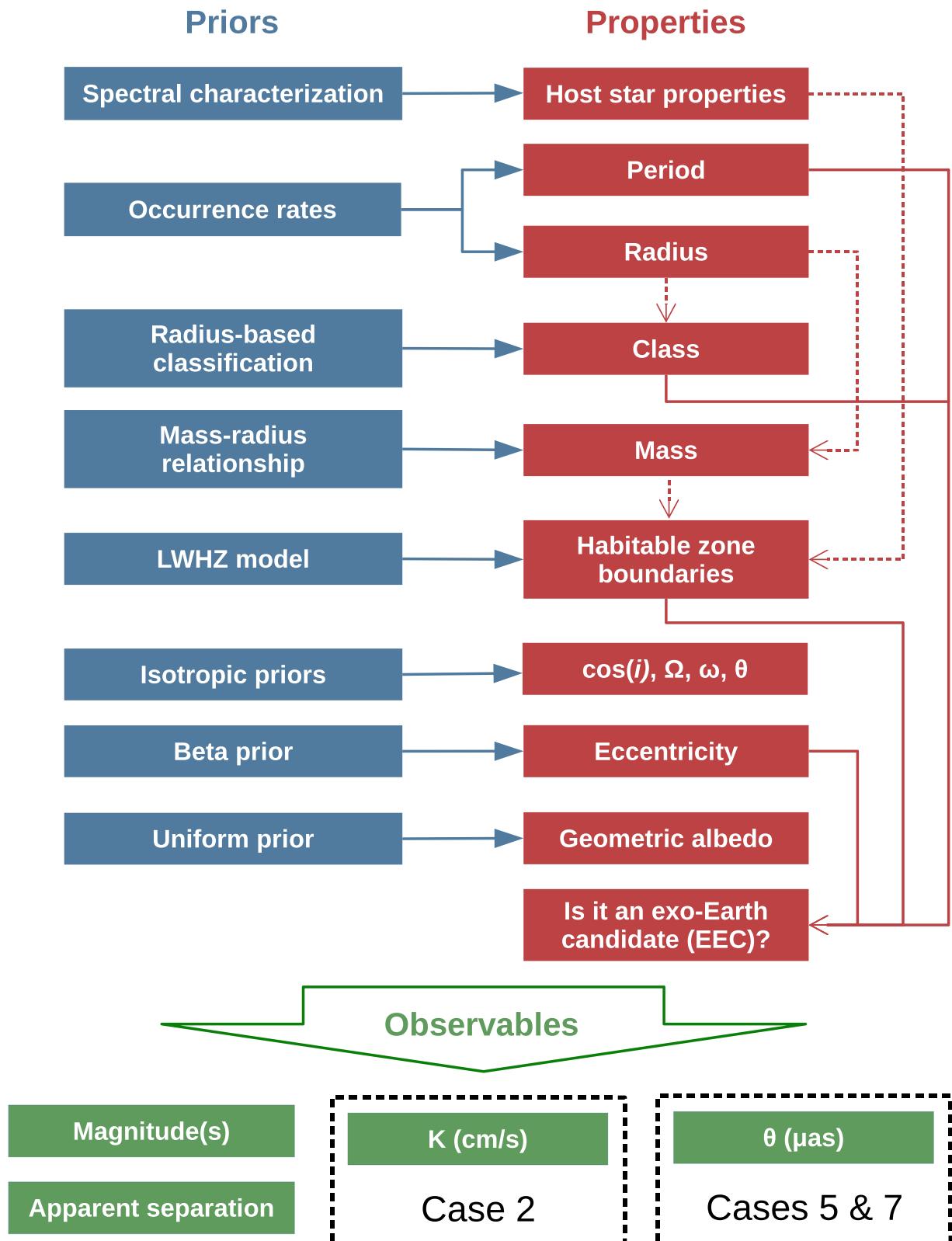


Figure 2. Flowchart illustrating our algorithm for simulating the prior sample of planets under Case 1. The red boxes represent the intrinsic planet properties which we simulate, and the red arrows indicate how they are used in the calculation of other properties. The blue boxes represent the priors which we use to simulate the properties, and the green boxes are the observable values which can be compared to the data.

Table 2
The Different Cases Under Which we Conduct our Mock Surveys, Representing the Different Data which the Observer Might have

Case	Description
1	Each planet has a monochromatic geometric albedo drawn uniformly from 0.2 to 0.7. The planet's monochromatic magnitude and separation angle are observed in a single epoch with a signal-to-noise ratio of 7 and centroid precision $\sigma_c = 3.8$ mas.
2	Prior radial velocity (RV) observations provide constraints of $\pm 10\%$ on the period, $\pm 5 \text{ cm s}^{-1}$ on the RV semi-amplitude, and $\pm 30^\circ$ on the mean anomaly, in addition to the data from Case 1.
3	Simultaneous observations of a debris disk provide a $\pm 1^\circ$ constraint on the inclination and longitude of the ascending node of the orbital plane. The planet's orbital elements may be further offset from these by ± 0.2 in Ω and $\cos(i)$.
4	Multiple epochs of direct imaging data permit constraints of $\pm 15^\circ$ on the phase angle, $\pm 10\%$ on the semimajor axis, and ± 0.05 on the eccentricity.
5	Case 4 with an additional $\pm 0.1 \mu\text{as}$ constraint on the semi-amplitude of the motion of the star due to the planet.
6	We assume prior knowledge about the albedo distribution of small planets. Each planet's spectral albedo is determined by its class and location with respect to the habitable zone. Random monochromatic and polychromatic offsets are also introduced. The planet's magnitude in three bands and separation angle are observed in a single epoch with a wavelength-integrated signal-to-noise ratio of 7.
7	Combination of Cases 5 and 6: three-band photometry is available along with constraints on the orbit of the planet and the astrometric motion of the star from multi-epoch observations.

planet could feasibly host liquid water on its surface. Kopparapu et al. (2014) find the boundaries of this zone to be a function of the mass of the planet and the effective temperature of the star, hence we cannot infer a planet's membership to the LWHZ based solely on its insolation. We calculate the LWHZ boundaries and determine whether the planet lies in the runaway greenhouse, temperate, or maximum greenhouse regimes, interpolating between the discrete planet masses modeled by Kopparapu et al. (2014) (0.1, 1.0, and $5.0 M_\oplus$), and taking the minimum or maximum mass values for planets outside of this range. For planets with noncircular orbits we use the mean flux approximation to determine whether they are in the habitable zone, which Bolmont et al. (2016) find to be valid for planets with low or moderate eccentricities receiving Earth-like insolations.

Finally, we assign a geometric albedo to each planet using one of two methods for each of the cases in Table 2. In Cases 1–5, we assign a monochromatic albedo from a broad uniform prior with no dependence on the class or orbital parameters of the planet. In Cases 6 and 7, we generate a spectrum for each planet by mapping it to a solar system analog with a comparable size and orbit. These two cases allow us to test the usefulness of color information for identifying potential exo-Earths. We dedicate more considerable discussion to the geometric albedos in Section 3.6.

2.3. Calculating the Observable Data Values

Once the full assortment of planet properties has been simulated, we can proceed to calculating the observable data values in order to compare them with those of a newly detected planet.

2.3.1. Apparent Separation

The angular separation vector has two components and can be calculated from the orbital elements and the distance to the system d . We adopt the same reference frame and notation as Murray & Correia (2010),⁴ where $i = 90^\circ$ is an “edge-on” inclination and the observer is at $z = \infty$, so the angular

separation components orthogonal to the line of sight are:

$$s_x = (r/d)[\cos(\Omega)\cos(\omega + f) - \sin(\Omega)\sin(\omega + f)\cos(i)]$$

$$s_y = (r/d)[\sin(\Omega)\cos(\omega + f) + \cos(\Omega)\sin(\omega + f)\cos(i)].$$

In most cases, the position angle has no meaningfully defined zero-point, so for simplicity we calculate only the net separation $|s| = \sqrt{s_x^2 + s_y^2}$. The exception is Case 4, where the coeval detection of a debris disk is used to constrain the orbital plane of the planet; therefore the position angle is meaningfully defined, and we compute s_x and s_y separately.

2.3.2. Apparent Magnitude(s)

Following Madhusudhan & Burrows (2012),⁵ we model the planet as a Lambertian sphere, in which case the planet-to-star contrast ratio when observed at orbital phase α is:

$$\frac{L(\lambda)}{L_*(\lambda)} = A_g \left(\frac{R_p}{a} \right)^2 \left[\frac{\sin(\alpha) + (\pi - \alpha)\cos(\alpha)}{\pi} \right].$$

The phase angle is $\alpha = \text{Cos}^{-1}[\sin(\omega + f)\sin(i)]$. We draw the geometric albedo from a prior distribution, and allow slightly super-Lambertian values ($A_g > 2/3$) as these are observed in some wavelength ranges in the solar system.

2.3.3. Radial Velocity and Astrometric Semi-amplitudes and Periods

The semi-amplitude of the star's periodic RV variation, assuming no other perturbers and $M_P \ll M_*$, is:

$$K = (8.95 \text{ cm s}^{-1}) \frac{(M_p/M_\oplus)\sin(i)}{(M_*/M_\odot)^{1/2}(a/\text{au})^{1/2}(1 - e^2)^{1/2}}.$$

The semi-amplitude of the star's astrometric motion is:

$$\theta = (3.00 \mu\text{as}) \frac{(M_p/M_\oplus)(a/\text{au})}{(M_*/M_\odot)(d/\text{pc})}.$$

In both cases the period of the stellar motion (and therefore the planet's orbit) is also measurable. However the astrometric mass measurement requires that the system be observed for

⁴ Figure 4 and Equations (53) and (54).

⁵ Figure 1 and Equations (4) and (33).

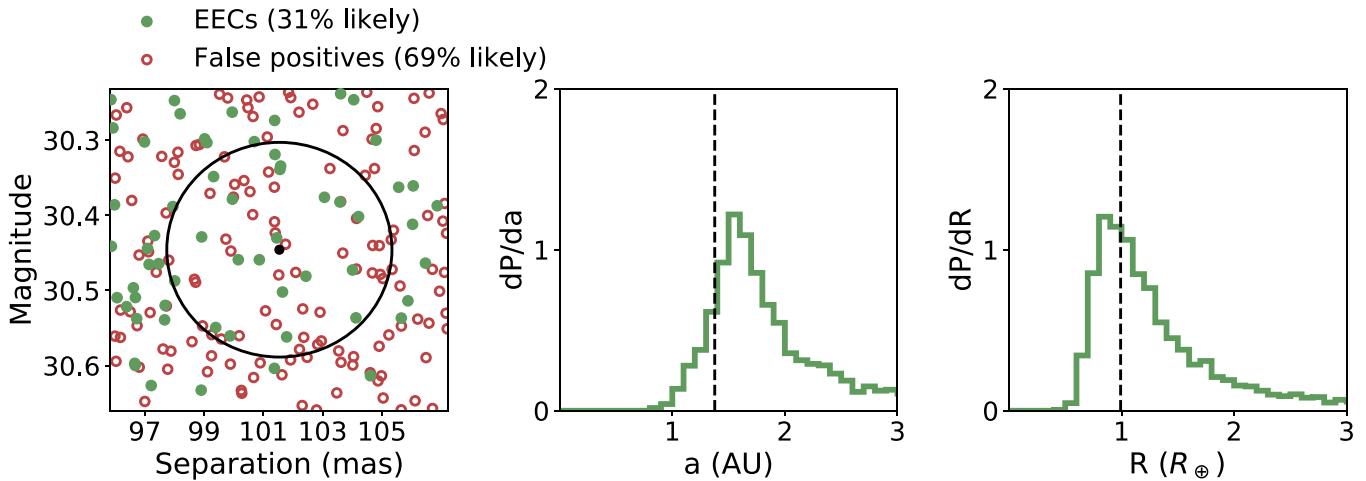


Figure 3. We simulate the detection of an “ideal” EEC from 15 pc. (Left) The separation- and magnitude-phase space populated by a host of simulated EECs (green) and non-EECs (red). The data point with uncertainties is marked in black. (Middle/right) Posterior distributions for the planet’s semimajor axis and radius, taking into account *Kepler* statistics and other priors. The true values are marked with dashed lines; this planet appears to the observer to have a wider orbit than it does in reality.

multiple epochs, in which case the planet’s orbital period can be derived from its apparent motion about the host star.

2.4. Calculating the Posterior Probability Distributions

We generate a posterior sample of simulated planets following the scheme in Section 2.1, where θ_{sim} are the simulated properties from Section 2.2, x_{sim} are the simulated data values from Section 2.3, and x_{obs} are the data values for the observed planet. This allows us to calculate a posterior probability distribution for each component of θ_{obs} , the physical properties of the observed planet. Based on the planet’s apparent separation and magnitude only, we are therefore able to place informative constraints on its semimajor axis, radius, and mass.

The posterior probability in which we are most interested is the probability that the planet is potentially habitable—i.e., a “terrestrial”-class planet ($0.8 \lesssim R_p \lesssim 1.6$) within the LWHZ. We designate these planets as “exo-Earth candidates,” following the terminology of The LUVOIR Study Team (2019) and The HabEx Study Team (2019), albeit including a slightly different range of sizes. Planets outside of this range we designate as “false positives,” with the *false positive probability* being the likelihood that the observed planet—as judged based solely on its observed data values—is a false positive instead of an EEC.

2.5. Example: An Exo-Earth Candidate Around a Solar Twin

As a demonstration of our method, we use the procedure in Section 2.2 to generate an Earth-sized planet in the center of the LWHZ of a Solar-type star at 15 pc—an ideal EEC. We simulate the detection of this planet by assuming a $S/N = 7$ measurement of its monochromatic magnitude, and a ± 3.8 mas measurement of its centroid.

Acting as the observer—who has no prior knowledge about the planet’s true size and orbit—we use the procedure outlined above to estimate the likelihood that that the planet is an exo-Earth candidate based on its observed magnitude and separation. The results of our analysis are summarized in Figure 3. We can confidently say that the planet is at least 1 au from its star, and is unlikely to be farther than 3 au—however, we cannot constrain its orbit to the habitable zone with

certainty. We can also tell that the planet is almost certainly larger than $0.5 R_{\oplus}$ and smaller than $3 R_{\oplus}$ —but this range includes planets which are too small (sub-terrestrial) or too large (sub-Neptune) to be EECs.

Most notably, we determine that it is only 31% likely that this planet is an EEC, as the majority of simulated planets with a similar separation and magnitude are not habitable. In this specific example, even though the planet appears 70% likely to have a size consistent with habitability, it also appears 60% likely to orbit beyond the maximum greenhouse limit. This example shows that while it will be difficult to discriminate between true EECs and false positives given only the data available on detection, it will still be possible to place meaningful probabilistic constraints on the planet’s properties.

3. Prior Assumptions

In this section, we discuss our priors in more detail by reviewing the relevant literature and discussing how they may be improved upon by future observations and modeling efforts.

3.1. Stellar Properties

It is likely that much effort will be dedicated to characterizing the stellar targets of a direct imaging mission in advance of its launch. Still, the stellar properties will only be constrained with finite precision—potentially several percent—so it is important that our prior sample includes host stars spanning the range defined by the relevant uncertainties.

Gaia DR2 (Gaia Collaboration et al. 2018) has already provided high-precision (10–100 mas) parallax measurements for nearby stars of spectral type F–M, so we treat this uncertainty as negligible. Optical/IR interferometry has allowed for the measurement of stellar radii to $\sim 3\%$ for targets at ~ 10 –100 pc (e.g., Ligi et al. 2016). Masses are more difficult to measure, so stellar atmosphere models are often used—as an example, Sharma et al. (2018) constrain model-dependent masses for more than 10,000 stars using high-resolution spectroscopy, with a median precision of 7%. Following these examples, we draw the radius and mass of the host star from normal distributions with widths of 3% and 7%, respectively.

3.2. Radius and Period

Kepler allowed for the precise calculation of planet occurrence for planets with periods shorter than 100 days. Mulders et al. (2018) find that the *Kepler* occurrence rates are well described by independent broken power laws in both radius and period:

$$\frac{dN_{\text{pl}}}{d \log P d \log R} \propto f_R(R) f_P(P)$$

where

$$f_P(P) = (P/P_{\text{break}})^{a_P} \quad (P < P_{\text{break}}) \\ = (P/P_{\text{break}})^{b_P} \quad (P \geq P_{\text{break}})$$

$$f_R(R) = (R/R_{\text{break}})^{a_R} \quad (R < R_{\text{break}}) \\ = (R/R_{\text{break}})^{b_R} \quad (R \geq R_{\text{break}}).$$

The best-fit parameters for this model are $(P_{\text{break}}, a_P, b_P) = (12, 1.5, 0.3)$ and $(R_{\text{break}}, a_R, b_R) = (3.3, -0.5, -6)$. The number of planets per system, N_{pl} , is found to be approximately five. Multiplicity will be an important factor for direct imaging surveys, as it could confuse the interpretation of the data or allow for simultaneous follow-up observations of multiple planets—but the topic is outside of the scope of this work. Here, we treat each detected planet independently, and normalize the above power laws so that $N_{\text{pl}} = 1$.

To properly simulate the abundance of planets in the habitable zone of stars of spectral type F, G, and K—as well as smaller or cooler planets which might be mistaken for them—requires us to extrapolate *Kepler* occurrence rates beyond the range of parameters within which they are well understood ($R \gtrsim 0.5 R_{\oplus}$, $P \lesssim 100$ days). This extrapolation could be problematic; for example, the results of Chen & Rogers (2016) suggest that planets on wider orbits can maintain thick volatile envelopes better against hydrodynamical escape, so we might find an over-abundance of large planets on wide orbits. We can gain some insight by studying the dependence of planet radii on insolation around low-mass stars, but these results cannot necessarily be extrapolated to Solar-mass regimes.

Furthermore, while *Kepler* was generally not sensitive to planets smaller than $0.5 R_{\oplus}$, it is likely that some such planets will be found by direct-imaging missions and could masquerade as EECs. It is therefore necessary that we extrapolate the power law of Mulders et al. (2018) down to $0.1 R_{\oplus}$ to ensure that the potentially large number of Mercury-sized objects are represented in our simulations. However, since our cutoff for EECs is $R \approx 0.8 R_{\oplus}$, planets smaller than $0.5 R_{\oplus}$ are less likely to be mistaken for EECs, so this extrapolation should not substantially influence our results.

3.3. Planet Classes

We employ a radius-based classification scheme to separate potentially habitable planets from those that are too small or too large to be habitable. This classification is motivated by two physical considerations affecting whether a planet can maintain a habitable atmosphere against irradiation over several gigayears.

Empirical evidence suggests a change in planet composition between 1.5 and $2 R_{\oplus}$. Multiple authors find evidence for a split in planet densities in this range, with planets larger than $\sim 1.5 R_{\oplus}$ mostly having densities much lower than that of the Earth (e.g., Weiss & Marcy 2014; Rogers 2015; Chen & Kipping 2017).

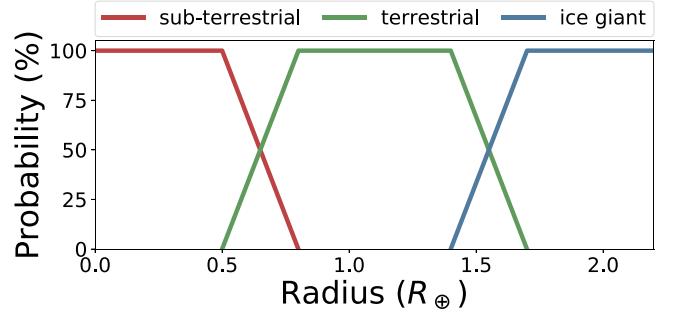


Figure 4. Our probabilistic scheme for classifying planets based on their radii. Sub-terrestrial planets are so small that they will lose their atmospheres to thermal escape under LWHZ levels of irradiation. Ice giants are so massive that they will form and maintain thick volatile envelopes. Only the terrestrial planets are neither too small nor too large to maintain a habitable atmosphere against irradiation.

Fulton et al. (2017) find a relative lack of *Kepler* planets with $R \sim 1.75 R_{\oplus}$ compared to smaller or larger radii; this “photoevaporation valley” was predicted by several authors who show that smaller planets would lose thick envelopes due to hydrodynamic escape (e.g., Owen & Wu 2013). We interpret both results as evidence that planets larger than ~ 1.4 – $1.7 R_{\oplus}$ have compositions more comparable to that of Neptune than to that of the Earth, and are therefore not habitable in the traditional sense.

Very small planets will also have trouble maintaining even small and dense atmospheres against Earth-like insulations. Zahnle & Catling (2017) find that a simple power law relationship between a body’s escape velocity and its effective insolation ($I \propto v_{\text{esc}}^4$) can predict whether planets in the solar system (and some exoplanets) have atmospheres. According to this relation, a planet with the same insolation and density as the Earth would need to be larger than $0.8 R_{\oplus}$ to maintain a habitable atmosphere. However, we note that the Earth lies toward the inner edge of the LWHZ as calculated by Kopparapu et al. (2014); it is possible that smaller planets could maintain Earth-like atmospheres further out.

Taking both of these considerations into account, we assign one of three classes to each planet based on its radius: “sub-terrestrial” planets which are too small to be habitable, “terrestrial” planets which could have Earth-like atmospheres, and “ice giant” planets which are too large. There is likely some overlap between these categories; for example, planets slightly larger than $1.4 R_{\oplus}$ or smaller than $0.8 R_{\oplus}$ might still have an Earth-like atmosphere. To simulate this overlap we probabilistically assign the class of each planet from its radius using the probabilities in Figure 4. The “terrestrial” class includes all planets with $0.8 < R < 1.4 R_{\oplus}$ and a fraction of planets with $0.5 < R < 0.8 R_{\oplus}$ or $1.4 < R < 1.7 R_{\oplus}$.

3.4. Mass

To calculate the mass of each planet, we rely on the empirical mass-radius relationships of Wolfgang et al. (2016, hereafter W16)⁶, which are calculated for smaller ($< 1.6 R_{\oplus}$) and larger ($> 4 R_{\oplus}$) planets, reflecting the bimodal split in planet composition. These relationships do not treat mass as a deterministic function of radius, but rather model a *distribution* of masses for each radius to capture the intrinsic variability in planet composition.

⁶ Equation (2) and Table 1.

We draw planet masses from truncated normal distributions defined by mean μ and variance σ^2 , with minimum values of 0.01μ and maximum values of $M_{\text{pure Fe}}$ —the mass of a pure iron composition as defined in W16. The parameters of the distributions are:

$$\mu, \sigma = \begin{cases} 2.7 M_{\oplus} (R/R_{\oplus})^{1.3}, 1.9 M_{\oplus} & (R \geq 1.6 R_{\oplus}) \\ 1.4 M_{\oplus} (R/R_{\oplus})^{2.3}, 0.3\mu & (0.8 < R < 1.6 R_{\oplus}) \\ 1.0 M_{\oplus} (R/R_{\oplus})^{3.0}, 0.3\mu & (R \leq 0.8 R_{\oplus}) \end{cases}$$

These are the values of μ and σ fitted by W16, with a few caveats:

1. Their data did not allow the authors to determine σ_M for the smaller planets; here, we arbitrarily choose $\sigma = 0.3\mu$ (i.e., a 30% spread in density).
2. Only a few uncertain data points and upper limits were available for planets with $R < 0.8 R_{\oplus}$, so we instead assume approximately Earth-like densities.
3. The large radius relationship was fitted for all planets with $R < 4 R_{\oplus}$, not just the ice giants; however, most of the precise data points nevertheless had $R > 1.6 R_{\oplus}$.
4. We model a few planets as large as $10 R_{\oplus}$, but this mass-radius relationship is likely not valid beyond $4 R_{\oplus}$; indeed, it underestimates the mass of Jupiter ($\sim 11 R_{\oplus}$) by a factor of five. Ultimately, the overlap in mass and magnitude between Earth-sized and Jupiter-sized planets is negligible when considering potential false positives for EEC detections.

Finally, we expect that the empirical mass–radius relationship will be improved upon in coming years by the discovery of transiting rocky planets around low-mass stars by *TESS* (Ricker et al. 2014), precision radius measurements from CHEOPS (Broeg et al. 2013), and mass measurements through TTV or RV. By the time an exo-Earth direct imaging mission begins, observers should have a better understanding of the relationship between the size, composition, and mass of a planet with which to interpret RV or astrometric mass measurements.

3.5. Eccentricity

Planets in the solar system tend to have eccentricities smaller than 0.1, but multiple authors find evidence for a wider distribution of exoplanet eccentricities in both transit (Kane et al. 2012) and RV data (Kipping 2013). Kipping (2013) determine that the eccentricity distribution of several RV-detected exoplanets is well described by a beta function, with $\alpha = 0.867$ and $\beta = 3.03$ —in which case $>50\%$ of planets have $e > 0.1$.

Eccentricity will have the effect of confusing the determination of a planet’s orbit from a single epoch of imaging data, as a wider range of eccentric orbits could be consistent with the observed separation. Furthermore, planets which orbit near the inner or outer edge of the LWHZ may spend a fraction of their orbit outside of the zone. To determine which of these planets are EECs we use the mean flux approximation—assuming that a planet is habitable if the average insolation of its orbit is the same as a circular orbit in the habitable zone; Bolmont et al. (2016) find this to be an adequate approximation for planets with low or modest eccentricities receiving a mean flux equal to that of the Earth, while highly eccentric planets tend to freeze

out. However, it is possible that this approximation is not valid for eccentric planets receiving a lower mean flux.

We use the beta distribution of Kipping (2013) to draw planet eccentricities and compute the projected separation accordingly, though we truncate the distribution beyond $e > 0.8$ (0.5% of planets) due to the additional computing time required to solve Kepler’s equation for highly eccentric orbits.

We note that multi-planet systems such as our own tend to have more circularized orbits (Van Eylen & Albrecht 2015). In principle, if a planet is detected with a companion then it is a priori less likely to have an eccentric orbit, and it should be easier to determine whether the planet is an EEC. In the scope of this paper however, we treat all planets as the only member of their system.

3.6. Albedo

We consider two different prescriptions for simulating the geometric albedo: in Cases 1–5 we assume a monochromatic albedo drawn from a broad uniform distribution, while in Cases 6 and 7 we assume a spectral albedo model which depends on the class and position of the planet with respect to the LWHZ.

Constraining the actual distribution of planet surface and atmospheric properties is one of the goals of future imaging missions, so it might seem backwards to interpret these observations by assuming the underlying distribution of geometric albedos as a prior. Nevertheless, such an assumption is necessary in order to infer a planet’s size from photometric data.

A uniform prior represents a conservative approach to the problem. We note that most planets in the solar system have geometric albedos ranging from 0.2 to 0.7 across the UV to NIR wavelength range, with the exception of the very low albedo of Mercury. Therefore for most cases we draw a monochromatic geometric albedo for each planet uniformly from 0.2 to 0.7.

However, the proposed designs of direct imaging missions allow for the simultaneous observations of a planet in two to three photometric bands, in which case color information would be available with no additional overhead. An efficient characterization strategy would make use of this color information to discriminate between EECs and their false positives, but to do so the observer must assume some prior knowledge about the diversity of planetary atmospheres and surfaces. As an example, we assume that all planets approximately reflect one of four solar system analogs—Earth, Venus, Mars, or Neptune—based on their radius-based classification and orbit as described in Table 3.

We accessed model spectra through the Virtual Planetary Laboratory⁷ for the Earth (Robinson et al. 2011, scaled from quadrature), Venus dayside, and Mars (no publications listed). For Neptune we use the planet’s observed geometric albedo as provided by Madden & Kaltenegger (2018) from 450 to 2500 nm, and set $A_g = 0.6$ from 300 to 450 nm (e.g., Mallama et al. 2017).

To simulate both model uncertainty and physical diversity among these solar-system analogs, we allow the spectral albedos to vary by ± 0.1 monochromatically and ± 0.05 in each photometric bandpass. Additionally, we enforce a lower limit of 0.001—to ensure that none of our planets are perfect

⁷ <http://depts.washington.edu/naipl/content/vpl-spectral-explorer>

Table 3

A List of the Solar System Analogs for which we Simulate Each Planet's Spectrum Under Cases 6 and 7

Class	Location	Assumed Model
Terrestrial	in LWHZ	Earth
Sub-terrestrial	anywhere	Mars
Terrestrial	exterior to LWHZ	Mars
Terrestrial	interior to LWHZ	Venus
Ice giant	anywhere	Neptune

Note. The assumed model for the planet's spectrum depends on its class or size, and location with respect to the liquid water habitable zone, as defined by the runaway and maximum greenhouse limits.

blackbodies—and an upper limit of 0.7, slightly more than the upper limit for the geometric albedo of a Lambertian sphere. The range of spectral models for each category of planet is demonstrated in Figure 5.

Finally, we consider the bandpasses in which the planets are observed. For reference, LUVOIR's proposed coronagraphic instrument would be able to observe simultaneously in 10% bandpasses of each of its three channels (L19). We choose wavelength ranges near the centers of each channel: 335–390, 715–830, and 1390–1610 nm.

It is likely that a wide variety of terrestrial planets exist; indeed, all four of the terrestrial planets in the solar system are distinct in surface reflectance and atmospheric absorption. Both of the ice giants however, have similar albedo distributions. This suggests that some understanding of the albedo distribution of ice giant analogs can be attained during the coming decade. Modeling efforts to understand the composition and appearance of sub-Neptune-type planets are already underway (e.g., Hu et al. 2015). New observatories such as the *James Webb Space Telescope* (*JWST*), the Wide-Field Infrared Survey Telescope (WFIRST), and ground-based Extremely Large Telescopes (ELTs) could provide observational tests of these models—the first through eclipse and transit spectroscopy of warm and hot Neptunes, and the latter two through direct imaging of ice giants at wide separations.

4. Mock Surveys: Methodology

The LUVOIR team has released an interim report (L18)⁸ which includes estimates for the number of planets which LUVOIR could detect as a function of planet radius, insolation, and host spectral type. In this section, we produce a candidate sample based on these estimates, characterize each planet therein, and report on the efficiency of each strategy outlined in Table 2 for properly identifying EECs. While we rely on the yield estimates and stellar targets of L18 for our mock survey, we use these only as baseline estimates for the yield of a generic, hypothetical direct imaging survey; we do not attempt to reproduce the results of the report or to assess the mission design.

4.1. Generating the Candidate List

To properly generate a list of planets which could be detected by a coronagraphic imaging mission requires careful treatment of the instrument design, survey strategy, and

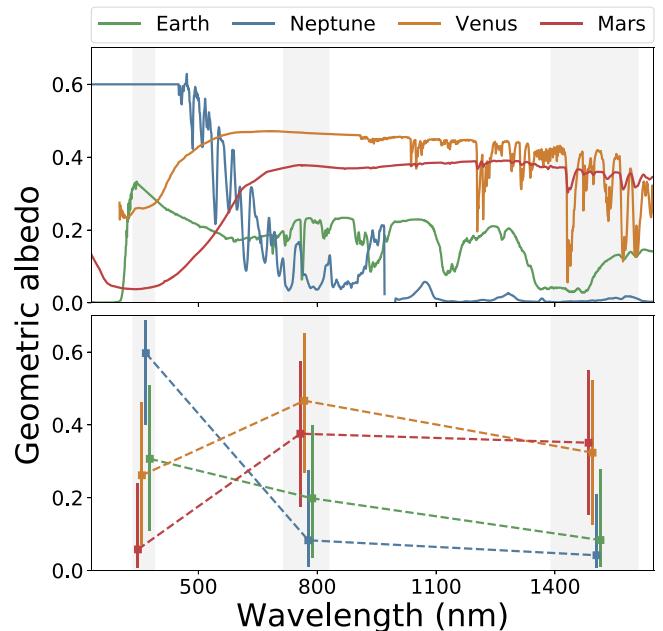


Figure 5. (Top) Set of base models used to simulate geometric albedos under Cases 6 and 7 only, according to the scheme in Table 3. (Bottom) The actual range of simulated spectra for each type of planet as observed in three photometric bandpasses (with wavelength offsets for visibility). We apply moderate differences to the spectrum of each planet to simulate the underlying physical diversity, but broad differences between the four groups can still be seen.

additional sources of noise (e.g., dust) which is outside of the scope of this work. L19 have performed such an analysis (following the methodology laid out by Stark et al. 2014) so we use their results to generate our candidate list.⁹

We acquired one of the simulated samples of host stars upon which the LUVOIR-A yield estimates are based, including masses, luminosities, and distances for 287 stars (C. Stark 2019, private correspondence). To assign a radius to each star we use a simple scaling relation (Hansen et al. 2004):

$$(R_*/R_\odot) = (M_*/M_\odot)^{0.8}.$$

While this simulated list will not be the final target list of LUVOIR-A, it generally represents the diversity of host star properties which such a survey would encounter. Next, we use the method outlined in Section 2.2 to generate a large sample of planets around these stars, and draw from that sample at random until the yield estimates for LUVOIR-A have been satisfied for each bin in spectral type, planet size, and insolation.

To ensure that the planets we simulate are detectable, we also enforce the same separation and brightness limits as L18. Namely, we only include targets which are brighter than a planet-to-star contrast ratio of 2.5×10^{-11} and which are detected between 24 and 440 mas—approximately the working angles of the proposed LUVOIR-A coronagraph at 500 nm.

The yield estimates of the LUVOIR-A architecture project the discovery of approximately 450 planets, about 50 of which would be EECs. For our purposes, the actual number of planets is not relevant—only their relative abundance by size, insolation, distance, and host spectral type—so we improve

⁸ The LUVOIR Final Report (L19) was released with slightly updated yield estimates while this paper was under review. The results here are based on the numbers given in the interim report.

⁹ See Figures 3.1 and 3.13 and Table 8.7 for details on the target list, yield estimates, and instrument parameters.

the accuracy of our results by inflating the yield estimates unilaterally by a factor of fifty.

4.2. Survey Cases

We run our mock surveys under seven cases representing the different data which could available to the observer. A brief description of each can be found in Table 2.

4.2.1. Case 1: Detection Data Only

Under Case 1, the planet's existence is entirely unknown before its direct imaging detection, and the observer only measures its apparent separation and monochromatic magnitude. Following L18, we assume a S/N of 7. If photon noise is dominant, the uncertainty on the planet's centroid position is described by:

$$\sigma_c = \frac{\text{FWHM}}{\text{S/N}}.$$

Assuming an effective wavelength of 500 nm, then $\sigma_c = 3.8$, 1.9, and 1.0 mas for the 4 m HabEx and 8 or 15 m LUVOIR architectures. This uncertainty will be further affected by the pointing stability of the telescope during observations and the wavelengths at which the planet is observed; to be conservative, we choose $\sigma_c = 3.8$ mas. Finally, we simulate the planet's detection by re-drawing its magnitude and apparent separation from normal distributions with widths defined by these uncertainties.

4.2.2. Case 2: Additional Radial Velocity Detection

It has been emphasized that a RV search for nearby Earth twins would be an important precursor to a space-based direct imaging mission (Dressing et al. 2019), but the detection of Earth analogs around Solar twins is beyond the reach of current instrumentation. Doing so would require a significant investment of time on major observing resources and new methods to correct for systematic sources of noise such as stellar jitter (Plavchan et al. 2015).

To investigate the potential benefits of a precursor RV search for interpreting planet detections, in Case 2 we simulate the direct imaging detection of a planet (Case 1) along with a measurement of its orbital parameters and RV semi-amplitude. We set a baseline uncertainty of 5 cm s^{-1} on the measurement of the RV semi-amplitude K —this value is chosen to be smaller than the value for an Earth twin ($\sim 10 \text{ cm s}^{-1}$), but not negligible in comparison. Additionally we allow a conservative 10% uncertainty on the measurement of the orbital period, and a loose $\pm 30^\circ$ constraint on the planet's mean anomaly.

Several planets—including small planets in the LWHZ which could masquerade as exo-Earths candidates—will have $K < 5 \text{ cm s}^{-1}$. In these cases we assume the planet is undetected, so no constraints on its orbit are available. An upper limit of 10 cm s^{-1} is enforced so that the nondetection also carries useful information about the planet's size.

4.2.3. Case 3: Constraining the Orbital Plane Using Debris Disks Measurements

It is expected that several nearby systems contain exozodiacal dust disks near their habitable zones (Ertel et al. 2018). This dust may be a significant source of background and confusion noise for future direct-imaging surveys and could be a key

driver of aperture size, with larger apertures collecting less background light within their resolution element (Roberge et al. 2012). However, these disks could also carry useful information about the orientation of the system which could help to establish the orbit of a newly detected planet, and their observation would require no additional resources.

Observations of larger protoplanetary disks have yielded tight (better than 1°) constraints on the disk inclination and orientation (e.g., HL Tau, ALMA Partnership et al. 2015). While the orientation of exozodiacal disks may be harder to constrain if they are faint or exceed the outer working angle of the coronagraph—and while it is possible that not all systems have substantial disks—here we consider the “optimistic” case where every system has a disk, and the orientation parameters ($\cos(i)$ and Ω) can be constrained to $\pm 1\%$.

In principle, if the orientation of the debris disk can be tightly constrained and the planet shares *exactly* its orbital plane with a circular orbit, then a single precise measurement of the planet's apparent separation could be sufficient to determine its phase and semimajor axis. However, the orbital inclinations of solar system planets deviate by up to 5° from the Solar spin axis—and larger misalignments could be possible in other systems—while several of our simulated planets have eccentric orbits.

We can still infer some information about the orbit of a planet given the orientation of its disk and some prior knowledge about how misaligned planetary systems tend to be. In Case 3, we use the orientation of a contemporaneously detected disk as a prior constraint on the mean orbital plane of the system, but allow for a difference of ± 0.2 between $\cos(i)$ and the longitude of the ascending node (Ω) of the two components. We then interpret the observations using the same data as in Case 1, but treating the two-dimensions of the separation vector separately.

4.2.4. Cases 4 and 5: Multiple Revisits and Astrometric Mass Measurements

Constraining the planet's orbital parameters will require multiple revisits spaced over the orbital period, so revisits will likely be folded into the observing strategy of future direct-imaging missions. However, even within the habitable zone there will be numerous potential false positives for EECs, a fact which may limit the practical benefit of revisiting every system.

Guimond & Cowan (2019) have shown that approximately three revisits will be sufficient to constrain the orbital parameters with better than 10% precision. In Case 4 we assume that the system has been revisited enough times for the orbital parameters of the planet to be constrained with comparable precision. Specifically, we assume that the semimajor axis is measured to $\pm 10\%$, the eccentricity to ± 0.05 , and the orbital phase to $\pm 15^\circ$.

An ancillary benefit of revisiting targets would be the ability to measure the star's astrometric motion about the system's center of mass. Measuring the astrometric semi-amplitude θ would allow observers to determine the planet's mass and potentially to identify which planets are too small or large to be habitable. In Case 5 we assume that, in addition to the constraints provided in Case 4, the observer can measure θ with a precision of $\pm 0.1 \mu\text{as}$. This is the targeted astrometric precision of the High Definition Imager instrument with the LUVOIR-A aperture (L19), and approximately half of the amplitude induced by an Earth twin at a distance of 15 pc. If $\theta < 0.1 \mu\text{as}$, an upper limit of $0.2 \mu\text{as}$ is enforced instead.

4.2.5. Case 6: Color Measurements

Future direct imaging missions could be able to observe simultaneously in multiple bandpasses, in which case measurements of the planets' colors would be available as soon as they are detected. Color information could be used for preliminary planet classification; for example, Batalha et al. (2018) predict that it will be possible to differentiate between cloudy and cloud-free Jovian planets with reflected light imaging in three filters using WFIRST or ELTs.

However, color information is only useful for inferring the planet's properties if we make a prior assumption about the spectral albedos of extrasolar planets. Some observational constraints are currently available through eclipse observations of close-in giant planets, but for smaller planets we can only rely on planets in the solar system and theoretical models of the surfaces and atmospheres of known transiting and RV-detected planets. It is therefore worthwhile to determine what effect an approximate prior understanding of spectral albedos would have on the interpretation of direct imaging data.

In Case 6 we simulate planets with spectral albedos that resemble (though are not identical to) those of solar system analogs; our detailed assumptions are described in Section 3.6. Each planet is observed in a 10% bandpass at the center of three wavelength channels, with a S/N weighted by the square root of the bandpass-integrated flux (i.e., photon noise). The S/N integrated across all three bandpasses is 7, as in Case 1. If $S/N < 2$ in a given bandpass (typically in the UV or infrared), then a 2σ upper limit is enforced instead.

4.2.6. Case 7: Maximum Information

In the final case we consider all of the information which a larger telescope would be able to acquire on a target after several revisits. These include measurements of the orbital parameters and astrometric semi-amplitude as well as brightness measurements in three bands. We assume the same measurement precision and sensitivity limits as in Cases 4 and 5, and we interpret the color information following the method of Case 6.

4.3. Classifying the Targets

Once we have constructed a sample of targets and simulated their detection, we classify them using the inference framework described in Section 2. This yields for each planet the likelihood, according to the observer, that the planet is an EEC or a false positive. In Figure 6 we plot the likelihood that the planet is an EEC for each simulated EEC in the sample. Since the observer does not know a priori that these planets are EECs, we see that they cannot make a confident identification using the initial detections of the planet, but given additional data or multiple revisits they can identify several EECs with confidence.

We can break down the false-positive probability by size and orbit, determining for each observed EEC the inferred likelihood that it is, for example, a sub-Neptune on an orbit exterior to the habitable zone. In Figure 7 we plot this probability for each combination of class and orbit averaged over the sample of observed EECs. This plot illuminates the key sources of ambiguity in classifying EECs. For example, we see that it is difficult to distinguish between an EEC and a sub-terrestrial planet with a temperate orbit, or a planet which has the proper size but lies just inside or outside the habitable zone.

5. Mock Surveys: Results

5.1. Which Planets are "False Positives"?

Figure 7 demonstrates that true EECs share the observable parameter space with a wide range of planets both within and outside of the LWHZ. These can be broadly separated into: (i) planets which are too small or large to be habitable, but are yet in the habitable zone, (ii) planets which are of the proper size to be habitable, but are not in the habitable zone, and (iii) planets which are both of the wrong size and not in the habitable zone. Each of these categories are approximately equal in their potential to masquerade as EECs.

As we demonstrate in Figure 6, the observer will typically be unable to distinguish between true EECs and their many potential false positives given just the data necessary for the planet's detection. Even when a planet which is *in fact* an EEC is detected, the observer will only be able to make this determination with $<50\%$ confidence.

5.2. Do Constraints on the Orbit Help to Identify EECs?

Since most of the "false positives" are planets outside of the habitable zone, it stands to reason that measurements which constrain the planet's orbit would be useful for identifying EECs. Indeed, if the planet can be independently detected through RV, and its period constrained with 10% precision, then it can typically be constrained to the habitable zone with $>80\%$ confidence. However, approximately 25% of EECs remain below our 5 cm s^{-1} detection limit, in which case the orbit cannot be established.

Observing the planet multiple times over an orbital period will help to rule out planets outside of the habitable zone with similar confidence, assuming 10% uncertainties on the orbital parameters. Nevertheless, even if a planet can be constrained to the habitable zone, it may yet be too large or small to be habitable. In general, constraints on the orbit and phase will only allow the observer to distinguish EECs from temperate sub-terrestrial planets or ice giants with $\sim 50\%$ confidence.

5.3. Do Constraints on the Mass Help to Identify EECs?

We find that measurements of the astrometric semi-amplitude θ , when combined with magnitude measurements, can modestly increase the ability of the observer to identify EECs. For example, our average confidence for identifying EECs given multiple revisits to determine the orbit (Case 4) improves by about 18% if we include an astrometric measurement or upper limit on the planet's mass (Case 5), and several individual EECs can be identified with very high confidence. These could be the highest-priority targets for deeper spectroscopic follow-up.

Under Case 2, if both the measured period and RV semi-amplitude K are used to constrain the properties of the planet then EECs can be identified with an average confidence of 52%. If the measured value of K is ignored however, then this confidence drops to 40%. In other words, the measurement of the planet's mass affords an extra 12% confidence that the planet is an EEC.

Constraining the orbit generally reduces the potential for false positives more than it constrains the mass, in part because constraints on the planet's size are available based on its brightness alone. On the other hand, the relationship between radius, mass, and composition is better understood than the

How confidently can we identify exo-Earth candidates in mock observations?

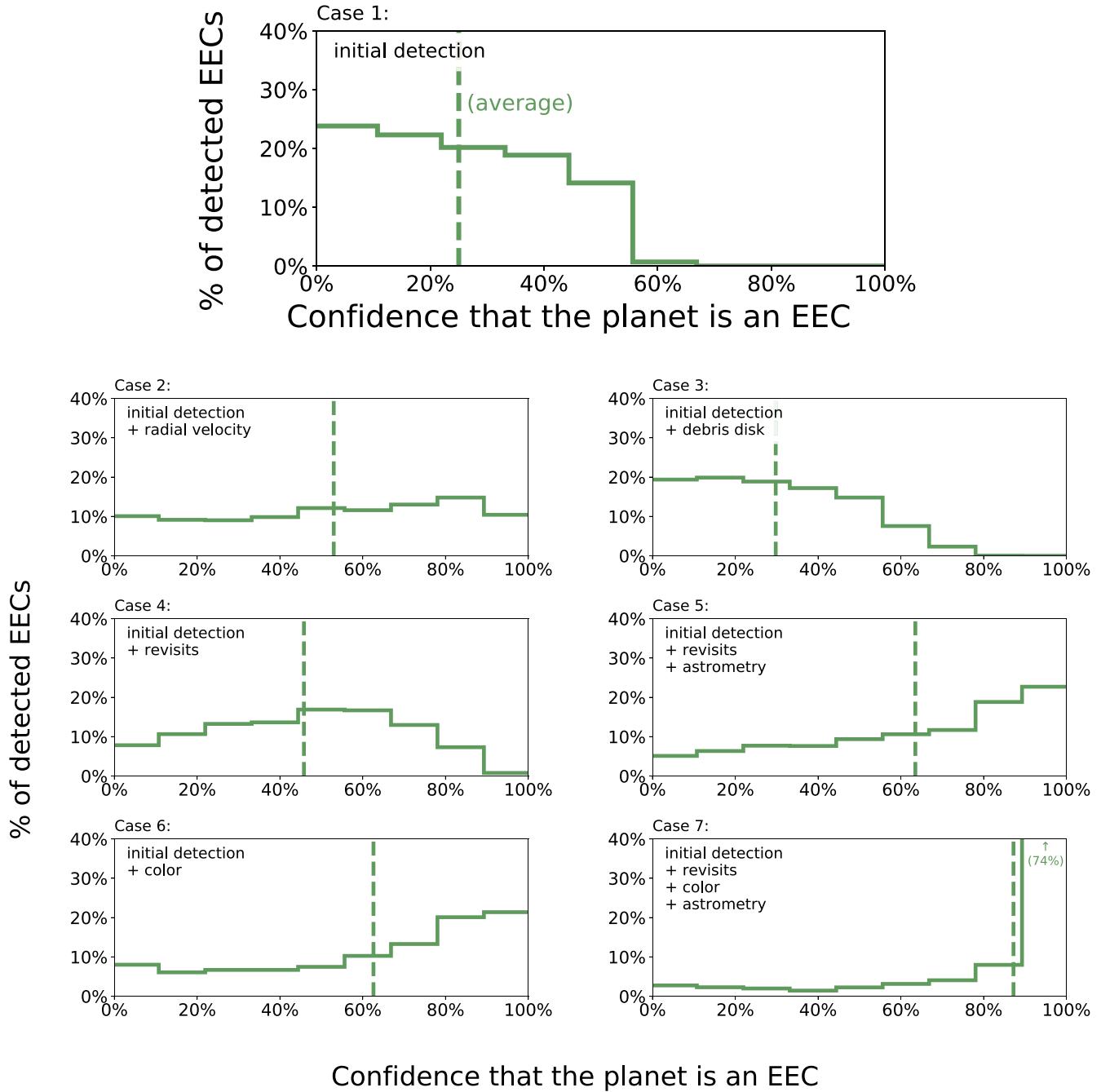


Figure 6. We simulate the detection of approximately 2500 EECs around nearby stars, then infer their properties from the mock data under each of the cases in Table 2. Above we plot the probability—as inferred by an uninformed observer—that each detected planet is an EEC, as well as the average value (dashed line). In the ideal case, this value would be 100% for all EEC targets, but typically it is smaller because of the limited data available to the observer. With additional data (Cases 2–7), the observer can be more confident that the detected planet is an EEC.

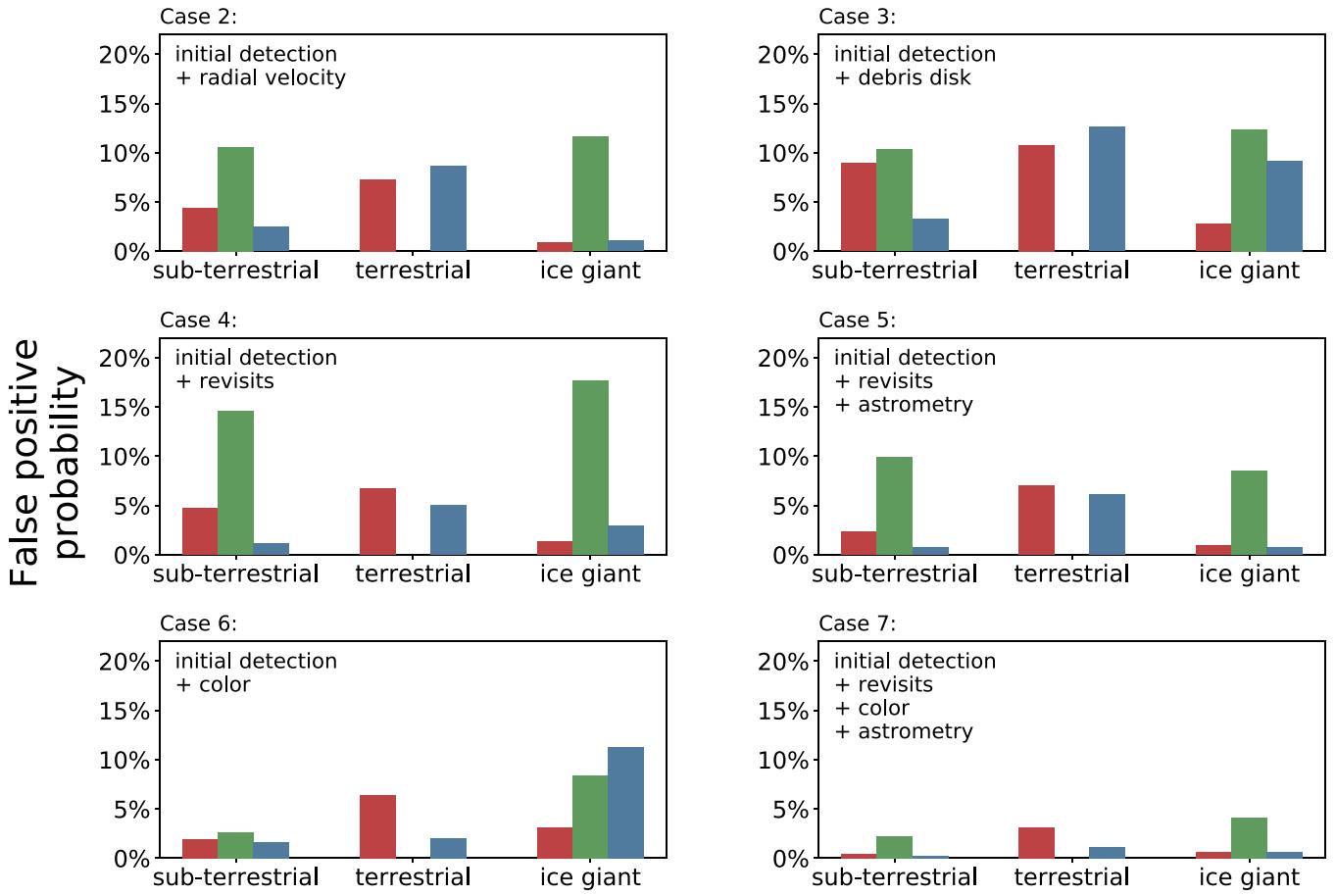
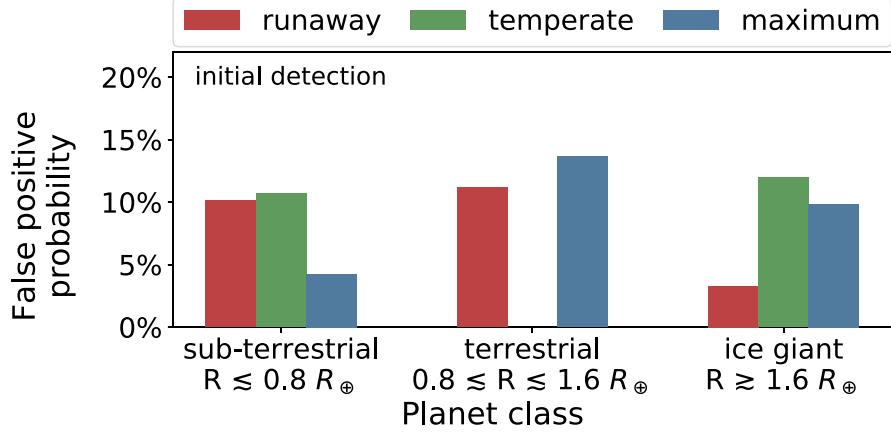
prior distribution of planet albedos (which almost certainly is *not* uniform). An inference about the composition of the planet made from its mass could be more reliable than one made from its apparent magnitude.

5.4. Do Color Measurements Help to Identify EECs?

In Figure 8 we plot the average inferred probability that an EEC is similar to the Earth, Neptune, Venus, or Mars for Cases 1

(no color) and 6 (color). We find that color information is useful for distinguishing between spectra with positive versus negative slopes between the UV and visible channels. Specifically, adding a color measurement allows the observer to distinguish between Earth and Martian analogs particularly well—if small planets in the habitable zone tend to look like Mars, color information will be a valuable discriminant. According to Figure 6, observing the color *and* constraining the orbit of a planet could allow the observer to identify most EECs with >80% confidence.

What kinds of planets could be mistaken for exo-Earth candidates?



Planet class

Figure 7. We simulate the detection of approximately 2500 EECs around nearby stars, then infer their properties from the mock data under each of the cases in Table 2. Above we plot the sample-averaged probability—as inferred by an uninformed observer—that the planet is instead a false positive with a nonhabitable class or orbit. Additional data will suppress the false-positive probability—for example, fitting the orbit with finite precision will reduce the likelihood that the planet is outside of the habitable zone.

We choose solar system planets as our templates as they cover a relatively broad range of insolations and planet sizes, and with the exception of very hot exoplanets they remain the only planets for which albedo measurements are presently

available. In our set of models shown in Figure 5, the Earth stands out due to its modest scattering slope in the optical versus a much stronger feature in Neptune's spectrum, or opposite features for Venus and Mars. Yet a habitable planet

Does color information help to identify exo-Earth candidates?

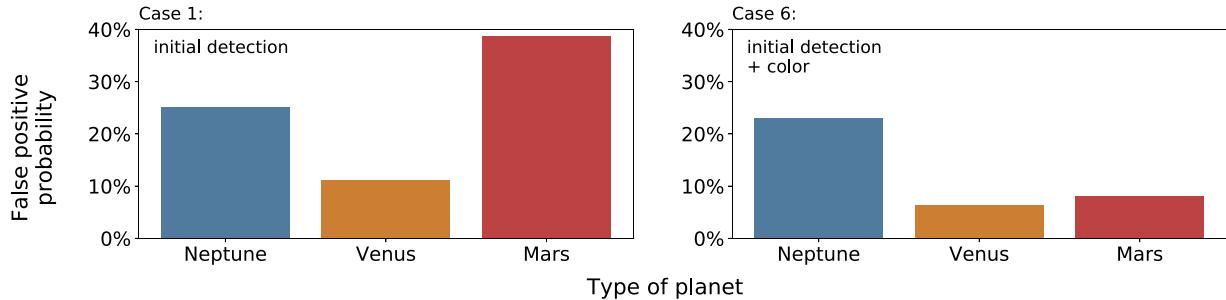


Figure 8. We test whether low-to-moderate S/N color measurements would help to identify EECs, assuming that all planets *approximately* resemble a solar system analog depending on their class and orbit (Table 3 and Figure 5). Plotted is the typical inferred probability that the EEC is actually a false positive with a Neptune, Venus, or Mars-type surface and atmosphere. Without color information (left) it is difficult to distinguish Earth analogs from smaller or cooler Mars-like planets, or larger Neptune-like planets. With color information (right), the slope of the optical spectrum provides a useful discriminant between Earth and Mars, but does little to reduce the ambiguity due to larger Neptunes.

does not need to look like the Earth, and indeed there is evidence that during the Archean the Earth had a substantially redder appearance due to organic haze particles (e.g., Arney et al. 2016, 2017). Similarly, the solar system provides no examples of an Earth-sized planet beyond the habitable zone. Here we assume such worlds have a Martian appearance, but this is likely inaccurate for icy worlds or planets with dense atmospheres.

We compare our range of simulated spectra to the models of Krissansen-Totton et al. (2016), who compute optimal photometric bandpasses to distinguish between several different examples of potential exoplanet reflectance spectra. Their optimized bandpasses are 431–531 nm (“blue”), 569–693 nm (“green”), and 770–894 nm (“red”)—different from those used in this work. In Figure 9 we place our models on a color–color plot similar to Figure 2 in the latter-mentioned study, along with a subset of the models considered therein. We find that our solar-system analogs with simulated physical diversity cover a comparable range in color space, so we believe that we adequately represent a diversity of planet appearances even though our range of base models is limited.

We stress that this result is sensitive both to our prior assumptions and to the bandpasses in which we choose to observe our targets. More work must be done to understand the potential diversity of terrestrial planets and to determine which photometric bandpasses are optimal for distinguishing them from EECs (e.g., Krissansen-Totton et al. 2016). In Section 5.9 we discuss how observational constraints on the albedos of potential false positives could be derived within the coming decade. These new discoveries can then be folded into our Bayesian framework, and the results of Case 6 suggest that doing so could allow for the confident distinction between false positives and true EECs on the basis of color information.

5.5. Can a Debris Disk be Used to Constrain the Orbital Plane?

We find that measuring the orientation of the debris disk provides relatively little information about the orbit of the planet. Specifically, we are on average only $\sim 5\%$ more confident that the observed EECs lie within the habitable zone when we have measured the orientation of the disk. The benefits are slightly greater in cases where the system is observed from a “pole-on” orientation ($|\cos(i)| > 0.5$), in which case the uncertainty in the centroid measurement

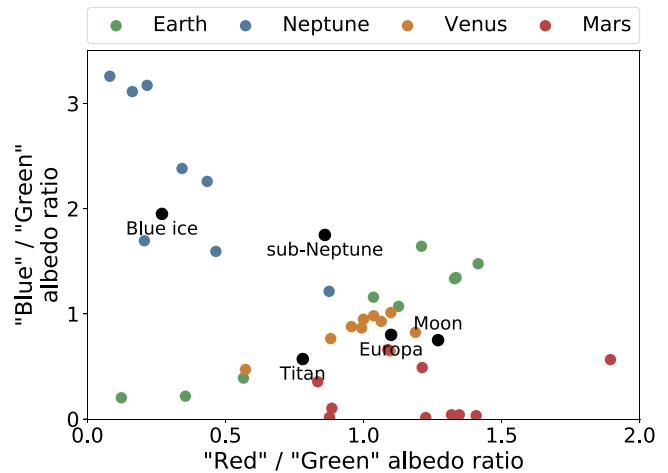


Figure 9. We plot a set of simulated Earth, Neptune, Venus, and Mars analogs in color–color space—each planet’s spectrum is modified from the solar-system model to simulate physical diversity. We also plot data points from Krissansen-Totton et al. (2016; Figure 2) who calculate these color ratios for a large diversity of solar-system bodies, materials, and exoplanet models. While we consider fewer base models than these authors, our simulated planets cover a similar range in color space.

translates to a smaller uncertainty in the semimajor axis versus the “edge-on” cases ($|\cos(i)| < 0.5$). Nevertheless, unless the centroid precision is much better than 3.8 mas and the observed systems are at least as well-aligned as the planets in the solar system, measuring the disk orientation will generally not allow an observer to constrain the orbit without revisits.

5.6. Can EECs be Identified given Maximal Photometric Information?

In Case 7 we assume that the observer has revisited the system multiple times to constrain the orbit with three-band photometry and has additionally measured (or placed upper limits on) the astrometric motion of the star due to the planet. This is the most information that could be obtained for the typical system without substantial follow-up or precursor observations, and we find that it would allow the observer to confidently identify most EECs, with an average confidence of 87%. This suggests a promising roadmap toward selecting targets for follow-up, but as in Case 6 this result is dependent on the observer’s prior assumptions about exoplanet reflectance spectra.

Efficiently surveying planets for spectroscopic H₂O absorption

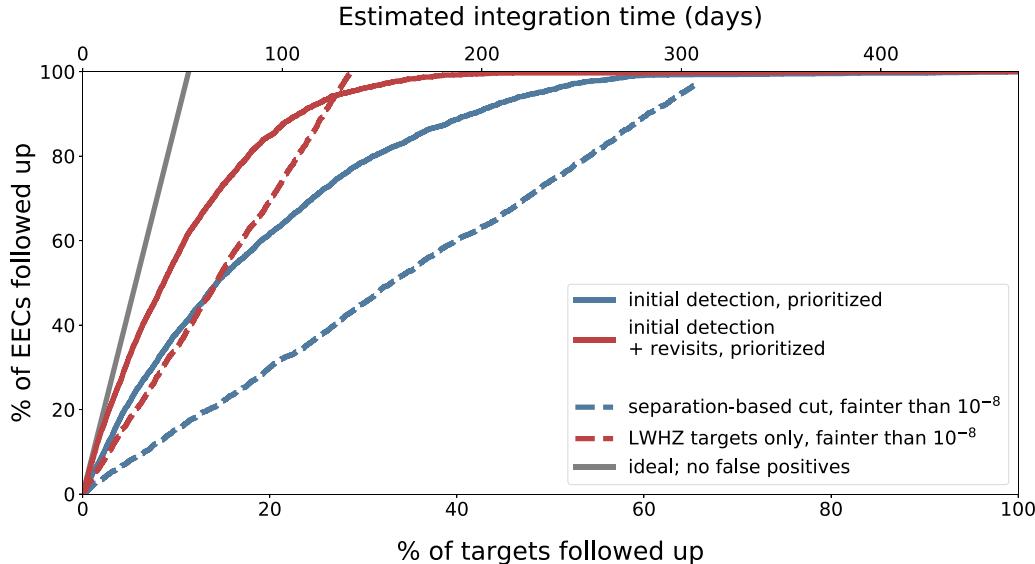


Figure 10. Efficiency of different post-detection follow-up strategies (e.g., to search for water absorption), quantified as the percentage of all targets that must be followed up before a given percentage of EECs have been covered; the ordinate is proportional to the amount of time required for follow-up. The gray line represents ideal survey efficiency, where no false positives are re-observed. The colored lines represent strategies which make use of the detection data only (blue) or multiple revisits to establish the orbit (red). The dashed lines are nonprioritized (blind) approaches which first remove very bright or widely separated planets, or planets whose orbits are constrained to be outside of the LWHZ. The solid lines are prioritized approaches in which we observe the planets which are most likely to be EECs first. The upper axis estimates the amount of time required to search the atmosphere of each planet for water absorption, assuming 25 hr of integration time per target.

5.7. Would Bayesian Prioritization Improve Follow-up Efficiency?

A practical way to frame our results is in terms of follow-up efficiency. A logical next step after detecting a directly imaged planet (and optionally constraining its orbit) would be to search for water absorption in a narrow part of the spectrum to further test its habitability. However, unless a strategy is employed for prioritizing or pruning the target list, a significant amount of time will be spent “following-up” nonhabitable planets. Indeed, Kawashima & Rugheimer (2019) have shown that 3–10 hr may be required to confidently detect H₂O absorption in an exo-Earth atmosphere from 5 pc using a 10 m telescope. For the typical target observed by the 15 m LUVOIR-A from 15 pc, approximately 25 hr may be required.

Our inference framework allows us to prioritize targets based on the likelihood that they are indeed EECs. After probabilistically classifying each observed target—including both EECs and nonEECs—we prioritize them by the probability that they are true EECs and submit them for follow-up observations with the most likely candidates first. In Figure 10 we summarize the efficiency of this prioritization strategy for Cases 1 (initial detection only) and 4 (multiple revisits).

We compare our Bayesian approach to two nonprioritized follow-up strategies: first, removing all targets whose projected separations are wider than the maximum greenhouse limit, and second, removing all targets that are found to be outside of the habitable zone after multiple revisits. In both of these cases we also remove targets with a contrast ratio brighter than 10⁻⁸. These separation- and magnitude-based cuts exclude $\lesssim 6\%$ of bright or eccentric EECs but several larger or cooler false positives. The remaining targets are then followed up blindly. Finally, we plot a line representing a perfectly efficient follow-up strategy with no false positives.

We find that prioritizing the targets using our Bayesian framework allows us to re-observe them with much greater efficiency—using the *same* data—than a blind approach. In fact, a prioritized approach using just the detection data is initially *as efficient* as taking a blind approach after each planet’s orbit has been characterized. We can discuss these efficiency gains in terms of integration time. Assuming 450 targets have been discovered in the initial census and 25 hr of integration time are required to search each planet for water, the results in Figure 10 suggest that probabilistic target prioritization could reduce the required amount of integration time to follow-up 50% of EECs by 28 days (if the orbits have been precisely constrained) to 95 days (if only the detection data are available). Since this prioritization scheme does not require additional data on the system, it could be naturally folded into the survey strategy of any direct imaging mission.

5.8. Is the Bayesian Approach Always Appropriate?

In principle, a Bayesian prioritization scheme should always be superior to a blind follow-up strategy because it leverages additional information about exoplanet demographics. In practice, this prior knowledge is always incomplete and potentially inaccurate, which could make Bayesian prioritization risky in the sense that it might fall prey to unexpected types of false positives. For example, to produce planets for our mock surveys we use the yield estimates of L18 which include more sophisticated treatments of the mission design and a different estimation and extrapolation of *Kepler* occurrence rates—but then we use our own algorithm to probabilistically classify each planet. Therefore, the prior assumptions we make when characterizing these simulated planets are not exactly representative of their true parent distribution, which makes our prioritization less efficient.

Indeed, we see in Figure 10 that if the planets’ orbits are known, then the blind approach to prioritizing targets is *more* efficient than the Bayesian prioritized approach for the last $\sim 5\%$ of EECs. However we argue that this is realistic—the observer’s prior assumptions will always differ from reality to a degree—and we note that our approach is still efficient despite this mismatch. For this reason we believe Bayesian prioritization will yield results that are better overall, but observers might prefer a blind approach for low-priority targets.

5.9. What Priors Can be Improved in the Coming Decades?

The accuracy of our probabilistic classification depends on the extent and accuracy of existing knowledge of exoplanet statistics. Considering the state of the field ~ 20 yr ago, it is reasonable to assume that an observer using this method to interpret observations ~ 20 yr from now will base their judgment on better prior assumptions than are currently available. Here we speculate on ways in which the prior inputs to our method could be refined in the coming decades.

5.9.1. What is the Largest Potentially Habitable Planet?

In our work we assume that planets larger than $\sim 1.6 R_{\oplus}$ will form and maintain volatile envelopes over gigayear timescales, thereby resembling “mini-Neptunes” more than “super-Earths.” The most compelling evidence for this comes from density measurements for a limited sample of small planets (Weiss & Marcy 2014; Rogers 2015; Chen & Kipping 2017) and the gap in *Kepler* radius occurrence rates near $\sim 1.7 R_{\oplus}$ (Owen & Wu 2013; Fulton et al. 2017), but the exact value of this transition radius and its dependence on parameters such as insolation and spectral type requires further research (e.g., Fulton & Petigura 2018; MacDonald 2019; Martinez et al. 2019). In the near future, *TESS* (Ricker et al. 2014) will detect hundreds of small planets orbiting bright stars (Sullivan et al. 2015; Bouma et al. 2017; Barclay et al. 2018; Ballard 2019). In combination with precise radii measurements by CHEOPS (Broeg et al. 2013) and ground-based RV measurements, these discoveries will enhance the sample of $1-2 R_{\oplus}$ planets with measured densities. Later on, PLATO (ESA 2017) will detect several hundred planets smaller than $2 R_{\oplus}$ on orbits as wide as that of the Earth, providing statistics for planet radii over a broader range of insolations than *Kepler*. Models of atmospheric evolution can be used to combine these lines of evidence into a more comprehensive understanding of which planets should have nonhabitable volatile envelopes (e.g., Owen & Wu 2013; Lopez & Fortney 2014; Gupta & Schlichting 2019).

5.9.2. What is the Period Distribution for Planets on Wide Orbits?

This work relies on extrapolation from *Kepler* occurrence rates for shorter-period planets, but ice giants in the habitable zone and beyond could be a significant source of false positives for EEC detection. WFIRST could detect hundreds of wide-orbit planets through its microlensing survey, some with masses lower than that of the Earth (Barry et al. 2011; Penny et al. 2019). PLATO will also detect a number of transiting planets on orbits as wide as that of the Earth (ESA 2017). Even if these data do not fully probe the relevant range of planet radii and periods, they may provide enough points for interpolation

to accurately predict the occurrence of terrestrial planets and ice giants within the habitable zone and beyond.

5.9.3. What do the Spectra of False Positives Look Like?

Our assumed distribution of planet albedos is the least well-vetted prior assumption in this work, but in Cases 6 and 7 we show that prior knowledge of planet albedos could greatly enhance survey efficiency. The most promising avenue for constraining planet albedos in the next decade is through direct imaging of super-Earths and sub-Neptunes in the habitable zone and on wider orbits with ELTs and WFIRST (e.g., Kasper et al. 2010; Savransky & Garrett 2016; Traub et al. 2016; Artigau et al. 2018; Weinberger et al. 2018; Akeson et al. 2019). These observations would be valuable for determining the optimum wavelength ranges for discriminating between ice giants and other types of planets.

Several authors have shown that post-runaway atmospheres like that of Venus could develop on extrasolar planets comparable in size to the Earth within the inner edge of the habitable zone, which itself depends on the spectral type of the host star and planetary factors such as mass, rotation rate, and ocean coverage (e.g., Kopparapu et al. 2013, 2014; Yang et al. 2014; Kodama et al. 2018). Transit spectroscopy with *JWST* could reveal the atmospheric composition of a small number of terrestrial planets within the runaway greenhouse limit of low-mass stars (Lustig-Yaeger et al. 2019). While not directly measuring the albedo, these observations could reveal whether Venus analogs are common, and modeling efforts could reconstruct their likely appearance in reflected light.

5.9.4. Testing Priors Against New Data

Finally, it will be possible for observers to validate their priors during the course of the direct imaging survey. As a simple example, if more faint planets are discovered than predicted under the assumed priors, then it is likely that either the occurrence rates under-predict at small radii or that there are more low-albedo planets than are present in the solar system. This information could then be forwarded into a probabilistic analysis of which planets are most likely to be Earth-like before substantial time is committed to follow-up observations.

6. Conclusions

We have developed a Bayesian framework with which to infer the properties of a directly imaged planet on the basis of limited photometric data, with the primary goal of identifying EECs for deeper spectroscopic follow-up. This framework is dependent on a multitude of priors drawn from observed exoplanet statistics and a few theoretical models. We use it to characterize the ability of future direct-imaging missions to identify potentially habitable planets upon their initial detection using only photometry. We determine the key ambiguities involved in this determination and explore potential solutions, such as constraining the orbit through multiple revisits.

Assuming a uniform prior on a monochromatic albedo, we find that the detection data alone is not sufficient to determine whether the planet has a potentially habitable size or orbit. In the best cases, a few EECs could be identified with $\sim 50\%$ confidence, but the average EEC would only be identified with $\sim 25\%$ confidence. This translates to a potential false discovery

rate of $\sim 75\%$, consistent with previous results (Guimond & Cowan 2018).

Constraints on the planet’s orbit could be achieved through a precursor RV survey or by revisiting the system multiple times. This would allow the observer to constrain the planet to the habitable zone with confidence, but would not eliminate the problem of false positives posed by very small or large planets in the habitable zone. A mass measurement could be somewhat useful for ruling out these false positives, but would by no means be definitive. By revisiting a system multiple times to establish its orbit and measuring the mass astrometrically, an observer could still only distinguish EECs from false positives with a typical confidence of $\sim 65\%$ —but could also identify several individual EECs with high confidence ($>90\%$).

The use of color information to discriminate between EECs and false positives could dramatically reduce these ambiguities, but requires that prior assumptions be made about the possible appearance of planets (e.g., as a function of their size and insolation). While current data and models do not allow such prior assumptions, in principle this could change over the next decade. In our example, we find that by revisiting a planet to establish its orbit and measuring its brightness in three bands, the strong majority of EECs could be identified with confidence.

Even though we are not always able to confidently identify EECs using our method, we show that a target prioritization strategy which leverages a Bayesian approach will be more efficient than a nonprobabilistic approach with the same data. Such an approach could reduce the time required for preliminary spectroscopic follow-up by a factor of two.

Our confidence in these results is dependent on our confidence in the priors we have chosen to use, and some of these—such as the radius and semimajor axis distribution of temperate planets around G dwarfs—are based on extrapolation. Many could be improved in the coming decade—such as the relationship between planet radius, mass, and bulk composition through transit and RV observations and advances in planet formation theory, or the reflected spectra of hot and cold sub-Neptunes with *JWST*, WFIRST, and ELTs. Near-term efforts to improve our prior knowledge will enable future observers to more efficiently survey nearby systems in search of potentially habitable worlds.

The authors are grateful to Christopher Stark for providing a sample of host stars from the LUVOIR yield estimates, and Ilaria Pascucci for offering feedback on the manuscript. A.B. acknowledges support from the NASA Earth and Space Science Fellowship Program under grant No. 80NSSC17K0470. The results reported herein benefited from collaborations and/or information exchange within NASA’s Nexus for Exoplanet System Science (NExSS) research coordination network sponsored by NASA’s Science Mission Directorate. This research has made use of NASA’s Astrophysics Data System.

Software: Matplotlib (Hunter 2007), NumPy (Oliphant 2006), and SciPy (Jones et al. 2001).

ORCID iDs

Alex Bixel  <https://orcid.org/0000-0003-2831-1890>
Dániel Apai  <https://orcid.org/0000-0003-3714-5855>

References

- Akeson, R., Armus, L., Bachelet, E., et al. 2019, arXiv:1902.05569
- ALMA Partnership, Brogan, C. L., Pérez, L. M., et al. 2015, *ApJL*, 808, L3
- Anglada-Escudé, G., Amado, P. J., Barnes, J., et al. 2016, *Natur*, 536, 437
- Apai, D., Ciesla, F., Mulders, G. D., et al. 2018, arXiv:1803.08682
- Arney, G., Domagal-Goldman, S. D., Meadows, V. S., et al. 2016, *AsBio*, 16, 873
- Arney, G. N., Meadows, V. S., Domagal-Goldman, S. D., et al. 2017, *ApJ*, 836, 49
- Artigau, E., Bernstein, R. A., Brandt, T., et al. 2018, arXiv:1808.09632
- Ballard, S. 2019, *AJ*, 157, 113
- Barclay, T., Pepper, J., & Quintana, E. V. 2018, *ApJS*, 239, 2
- Barry, R., Kruk, J., Anderson, J., et al. 2011, *Proc. SPIE*, 8151, 81510L
- Batalha, N. E., Smith, A. J. R. W., Lewis, N. K., et al. 2018, *AJ*, 156, 158
- Bixel, A., & Apai, D. 2017, *ApJL*, 836, L31
- Bolmont, E., Libert, A.-S., Leconte, J., & Selsis, F. 2016, *A&A*, 591, A106
- Bouma, L. G., Winn, J. N., Kosiarek, J., & McCullough, P. R. 2017, arXiv:1705.08891
- Broeg, C., Fortier, A., Ehrenreich, D., et al. 2013, *EPJWC*, 47, 03005
- Catling, D. C., Krissansen-Totton, J., Kiang, N. Y., et al. 2018, *AsBio*, 18, 709
- Chen, H., & Rogers, L. A. 2016, *ApJ*, 831, 180
- Chen, J., & Kipping, D. 2017, *ApJ*, 834, 17
- Dressing, C., Stark, C. C., Plavchan, P., & Lopez, E. 2019, *BAAS*, 51, 268
- Ertel, S., Defrère, D., Hinz, P., et al. 2018, *AJ*, 155, 194
- ESA 2017, PLATO Definition Study Report, <https://sci.esa.int/web/plato>
- Fujii, Y., Angerhausen, D., Deitrick, R., et al. 2018, *AsBio*, 18, 739
- Fulton, B. J., & Petigura, E. A. 2018, *AJ*, 156, 264
- Fulton, B. J., Petigura, E. A., Howard, A. W., et al. 2017, *AJ*, 154, 109
- Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2018, *A&A*, 616, A1
- Guimond, C. M., & Cowan, N. B. 2018, *AJ*, 155, 230
- Guimond, C. M., & Cowan, N. B. 2019, *AJ*, 157, 188
- Gupta, A., & Schlichting, H. E. 2019, *MNRAS*, 487, 24
- Hansen, C. J., Kawaler, S. D., & Trimble, V. 2004, *Stellar Interiors: Physical Principles, Structure, and Evolution* (2nd ed.; New York: Springer-Verlag)
- Hu, R., Seager, S., & Yung, Y. L. 2015, *ApJ*, 807, 8
- Hunter, J. D. 2007, *CSE*, 9, 90
- Jones, E., Oliphant, T., Peterson, P., et al. 2001, SciPy: Open Source Scientific Tools for Python, Scipy v1.1.0, <http://www.scipy.org/>
- Kane, S. R., Ciardi, D. R., Gelino, D. M., & von Braun, K. 2012, *MNRAS*, 425, 757
- Kasper, M., Beuzit, J.-L., Verinaud, C., et al. 2010, *Proc. SPIE*, 7735, 77352E
- Kawashima, Y., & Rugheimer, S. 2019, *AJ*, 157, 213
- Kipping, D. M. 2013, *MNRAS*, 434, L51
- Kodama, T., Nitta, A., Genda, H., et al. 2018, *JGRE*, 123, 559
- Kopparapu, R. K., Ramirez, R., Kasting, J. F., et al. 2013, *ApJ*, 765, 131
- Kopparapu, R. K., Ramirez, R. M., SchottelKotte, J., et al. 2014, *ApJL*, 787, L29
- Krissansen-Totton, J., Schwertner, E. W., Charnay, B., et al. 2016, *ApJ*, 817, 31
- Ligi, R., Creevey, O., Mourard, D., et al. 2016, *A&A*, 586, A94
- Lopez, E. D., & Fortney, J. J. 2014, *ApJ*, 792, 1
- Lustig-Yaeger, J., Meadows, V. S., & Lincowski, A. P. 2019, *AJ*, 158, 27
- MacDonald, M. G. 2019, *MNRAS*, 487, 5062
- Madden, J. H., & Kaltenegger, L. 2018, *AsBio*, 18, 1559
- Madhusudhan, N., & Burrows, A. 2012, *ApJ*, 747, 25
- Mallama, A., Krobusek, B., & Pavlov, H. 2017, *Icar*, 282, 19
- Martinez, C. F., Cunha, K., Ghezzi, L., & Smith, V. V. 2019, *ApJ*, 875, 29
- Meadows, V. S., Reinhard, C. T., Arney, G. N., et al. 2018, *AsBio*, 18, 630
- Mulders, G. D., Pascucci, I., Apai, D., & Ciesla, F. J. 2018, *AJ*, 156, 24
- Murray, C. D., & Correia, A. C. M. 2010, in *Exoplanets*, ed. S. Seager (Tucson, AZ: Univ. Arizona Press), 15
- Oliphant, T. E. 2006, *Guide to NumPy* (CreateSpace Independent Publishing Platform: USA)
- Owen, J. E., & Wu, Y. 2013, *ApJ*, 775, 105
- Penny, M. T., Gaudi, B. S., Kerins, E., et al. 2019, *ApJS*, 241, 3
- Plavchan, P., Latham, D., Gaudi, S., et al. 2015, arXiv:1503.01770
- Ricker, G. R., Winn, J. N., Vanderspek, R., et al. 2014, *Proc. SPIE*, 9143, 914320
- Roberge, A., Chen, C. H., Millan-Gabet, R., et al. 2012, *PASP*, 124, 799
- Robinson, T. D., Meadows, V. S., Crisp, D., et al. 2011, *AsBio*, 11, 393
- Rogers, L. A. 2015, *ApJ*, 801, 41
- Savransky, D., & Garrett, D. 2016, *JATIS*, 2, 011006
- Seager, S., Bains, W., & Petkowski, J. J. 2016, *AsBio*, 16, 465

- Sharma, S., Stello, D., Buder, S., et al. 2018, *MNRAS*, **473**, 2004
Stark, C. C., Roberge, A., Mandell, A., & Robinson, T. D. 2014, *ApJ*, **795**, 122
Sullivan, P. W., Winn, J. N., Berta-Thompson, Z. K., et al. 2015, *ApJ*, **809**,
77
The HabEx Study Team 2019, The HabEx Final Report, <https://www.jpl.nasa.gov/habex/>
The LUVOIR Team 2018, arXiv:1809.09668
The LUVOIR Study Team 2019, The LUVOIR Final Report, <https://asd.gsfc.nasa.gov/luvoir/>
Traub, W. A., Breckinridge, J., Greene, T. P., et al. 2016, *JATIS*, **2**, 011020
Van Eylen, V., & Albrecht, S. 2015, *ApJ*, **808**, 126
Weinberger, A., Berkby, J., Bowler, B., et al. 2018, in Giant Magellan Telescope Science Book 2018, ed. R. Bernstein et al. (Pasadena, CA: GMTO Corporation), 31
Weiss, L. M., & Marcy, G. W. 2014, *ApJL*, **783**, L6
Wolfgang, A., Rogers, L. A., & Ford, E. B. 2016, *ApJ*, **825**, 19
Yang, J., Boué, G., Fabrycky, D. C., & Abbot, D. S. 2014, *ApJL*, **787**, L2
Zahnle, K. J., & Catling, D. C. 2017, *ApJ*, **843**, 122