The Astrophysical Journal, 738:170 (12pp), 2011 September 10 doi:10.1088/0004-637X/738/2/170

ON THE LOW FALSE POSITIVE PROBABILITIES OF *KEPLER* PLANET CANDIDATES

Timothy D. Morton[1] and John Asher Johnson[1,2]
[1] Department of Astrophysics, California Institute of Technology, MC 249-17, Pasadena, CA 91125, USA; tdm@astro.caltech.edu
[2] NASA Exoplanet Science Institute (NExScI), California Institute of Technology, Mail Code 100-22, 770 South Wilson Avenue, Pasadena, CA 91125, USA; johnjohn@astro.caltech.edu
Received 2011 February 2; accepted 2011 June 11; published 2011 August 23

ABSTRACT

We present a framework to conservatively estimate the probability that any particular planet-like transit signal observed by the *Kepler* mission is in fact a planet, prior to any ground-based follow-up efforts. We use Monte Carlo methods based on stellar population synthesis and Galactic structure models, and report false positive probabilities (FPPs) for every *Kepler* Object of Interest, assuming a 20% intrinsic occurrence rate of close-in planets in the radius range $0.5\,R_\oplus < R_p < 20\,R_\oplus$. Nearly 90% of the 1235 candidates have FPP < 10%, and over half have FPP < 5%. This probability varies with the magnitude and Galactic latitude of the target star, and with the depth of the transit signal—deeper signals generally have higher FPPs than shallower signals. We establish that a single deep high-resolution image will be an effective follow-up tool for the shallowest (Earth-sized) transits, providing the quickest route toward probabilistically validating the smallest candidates by potentially decreasing the FPP of an Earth-sized transit around a faint star from >10% to <1%. Since *Kepler* has detected many more planetary signals than can be positively confirmed with ground-based follow-up efforts in the near term, these calculations will be crucial to using the ensemble of *Kepler* data to determine population characteristics of planetary systems. We also describe how our analysis complements the *Kepler* team's more detailed BLENDER false positive analysis for planet validation.

Key words: methods: statistical – planets and satellites: general – stars: statistics

Online-only material: color figures, machine-readable table

1. INTRODUCTION

In the wake of the first full release of planet candidates from the *Kepler* mission (Koch et al. 1998; Borucki et al. 2008, 2011), the study of the properties of exoplanetary systems has entered a new era. For the first time there exists a large uniform sample of transiting planets largely unaffected by the detection challenges and selection effects inherent in ground-based searches (e.g., Gaudi 2005; Gaudi et al. 2005), enabling the first clear glimpse of the population of exoplanets down to the size of Earth as well as the first opportunity to study planet radii at large orbital separations. However, follow-up observations to unambiguously confirm individual signals are time-consuming and difficult (or impossible), especially for fainter stars and smaller planets. Consequently, in order to understand what the population of *Kepler* transit-like signals can tell us about the population of exoplanets in general, the problem of astrophysical false positives must be understood.

From the early days of planet transit searches, eclipsing binary systems masquerading as transit signals have plagued detection efforts (Konacki et al. 2003; O'Donovan et al. 2006; Poleski et al. 2010; Almenara et al. 2009). Generally speaking, there are three types of astrophysical false positives: a grazing eclipsing binary, a dwarf star eclipsing a giant star, and a blended eclipsing binary (BB) system, which may be either a hierarchical triple system or an unassociated binary blended within the aperture of a target star (Torres et al. 2004).[3]

The remarkable photometric precision that *Kepler* is delivering (∼30 ppm; Jenkins et al. 2010b) allows for an immediate simplification of the false positive landscape. Batalha et al. (2010b) explain the multitude of ways that certain common false positive scenarios can be identified from *Kepler* photometry alone. For example, grazing eclipsing binaries can be identified by their V-shaped transits, and the giant-eclipsed-by-a-dwarf scenario can be avoided both by the comprehensive work that went into assembling the Kepler Input Catalog (KIC; Latham et al. 2005; Batalha et al. 2010a) and by the ability to photometrically identify giants by their elevated levels of stellar variability compared with dwarf stars (Basri et al. 2010). Even many blended binaries can be identified from the *Kepler* photometry and astrometry alone, by looking for a shift in the center of light, e.g., the "rain diagrams" of Jenkins et al. (2010a). However, some blended binary scenarios remain undetectable by this technique, especially those in hierarchical triple systems, and so a detailed understanding of the false positive problem for *Kepler* requires a detailed understanding of the probability of encountering such blend scenarios.

The *Kepler* team has proven that extremely careful and detailed analyses of individual systems can "validate" planets probabilistically by combining various follow-up observations with modeling the light curves of all possible false positive scenarios with the so-called BLENDER software (Torres et al. 2011). However, this method is computationally expensive and labor-intensive, rendering it a time-consuming process. As a result, only three BLENDER-validated planets having been revealed to date: Kepler-9d (Torres et al. 2011), Kepler-11g (Lissauer et al. 2011), and Kepler-10c (Fressin et al. 2011). With dedicated supercomputer resources coming online for the *Kepler* team's use, this number will certainly rise, but the fact remains that it will be a long time before the BLENDER method can be applied to any large number of the *Kepler* candidates (*Kepler* team 2011, private communication); in the meantime,

[3] In this paper, we consider "false positives" to be purely stellar configurations mimicking transiting planet signals. For discussion of scenarios involving "blended planets," see the Appendix.

statistical interpretations of the candidate sample will rely on statistical assumptions of the false positive rate.

There has been significant previous effort in the literature dedicated to predicting the expected rate of false positive transit signals. Brown (2003) pioneered this work by predicting the rates of different types of false positives and Jovian planet detections for a variety of different surveys, including the then-future *Kepler* mission. Evans & Sackett (2010) greatly extend this work by deriving detection and false positive rates from full-scale bottom-up simulations of synthetic ground-based transit surveys, taking into account all false positive possibilities and many details not included by Brown (2003). We continue in the tradition of these authors with an analysis directly applicable to the *Kepler* mission, approaching from a slightly different angle. Instead of focusing on predicting an overall number or expected rate of planet detections or false positives, we instead seek a simple answer to the following question: "What is a conservative estimate of the probability that an observed apparent transit signal is in fact a true transiting planet?" By framing the issue in this manner, we are able to sidestep the complex issue of detectability, as our analysis assumes a transit-like signal has been detected.

Our philosophy in this work is not to take into account all conceivable details of transit signals, but rather to consider only those which are most salient: the brightness of the *Kepler* target star, its location in the field, and transit signal depth. The details we choose not to address in this work (notably transit period and duration) are those we judge would add uncertainty to our calculations while tending to only decrease our estimates of the false positive probability (FPP). Thus, we are able to keep our analysis straightforward, yet remain confident that we are calculating conservative upper limits to the probability that any given *Kepler* transit signal might be a false positive. As we show in Section 2 and again in Section 3, even these conservative upper limits are enough to indicate that *Kepler* planet candidates will only rarely turn out to be false positives.

2. BASIC BAYESIAN FRAMEWORK

The probability that a given transit signal is of planetary origin may be expressed as the following, according to Bayes' theorem

$$\Pr(\text{planet}|\text{signal}) = \frac{\Pr(\text{signal}|\text{planet})\Pr(\text{planet})}{\Pr(\text{signal})}. \quad (1)$$

In this framework, Pr(signal|planet) is the probability of obtaining the observed signal given that there is a transiting planet on an orbit of a particular period. This factor is known as the *likelihood* of the signal under the planet hypothesis, and we will abbreviate it as $\mathcal{L}_{\rm pl}$. Pr(planet) is the probability of a star hosting a transiting planet (the occurrence rate of planets times the transit probability), which must enter the calculation as an a priori assumption. Thus we call this factor, according to Bayesian convention, the prior on planets and designate it $\pi_{\rm pl}$.

Since there are only two possible origins of a transit-like signal (planet or false positive), the denominator of Equation (1) can be rewritten as marginalizing over the possible models:

$$\Pr(\text{signal}) = \mathcal{L}_{\rm pl}\pi_{\rm pl} + \mathcal{L}_{\rm FP}\pi_{\rm FP}. \quad (2)$$

Using our convention, $\mathcal{L}_{\rm FP}$ and $\pi_{\rm FP}$ are the likelihoods and priors for a false positive signal. The false positive term can be further broken down accounting for the two specific false positive scenarios we are exploring: the BB and the hierarchical eclipsing triple (HT), allowing Equation (1) to be rewritten as the following:

$$\Pr(\text{planet}|\text{signal}) = \frac{\mathcal{L}_{\rm pl}\pi_{\rm pl}}{\mathcal{L}_{\rm pl}\pi_{\rm pl} + \mathcal{L}_{\rm BB}\pi_{\rm BB} + \mathcal{L}_{\rm HT}\pi_{\rm HT}}. \quad (3)$$

In general, the likelihoods depend on the particularities of the transit signal and enable discrimination between models depending on the transit depth, shape, or period. For now we ignore these details, assuming for the moment that we have no knowledge of the differences between the kind of transit signals to expect from planets and from false positives. This enables us to write a simplified version of Equation (3):

$$\Pr(\text{planet}|\text{signal}) \approx \frac{\pi_{\rm pl}}{\pi_{\rm pl} + \pi_{\rm BB} + \pi_{\rm HT}}. \quad (4)$$

We then define the "FPP" as the complement of this probability:

$$\text{FPP} = 1 - \Pr(\text{planet}|\text{signal}). \quad (5)$$

Thus, before considering any detailed information of a particular light curve, the probability that an observed transit signal is actually a false positive depends only on the relative occurrence rates of planets and the false positive scenarios. As mentioned above, $\pi_{\rm pl}$ is simply an assumed occurrence rate of planets times the transit probability; we explain how we determine $\pi_{\rm BB}$ and $\pi_{\rm HT}$ in the following subsections. We describe first this priors-only framework in order to elucidate what dominates our final results, but in Section 3 we will include the likelihoods we removed in Equation (4), taking into account dependence on the depth of the transit signal.

2.1. Blended Binaries

The probability of a transit-mimicking binary system to be blended within the aperture of a *Kepler* target star ($\pi_{\rm BB}$) can be broken down in the following way:

$$\pi_{\rm BB} = \Pr(\text{blend}) \cdot \Pr(\text{appropriate eclipsing binary}). \quad (6)$$

The first factor is the probability for a potentially blending star to be projected within a given radius of a *Kepler* star, and the second is the probability for that star to be an eclipsing binary system that can appropriately mimic a planetary transit.

To calculate these probabilities, we use the stellar population synthesis and Galactic structure code TRILEGAL (TRIdimensional modeL of thE GALaxy; Girardi et al. 2005), which is publicly available on the Web site.[4] TRILEGAL simulates the physical and photometric properties of the stars along a given line of sight, using various stellar evolution grids (Girardi et al. 2002; Chabrier et al. 2000) and a Galactic model that includes a halo, thin and thick disks, and a bulge. All of our simulations use a Chabrier log-normal initial mass function (Chabrier 2001) and default TRILEGAL values for the Galactic structure parameters, including a squared hyperbolic secant structure for the thin disk, an exponential structure for the thick disk, and an oblate spheroid for the halo.

[4] http://stev.oapd.inaf.it/cgi-bin/trilegal

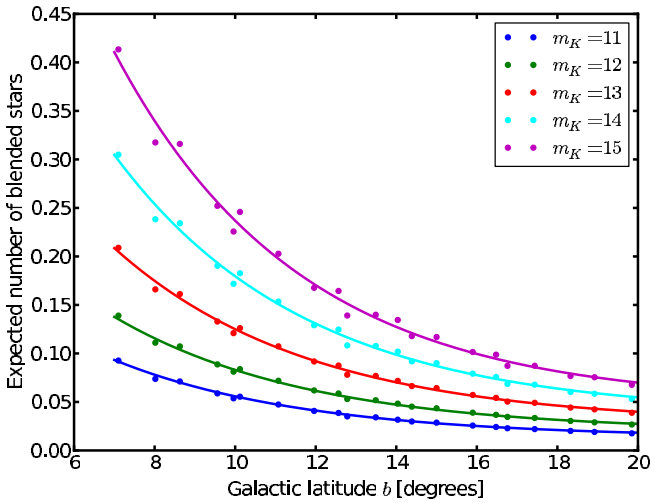

Figure 1. Probability for a possibly blending star to be projected within 2″ of a *Kepler* target star, as a function of Galactic latitude, as determined by TRILEGAL simulations. The plotted points are simulations; the lines are the exponential fits as described in Equation (8).

(A color version of this figure is available in the online journal.)

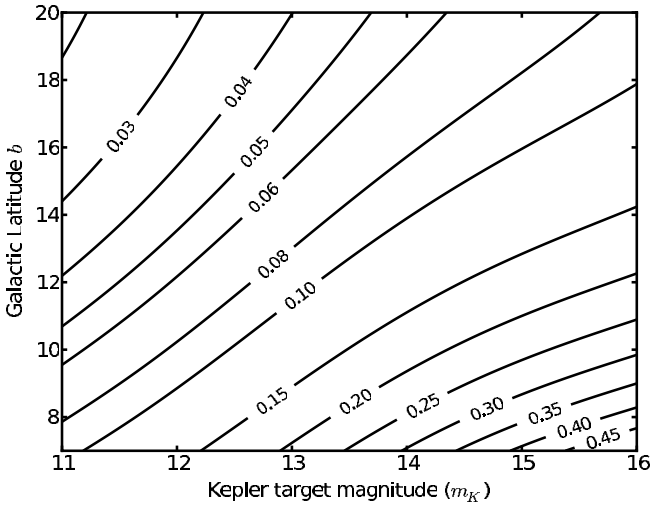

Figure 2. Probability for a possibly blending star to be projected within 2″ of a *Kepler* target star, as a function of both Galactic latitude and target star magnitude, as determined by TRILEGAL simulations.

2.1.1. Probability of a Blend

The blend probability can be calculated by determining the average sky density (e.g., stars per square arcsec) of stars faint enough so as not to be obviously present in *Kepler* data yet bright enough to possibly mimic a transit. The first condition is somewhat subjective, and we conservatively say that a star must be more than 1 mag fainter than the *Kepler* primary in order to be able to hide undetected within the *Kepler* aperture. In practice the true value is probably significantly fainter, but this approximation will lead to only a small overestimate of the blended star probability, as there are many more faint than bright stars.

The faint condition can be determined by noting that in order for a BB system to mimic a transit of fractional depth δ, the blended system must comprise more than a fraction δ of the total flux within the *Kepler* aperture. This condition may be expressed as the following:

$$m_{K,\mathrm{bin}} - m_{K,\mathrm{target}} = \Delta m_K \leqslant -2.5 \log_{10}(\delta), \tag{7}$$

where $m_{K,\mathrm{bin}}$ is the total apparent *Kepler* magnitude of the blended binary system and $m_{K,\mathrm{target}}$ is the magnitude of the *Kepler* target star. A transit depth of $\delta = 0.01$ corresponds to $\Delta m_K = 5$; for $\delta = 10^{-3}$, $\Delta m_K = 7.25$; and for $\delta = 10^{-4}$ (approximately an Earth-sized transit of a solar-radius star), $\Delta m_K = 10$. This means that no binary system fainter than $m_K = 24$ can possibly mimic a $\delta = 10^{-4}$ transit around an $m_K = 14$ star, which is a typical magnitude for a *Kepler* target.

Using TRILEGAL, we determine the sky density of stars in this magnitude range within the *Kepler* field, and thus the probability of one by chance being projected close to a *Kepler* target star, by simulating a 10 deg^2 field centered on the center of the *Kepler* field. We then simply count the stars within the desired range of *Kepler* magnitude (which TRILEGAL provides). As a fiducial example, the average density of stars between $m_K = 15$ and $m_K = 24$ (the range corresponding to a $\delta = 10^{-4}$ transit of an $m_K = 14$ star) is 0.0086 stars arcsec^{-2}. The probability of any given small circle on the sky containing one of these stars is then simply the area of the circle multiplied by this density. Continuing this example ($m_K = 14$, $\delta = 10^{-4}$), the probability of such a star being within 2″ of a *Kepler* target star is 0.11.

However, because the *Kepler* field is quite extended and centered only a few degrees off the Galactic plane, there is a considerable gradient in background stellar density across the field that must be accounted for. To accomplish this, we simulate 21 different 5 deg^2 fields, each centered on one of the *Kepler* double-CCD squares. The resulting probabilities are plotted in Figure 1 as a function of Galactic latitude, for the magnitude ranges corresponding to $m_K = 11, 12, 13, 14$, and 15. Recognizing that this blend probability appears to be exponentially related to Galactic latitude b and that the nature of the exponential depends on m_K, we fit an analytic expression of the following form:

$$p_{\mathrm{blend}}(b, m_K) = C(m_K) + A(m_K) e^{-b/B(m_K)}, \tag{8}$$

where A, B, and C are all polynomial functions of *Kepler* magnitude, with the coefficients listed in Table 1. These fits are valid between m_K values of 11 and 15, and b values between 7° and 20° (the approximate extent of the *Kepler* field). Figure 2 graphically illustrates the behavior of Equation (8).

2.1.2. Probability of an Appropriate Eclipsing Binary

The probability that a blended star is an appropriately configured eclipsing binary system depends first on the binary fraction of blending stars and second on both the distribution of binary properties and the magnitude of the *Kepler* target star. Of central importance is that in order for a blended binary to successfully mimic a *Kepler* planet transit candidate, it must both have a diluted primary eclipse shallow enough to look like a planet and a diluted secondary eclipse either shallow enough so as not to be detected or geometrically aligned so as not to occur.

The apparent fractional "transit" depth of a blended binary system depends on the intrinsic binary system eclipse depth δ_b, and the relative apparent magnitudes of the *Kepler* target star and the blended system:

$$\delta = \delta_b \cdot 10^{-0.4(m_{K,\mathrm{bin}} - m_{K,\mathrm{target}})}. \tag{9}$$

The primary and secondary eclipse depths of the binary system are the following:

$$\delta_{b,\mathrm{pri}} = \frac{\left(\frac{R_2}{R_1}\right)^2 F_1}{F_1 + F_2}, \tag{10}$$

and

$$\delta_{b,\mathrm{sec}} = \frac{F_2}{F_1 + F_2}, \tag{11}$$

where R_1 and F_1 are the stellar radius and flux, respectively, in the *Kepler* band of the larger of the two stars, and R_2 and F_2 are of the smaller star.

If the binary system has an eccentric orbit, then there is a chance that the orbit may be aligned such that only one eclipse occurs. The probability that *both* eclipses occur, given that at least *one* occurs, can be shown to be the following (using Equations (9) and (10) from Winn 2010 and guided by Figure 3 in that work):

$$\Pr(\text{both eclipses}) = \frac{1-e^2}{2\pi}\left[\int_0^{\pi}\frac{1-e\sin\omega}{1-e^2}d\omega + \int_{\pi}^{2\pi}\frac{1-e\sin\omega}{1-e^2}d\omega \cdot \frac{1}{\pi}\int_0^{\pi}\frac{1-e\sin\omega}{1+e\sin\omega}d\omega\right], \tag{12}$$

where e is the eccentricity and ω is the argument of periastron (defined such that the transit occurs when the true anomaly $f = \pi/2 - \omega$). The first term represents the probability that the one given eclipse occurs on the "apastron half" of the orbit (thus guaranteeing the other eclipse will occur); the second term is the probability that the given eclipse occurs on the periastron half, times the fraction of orbital orientations that give the second eclipse as well as the first. One can see that when $e = 0$ this probability becomes 1. When $e \to 1$, the probability $\to 4/\pi^2$, or ~0.405.

The conditions we define for a binary to be "appropriate" are for the diluted primary eclipse depth to be between 0.02 and 10^{-4} (shallow enough to look like a planet, but still detectable), and for the diluted secondary to either be shallower than 10^{-4} (undetectable) or not to occur. We recognize that "detectability" of a transit is a function of more than just the transit depth, but for our purposes we use a depth of 10^{-4} as the detection threshold. A more detailed population study based on *Kepler* candidates should use rather the signal-to-noise ratio (S/N) of a transit as the criterion for detectability (Beatty & Gaudi 2008). However, as our framework deals with how to interpret signals once they are detected, careful detectability analysis is unnecessary.

To calculate the probability of all these conditions being met (a star being binary and being "appropriate"), we use the TRILEGAL simulations and assume binary properties according to the work of Raghavan et al. (2010). That is, we assume a flat mass ratio distribution between 0.1 and 1 (Raghavan et al. 2010 actually observe the distribution to be flat between about 0.2 and 1, but we extend it to 0.1 to be more conservative). We randomly assign eccentricities following the distribution in the Multiple Star Catalog (Tokovinin 1999).

For each star in a particular TRILEGAL line-of-sight simulation that lies in the appropriate magnitude range (Section 2.1.1), we first randomly assign it to be a binary or not and then calculate what the primary and secondary diluted depths would be if the system were eclipsing and blended with a *Kepler* target star of a particular magnitude.

R_1 and F_1 are provided by TRILEGAL,[5] and we determine R_2 and F_2 based on a randomly assigned mass ratio and the Padova models at the age of the primary. Given these system parameters, we can then randomly determine if each system undergoes a non-grazing eclipse, according to the probability that each system will be in such an orientation:

$$\Pr(\text{eclipse}) = \frac{R_1 - R_2}{a}\cdot\frac{1}{1-e^2}, \tag{13}$$

where a is the orbital semimajor axis, determined from Kepler's law, and e is the orbital eccentricity. We then determine whether one or both eclipses occur (according to Equation (12)) and if only one occurs, then we randomly assign whether the lone eclipse is the "primary" or "secondary."

From this procedure, using a *Kepler* target star of $m_K = 14$, an orbital period of 10 days, and a line-of-sight simulation at the center of the *Kepler* field, we find that 1.4% of binaries have non-grazing eclipses and about 27% of those eclipsing binaries are "appropriate." For the binary fraction, we assume that ~40% of stars have binary companions[6] and then consider as potential false positives only the fraction of those binaries that are "short period;" for our purposes, $P < 300$ days, which comes out to about 1/8 of binary systems (again, according to the observed distribution of binary periods from Raghavan et al. 2010). This gives an effective binary fraction of about 5%. This results in a probability of 2.5×10^{-4} for a star to be an appropriate eclipsing binary, giving a value of $\pi_{\mathrm{BB}} = 0.11 \times 2.5 \times 10^{-4} = 2.6 \times 10^{-5}$ for the center of the *Kepler* field.

As in Section 2.1.1, we empirically investigate how this probability changes as a function of galactic latitude and target star magnitude. We find that the behavior for any particular magnitude is well described by a shallow linear relation in b:

$$\Pr(\text{appropriate ecl. binary}) = bD(m_K) + E(m_K), \tag{14}$$

where again the variation of the values of the coefficients D and E is modeled well with a polynomial in m_K (Table 1).

Multiplying Equation (14) with Equation (8) then gives a full analytic expression for the probability of a star of given Kepler magnitude at a given Galactic latitude to be blended with an eclipsing binary system able to mimic a planetary transit:

$$\pi_{\mathrm{BB}}(m_K, b) = [C(m_K) + A(m_K)e^{-b/B(m_K)}] \times [bD(m_K) + E(m_K)], \tag{15}$$

where A, B, C, D, and E are polynomial functions of m_K with coefficients given in Table 1.

Table 1
Polynomial Coefficients[a] for Equation (15)

	c_0	c_1	c_2	c_3	c_4
A	−2.5038e-3	0.12912	−2.4273	19.980	−60.931
B	3.0668e-3	−0.15902	3.0365	−25.320	82.605
C	−1.5465e-5	7.5396e-4	−1.2836e-2	9.6434e-2	−0.27166
D	7.2607e-8	−3.9902e-6	8.1050e-5	−7.1853e-4	2.3345e-3
E	−1.5853e-6	8.6394e-5	−1.7333e-3	1.5099e-2	−4.7598e-2

Note.
[a] This table lists the polynomial coefficients for the empirical fits to how the blended binary false positive probability as a function of Galactic latitude changes with *Kepler* magnitude m_K. A, B, C, D, and E are functions of m_K, valid between $m_K = 11$ and $m_K = 16$. The polynomials are of the form $c_0 + c_1 m_K + c_2 m_K^2 + c_3 m_K^3 + c_4 m_K^4$.

[5] This properly accounts for the possibility that the blend might be an evolved system, e.g., a dwarf star eclipsing a giant.

[6] To be precise, we actually use a binary fraction function that increases with stellar mass: 40% for $M < M_\odot$, 50% for $M_\odot < M < 1.5\,M_\odot$, and 75% for $M > 1.5\,M_\odot$, roughly adapted from Figure 12 in Raghavan et al. (2010). This is a conservative estimate of the binary fraction, as the Raghavan figure includes multiple systems as well as binaries.

2.2. Hierarchical Triples

The probability that a *Kepler* target star is in fact a hierarchical triple system configured such that it might be able to mimic a planetary transit ($\pi_{\rm HT}$) can be broken down as follows:

$$\pi_{\rm HT} = \Pr(\text{triple}) \cdot \Pr(\text{eclipsing and appropriate}). \quad (16)$$

The first factor is simply the frequency of triple systems. Raghavan et al. (2010) determine the frequency of multiple systems to be 12% for Sun-like stars, and the Multiple Star Catalog (Tokovinin 1999) suggests that about half of multiple systems have $P < 300$ days, and so we adopt 6% as our triple fraction. The fraction of triple systems that are of appropriate configuration can be determined by using the same conditions as we used above in Section 2.1.2. That is, we require the diluted eclipse depths (Equations (9)–(11)) to be between 0.02 and 10^{-4} (again accounting for the possibility that an eccentric orbit might provide only a single eclipse), except this time one of the three triple components provides the diluting flux.

We assume two different hierarchical possibilities for triple systems. Referring to the three components in order of descending mass as A, B, and C, the triple system may either be set up as A + BC, where B and C are the closer potentially eclipsing pair and A is the diluting star, or as AC + B, with A and C as the closer pair and B diluting. We ignore the case AB + C because the faintest component being the diluting star would be unable to mimic a planet transit.

We calculate the probability that a triple system will be eclipsing and "appropriate" (again assuming a 10 day orbit) as follows:

$$p_a = \int\int A(M_A, q_1, q_2)\Phi_q dq_1 \Phi_q dq_2. \quad (17)$$

$A(M_A, q_1, q_2)$ equals 1 if the system is eclipsing and can mimic a transit and 0 if not, and the mass ratios $q_1 \equiv M_B/M_A$ and q_2 (either M_C/M_A or M_C/M_B, with 50/50 odds) determine the architecture of the triple system. Φ_q is the mass ratio distribution that we used in Section 2.1.2 (flat between 0.1 and 1). We assign the radius and flux of each component according to the Padova model grids in order to calculate both the non-grazing eclipse probability and the diluted eclipse depths, taking into account the effect of eccentric orbits. Evaluating this integral numerically we obtain $p_a = 0.12$, which results in $\pi_{\rm HT} = 9.8 \times 10^{-5}$.

Unlike the BB scenario, the probability of a target being an HT does not depend either on galactic latitude or apparent magnitude. There is a very weak dependence on stellar mass of the primary, but for our calculations we just assume that all target stars have masses close to 1 $M_\odot$, which is reasonable as *Kepler* is specifically targeting solar-type stars.

2.3. Basic Framework: Summary and Discussion

Now that we have determined the priors for both false positive scenarios, we are able to evaluate the FPP (Equations (4) and (5)) by assuming a frequency of close-in planets. We adopt a 20% frequency according to the results of the *NASA-UC Eta-Earth Survey* of Howard et al. (2010). This conservative estimate of 20% combined with a 5% transit probability for a planet on a 10 day orbit (the period we have been assuming up to now) gives $\pi_{\rm pl} = 0.01$, and thus an FPP of ~0.01. From a planet detection standpoint, this result is quite promising, as it gives a 99% probability that an observed planet-like transit signal around an $m_K = 14$ star in the middle of the *Kepler* field is authentic. Because of the variation of the background stellar density across the field, this value varies with Galactic latitude and m_K, as shown in Figure 3. This is a remarkable result, as it indicates that almost every signal that passes the *Kepler* astrometric and photometric false positive tests is likely a planet transit, before any radial velocity (RV) confirmation attempts.

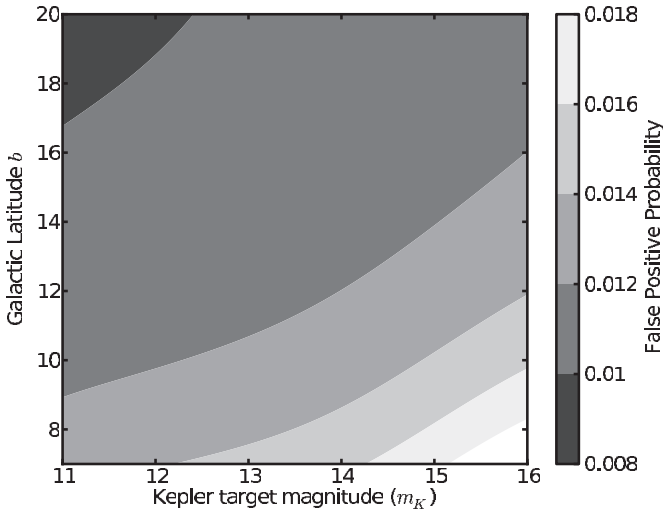

Figure 3. False positive probability of a *Kepler* candidate, according to our basic framework (i.e., independent of δ), as a function of target star magnitude m_K and galactic latitude. A planet occurrence rate of 20% is assumed. This plot assumes that *Kepler* is able to internally restrict the radius inside which a possible blended binary might reside to 2″. There is a small gradient across the field, but the false positive probability is uniformly low.

One might rightly pause at this juncture and wonder how the FPP for *Kepler* can be so low. After all transit searches up until now, both ground-based (e.g., HAT, WASP) and space-based (e.g., *CoRoT*) been plagued by false positives (Konacki et al. 2003; O'Donovan et al. 2006; Poleski et al. 2010; Almenara et al. 2009). To address this, we consider what Equation (4) would say about the probability of a transit signal being true for those experiments.

Taking the Hungarian-made Automated Telescope Network as an example of a ground-based survey, we note that its 11 cm telescopes produce a point-spread function of about 14″ in radius (Hartman et al. 2004), and thus a photometric aperture of ~25″ in radius. Using this radius and a depth of 0.5% as a detection threshold, we repeat the analysis of Section 2.1, using the line-of-sight simulation at the center of the *Kepler* field for the sake of comparison. For the probability of a possibly blending star to be within the aperture, assuming a target star of 12th magnitude, we obtain 1.10, which must obviously now be interpreted as an average number of blending stars per aperture instead of a probability. For the probability of a blending star to be an appropriate eclipsing binary[7] we obtain 9.7×10^{-4}, giving $\pi_{\rm BB} = 1.10 \times 9.7 \times 10^{-4} = 0.0011$. Following Section 2.2 we calculate $\pi_{\rm HT} = 2.1 \times 10^{-4}$. Finally, taking into account that the probability of a Sun-like star hosting a planet easily detectable by this survey is only about 1%,[8] then $\pi_{\rm pl} = 0.01 \times 0.05 = 5 \times 10^{-4}$ for this survey. This results in an FPP of 0.71 for a hot Jupiter-like transit signal for a

[7] For both HAT and *CoRoT*, we relax the assumption that V-shaped transit signals are pre-vetted, so we allow for grazing blended binary orbits.

[8] For $P < 11.5$ days and $M > 0.5\,M_J$; Cumming et al. (2008).

HAT-like ground-based search, according well with Latham et al. (2009), who describe the results of follow-up efforts of a sample of transit candidates, eight of which turned out to be blended binaries and one to be a planet.

The space-based mission *CoRoT* (Baglin 2003) has also had difficulties with false positives. Though it obtains much better photometric precision than a ground-based search and benefits from uninterrupted observing, its large, 320 arcsec^2 aperture (Almenara et al. 2009) results in an expected number of 2.98 blended stars for an $m_K = 14$ target star (according to a simulation of one of the *CoRoT* lines of sight). In addition, its photometric precision is about one part in 10^3, resulting in $\pi_{\rm BB} = 0.0031$, and $\pi_{\rm HT} = 2.9 \times 10^{-4}$. Assuming then a 20% occurrence rate of planets detectable by *CoRoT*, this gives an FPP of 0.25. At first this appears to somewhat contradict Almenara et al. (2009), who reported 6 planets and 25 diluted binaries among *CoRoT*'s "solved candidates" (ignoring the "undiluted binary" category, as we are not considering that possibility for *Kepler*). However, if one considers how much easier (and faster) it is to identify a false positive than to positively confirm a planet, this prediction can certainly be consistent with these results, as only 49 of their 122 candidates had been solved at the time. In fact, a prediction of our methods is that many of the unsolved *CoRoT* candidates are indeed planets.

Another reasonable question to ask is how uncertainties in our models and assumptions propagate through to uncertainties in FPP. This is challenging to address exactly, as our analysis rests on the results from TRILEGAL simulations, stellar model grids, and various assumptions about multiple star systems. Rather than attempt a detailed start-to-finish treatment of all the uncertainties, we instead investigate what happens if we artificially inject fractional uncertainties into our prior calculations and simulate the results according to our analytic fits. We find that 20% fractional uncertainties in background stellar density, appropriate eclipse probability, and HT probability lead to 17% fractional uncertainty in FPP. This is a fiducial example, and the uncertainty in FPP scales linearly with these component uncertainties.

One might also wonder how sensitive our derived FPP for *Kepler* is to the assumption that 20% of stars host planets, as well as how justifiable such an assumption may be. We address these questions in Figure 4. A 20% occurrence rate lies in the middle of the measured occurrence rate of planets with minimum masses $>3\,M_\oplus$ and periods <50 days from the *NASA-UC Eta-Earth Survey* of Howard et al. (2010). In addition, even the most pessimistic interpretation of the results from $\eta_{\rm earth}$ allows for a minimum of a 9% occurrence rate, which would still imply an FPP of only 3%. More likely, the true occurrence rate is somewhat higher than our assumption, if not as high as the ∼40% implied by a naïve extrapolation of the observed power-law-like distribution down to 0.5 $M_\oplus$. We note that the *NASA-UC Eta-Earth Survey*, as with all RV surveys, is only able to measure minimum masses and thus that the interpretation of the true mass of any individual detection is dependent on an assumption of the overall form of the planet mass function (Ho & Turner 2010). However, when an ensemble of minimum mass measurements is available and its distribution resembles a power law with index $\alpha < -1$, the most likely explanation is that the true mass function follows a similar power-law shape.

In summary, we may say that several factors contribute to *Kepler* being able to minimize the false positive problem compared to previous transit surveys. First, its ability to astrometrically rule out wide blend scenarios helps mitigate the issue of blended binaries. Second, its photometric precision enables it to identify many false positives based on their secondary eclipses. And lastly, *Kepler* is sensitive to lower-mass planets, which are significantly more common than the larger planets to which ground-based surveys are sensitive.

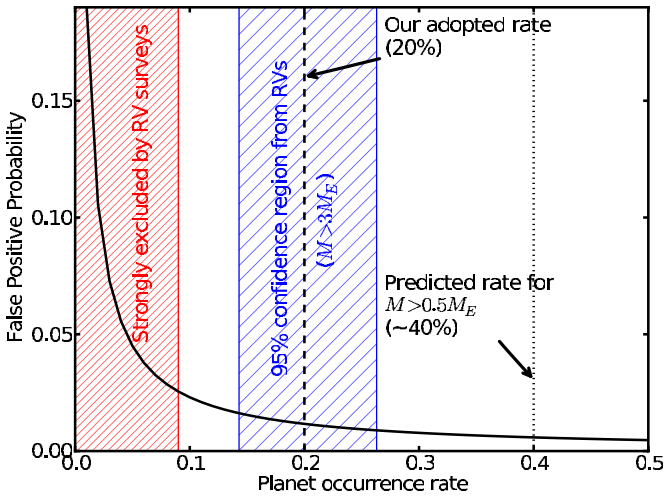

Figure 4. False positive probability as a function of assumed planet occurrence rate, for an $m_K = 14$ target star in the center of the *Kepler* field. The occurrence rate of planets detectable by *Kepler* is not known for sure, but RV surveys, especially the *NASA-UC Eta-Earth Survey* of Howard et al. (2010), have made inroads in measuring the fraction of stars hosting low-mass planets. The hashed area below 9% represents the occurrence rate of planets with $P < 50$ days that is ruled out with 95% confidence by $\eta_{\rm earth}$, counting only the firm detections, and not correcting for completeness. The central hashed area represents the 95% confidence region calculated including candidate planets and completeness correction, for minimum masses greater than 3 $M_\oplus$. Extrapolating their observed mass distribution down to 0.5 $M_\oplus$ brings their total estimated planet occurrence rate to 43%. Overall, this plot shows that our derived FPP cannot reasonably be any higher than 3% if our planet occurrence estimate is incorrect and will likely be lower.

(A color version of this figure is available in the online journal.)

3. DETAILED FRAMEWORK: CONSIDERING TRANSIT DEPTH

We note that we have not yet discussed any details of the transit signal besides its existence, though some of these details may be important. For example, one might expect false positive blended binaries to be more common at shallower depths (since faint stars are more common than bright stars and thus more likely to be blended), which might make the BB scenario more of a problem for Earth-sized transit signals. We have also assumed that planets and eclipsing binaries have the same eclipse probability (allowing us to cancel the likelihood factors in Equation (3)), though this is not exactly true either, as both the orbital separations of the systems and the radii of the objects are different for a given fixed period. The eccentricity distribution of binaries is a function of period—longer period systems have larger eccentricities and thus false positive scenarios might be more likely to show only a single eclipse. And finally, for fainter stars and shallower eclipses, it may be more difficult for internal *Kepler* procedures to astrometrically identify blends.

With these concerns in mind, we may pursue a more detailed analysis of any particular transit. There are many features of transit light curves that might all be used in this exercise, but for now we only take into account the depth of the signal, as that is the most easily measured and easily understood quantity. In

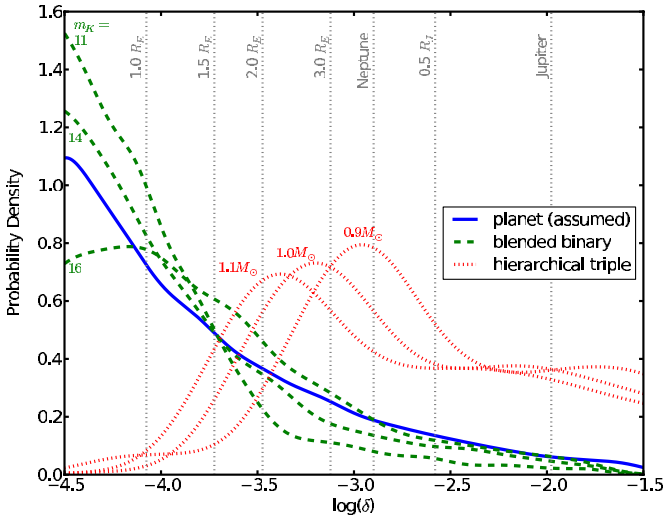

Figure 5. Distributions of apparent "transit" depths δ for different scenarios. The blended binary and hierarchical triple distributions are based on TRILEGAL simulations with the binary distribution assumptions discussed in Section 2. Examples of δ distributions are given for different target star properties, showing how the blended binary scenario depends on target star apparent magnitude and how the hierarchical triple distribution depends on intrinsic target star mass. The planet distribution comes from an assumption of a continuous power law in planet radius $dN/dR_p \propto R_p^{-2}$, including random statistical dilution by binary companions. Note how blended binaries become less significant for deep signals and how eclipsing triples become insignificant for the shallowest signals. The deep end of the false positive δ distributions is mostly due to eccentric binaries that are oriented such that only a single eclipse occurs.

(A color version of this figure is available in the online journal.)

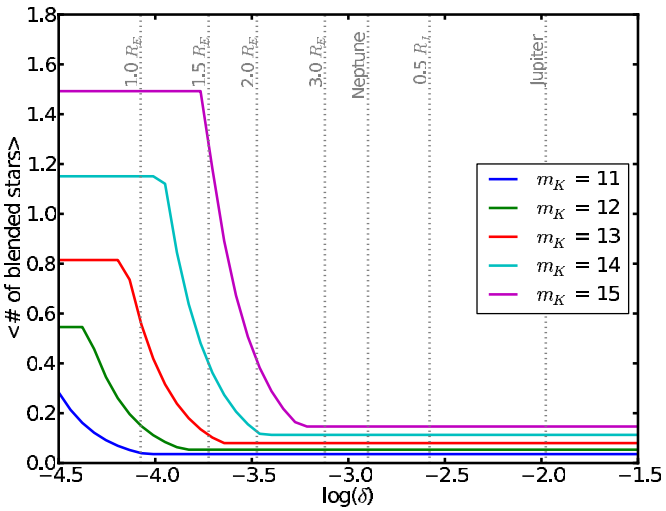

Figure 6. As stars get fainter and transit signals get shallower, the ability for *Kepler* to observe a centroid shift indicative of a displaced blended eclipsing binary decreases. We parameterize this effect according to Equation (19). The plateau toward shallow depths is a result of the maximum blending area for this example being set to an aperture of eight *Kepler* pixels; the location of this plateau for any particular target will depend on its aperture size. This plot is made according to a galactic latitude in the middle of the *Kepler* field; other latitudes will scale appropriately according to the varying stellar density. The planet radii are marked assuming a solar-radius star.

(A color version of this figure is available in the online journal.)

this case, Equation (3) becomes

$$\Pr(\mathrm{pl}|\delta) = \frac{\mathcal{L}_{\mathrm{pl}}(\delta)\pi_{\mathrm{pl}}}{\mathcal{L}_{\mathrm{pl}}(\delta)\pi_{\mathrm{pl}} + \mathcal{L}_{\mathrm{BB}}(\delta)\pi_{\mathrm{BB}} + \mathcal{L}_{\mathrm{HT}}(\delta)\pi_{\mathrm{HT}}}. \quad (18)$$

Here the likelihood functions provide a means to quantify the extent to which the conclusions of our simple framework may change as a function of transit depth δ.

Figure 5 shows the likelihoods that we estimate for the three different scenarios as a function of depth. The distribution of depths for the blended binary and hierarchical triple scenarios are determined from the same calculations that we used to compute the priors, except rather than just counting all the systems that give depths that are both planetary and detectable, we keep track of the depth of each simulated false positive and build up δ distributions.

We calculate the δ distribution for planets assuming a simple continuous power-law distribution of planet radii ($dN/dR_p \propto R_p^{-2}$) between 0.5 and 20 $R_\oplus$, and setting $\delta = (R_p/R_\star)^2$. While a more sophisticated treatment might involve adopting a planet mass distribution according to RV surveys and theoretical mass–radius relations (e.g., Fortney et al. 2007; Seager et al. 2007), the number of assumptions required for these models and the fact that they do not generally include significant atmospheres for super-Earth-type planets suggest that such efforts are not warranted; for example, Rogers et al. (2011) show that Neptune-sized planets can have a wide range of possible masses. In addition, the current uncertainties in stellar radius of the *Kepler* candidate host stars further blur the mapping from δ to R_p. Thus the main role of the δ distribution we adopt for planets is to encapsulate the assumption that smaller planets are more common than large ones, which is consistent with RV surveys (Howard et al. 2010).

Another consideration that should vary with δ is the ability of *Kepler* to astrometrically identify displaced blends. In Section 2, we assumed a radius of 2″ inside which a blend might reside. However, this radius should increase as transits get shallower and stars get fainter, which should cause the signal to noise of the centroid shift signal to decrease. This is a question that the *Kepler* team should be able to address using simulations of its offset-detecting procedures, but for our purposes we use the radius that the *Kepler* team obtained for Kepler 10b (1.″17; Batalha et al. 2011) and assume scaling with δ and m_K as follows:

$$r = 1.''17\sqrt{10^{-0.4(11-m_K)}}\left(\frac{\delta}{1.5\times10^{-4}}\right)^{-1}, \quad (19)$$

with 11 being the m_K value for Kepler 10. To be conservative we set the minimum r to be 2″ if this expression gives a smaller value. On the high end, we cap the radius at 6.″4, corresponding an area equivalent to eight *Kepler* pixels, a typical aperture size (though for any particular target this will vary). The square root factor accounts for a diminishing number of photons received as the target star gets fainter, and the inverse relationship with delta is because the centroid shift scales as δ: $\Delta C \sim \delta \cdot r$. Figure 6 illustrates this effect; bright stars and deeper transits give Pr(blend) as determined in Section 2.1.1, but as the target star gets fainter and the signal shallower, the expected number of possibly blending stars begins to increase substantially, up to the point at which our calculated blend radius exceeds the maximum assumed 8 pixel aperture area.

4. RESULTS

4.1. General

The adoption of these more detailed considerations enables us to estimate the FPP as a function of δ for a star of given apparent *Kepler* magnitude, Galactic latitude, stellar radius, stellar mass, and aperture size. This is illustrated in Figure 7 for fiducial

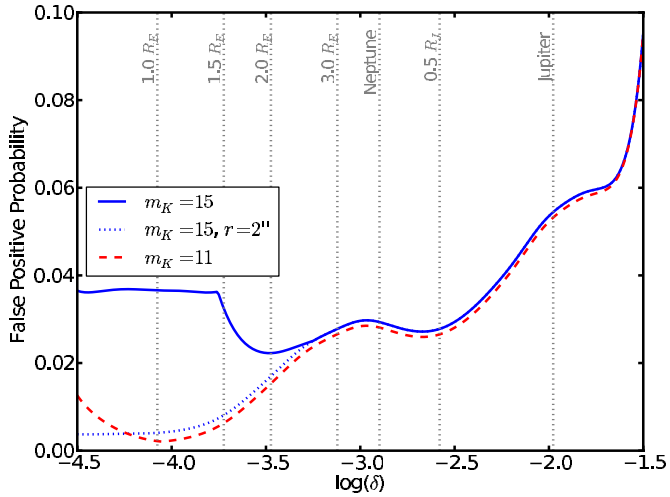

Figure 7. Probability that a 10 day signal of a given depth will be a false positive, shown for both an $m_K = 11$ and an $m_K = 15$ star with solar properties, Galactic latitude in the center of the *Kepler* field, and an 8 pixel *Kepler* aperture. An overall planet occurrence rate of 20% and a planet radius function $dN/dR \propto R^{-2}$ are assumed. The false positive probability increases toward deeper signals mostly because of the decreasing planet radius function combined with the significant tails of the false positive δ distributions. For the fainter star the false positive probability begins to pick up again at the shallowest depths as it becomes more difficult for *Kepler* to rule out displaced blended binaries via astrometry. The plateau corresponds to the point at which the blending radius becomes equal to the aperture size. The dotted line represents the effect of restricting the blending radius to 2″ with a single high-resolution image: this can decrease the false positive probability for Earth-sized signals from ∼4% to <1%.

(A color version of this figure is available in the online journal.)

examples of a signal with a 10 day period around $m_K = 15$ and $m_K = 11$ Sun-like stars in the middle of the *Kepler* field, both assumed to have 8 pixel apertures. We first note that over the whole range of δ for these examples, the FPP remains <0.10, indicating that these additional considerations do not significantly change the qualitative conclusions we reached within the simple framework. The majority of transit signals in the *Kepler* data release will be actual planets.

We note that FPP generally tends to *increase* with increasing signal depth. This may be understood by considering that the planet radius function decreases toward larger signals while the hierarchical triple false positive δ distribution remains relatively flat (see Figure 5). Toward shallow signals FPP begins to increase again; this is due to the growing influence of the blended binary scenario as the radius inside which *Kepler* is able to rule out displaced blends decreases. The plateau at the shallowest depths is a result of the "blending radius" becoming equal to the size of the aperture; the maximum value it can attain. If this blending radius is able to be decreased to 2″ by a deep high-resolution image, that reduces the FPP at shallow depths to below 1%.

We note that the plots in Figures 5–7 are only for particular chosen values of magnitude and a single Galactic latitude in the middle of the *Kepler* field, as well as for particular choices of stellar properties. We present a more comprehensive illustration of the FPP manifold in Figure 8, choosing three specific values of δ to illustrate how the FPPs for different types of signals vary with target star magnitude and Galactic latitude. We fix the target star to have solar properties in these examples.

Earth-sized transits show a steep gradient across the field and toward fainter stars; this is a result of increasing contribution to the FPP from blended binaries (see Figure 2), combined with the increased blend radius for a shallow transit (Figure 6). This gradient is shallower for a Neptune-sized signal and almost disappears for a Jupiter-sized signal, because of the growing contribution of the hierarchical triple scenario and decreasing influence of chance-alignment blended binaries. These plots also illustrate the potential power of deep high-resolution imaging follow-up observations. If such an image is taken and no companion is found outside a radius of a few arcseconds, then that dramatically reduces the FPP for shallow signals, as illustrated with the dotted contours.

4.2. Application to Kepler Candidates

We apply the framework discussed above to calculate the FPP for every *Kepler* Object of Interest (KOI) published in Borucki et al. (2011); these results are summarized in Table 2. For each KOI we generate individualized δ distributions for the different false positive scenarios using the relevant *Kepler* magnitude, Galactic latitude, and stellar parameters from the KIC. We also match the eccentricity distribution of false positive scenarios to the period of the signal (according to

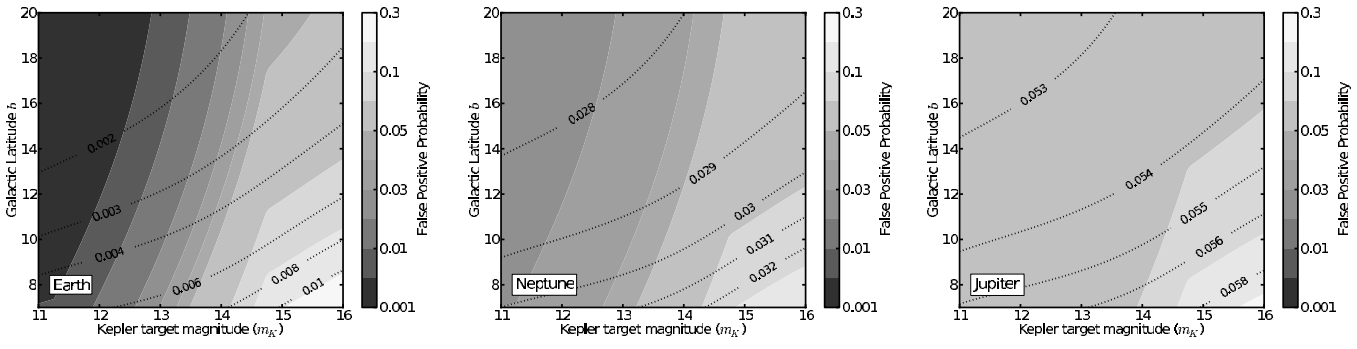

Figure 8. These plots illustrate the behavior of *Kepler* false positive probability (FPP) as a function of target star magnitude (m_K) and Galactic latitude, for three particular choices of transit depth δ, all plotted with the same color scale. A planet occurrence rate of 20% is assumed, and the target star is fixed to have solar mass, solar radius, and a photometric aperture of 8 pixels. These plots are similar to Figure 3 except for they take into account both the changing blend radius as a function of m_K and δ (Equation (19)) and the relative likelihoods of false positives and planets at the chosen values of δ. All three δ values show increasing FPP toward fainter target stars and lower galactic latitudes, though the strength of the gradient decreases for the deeper signals, as the relative importance of the hierarchical triple scenario increases. Dotted lines show the FPP contours if the blend radius were restricted to 2″, illustrating the power of a single deep high-resolution observation. Note that only in the Earth-sized transit case does the high-resolution observation result in FPP <0.01, as the FPP for shallower signals is dominated by chance-alignment blended binaries while the FPP for deeper signals is dominated by hierarchical triple scenarios.

Table 2
False Positive Probabilities for *Kepler* Planet Candidates[a]

KOI (1)	δ (2)	Threshold (3)	m_K (4)	No. of Pixels (5)	b (6)	P (7)	$M_\star$ (8)	$R_\star$ (9)	$L_{\rm BB}$ (10)	$L_{\rm HT}$ (11)	$L_{\rm pl}$ (12)	FPP (13)
370.01	3.21e−04	2.09e−05	11.93	16	13.04	42.9	1.30	2.29	7.83e−06	2.74e−05	0.00145	0.02
371.01	1.11e−03	5.73e−05	12.19	19	5.94	278.0	1.33	3.01	3.04e−06	6.37e−06	0.000246	0.04
372.01	7.64e−03	4.98e−04	12.39	12	6.82	125.6	1.05	0.95	1.5e−06	1.26e−05	0.000144	0.09
373.01	5.97e−04	5.12e−05	12.77	8	11.79	135.2	1.11	1.30	5.52e−06	1.42e−05	0.000506	0.04
374.01	5.95e−04	4.06e−05	12.21	13	13.68	172.7	1.11	1.26	2.57e−06	1.07e−05	0.000411	0.03
375.01	4.70e−03	1.16e−04	13.29	13	15.91	220.0	1.07	1.04	3.23e−07	6.79e−06	0.000137	0.05
377.01	6.94e−03	1.61e−04	13.80	6	14.49	19.3	1.00	0.68	1.25e−06	2.24e−05	0.000497	0.04
377.02	6.24e−03	2.26e−04	13.80	6	14.49	38.9	1.00	0.68	1.55e−06	2.09e−05	0.000337	0.06
377.03	2.25e−04	3.07e−05	13.80	6	14.49	1.6	1.00	0.68	0.000344	0.0001	0.0154	0.03
379.01	2.51e−04	3.14e−05	13.32	10	9.61	6.7	1.19	1.59	0.000125	7.73e−05	0.00593	0.03
384.01	1.76e−04	1.76e−05	13.28	8	8.46	5.1	1.09	1.22	0.000286	5.19e−05	0.00823	0.04
385.01	2.69e−04	4.25e−05	13.44	5	9.85	13.1	1.04	1.04	9.62e−05	4.11e−05	0.00352	0.04
386.01	8.45e−04	5.51e−05	13.84	5	8.61	31.2	1.11	1.12	3.2e−05	3.29e−05	0.00116	0.05
386.02	6.60e−04	8.25e−05	13.84	5	8.61	76.7	1.11	1.12	2.87e−05	2.03e−05	0.000686	0.07
387.01	9.41e−04	8.82e−05	13.58	9	13.50	13.9	0.69	0.74	1.89e−05	8.79e−06	0.0018	0.01

Notes.
[a] All 1235 candidates are listed in the full version of the table, available at exoplanets.org/data/KOIFPPtable.txt and in the online version of the journal. The table columns are described below.
(1) KOI identifier, from Borucki et al. (2011).
(2) Transit depth.
(3) Detection threshold, chosen to be S/N = 3, according to Borucki et al. (2011).
(4) *Kepler* magnitude.
(5) Size, in *Kepler* pixels (4″ square each) of the photometric aperture, according to the publicly available pixel data.
(6) Galactic latitude of target star, in degrees.
(7) Period of candidate, in days.
(8) Stellar mass, according to the Kepler Input Catalog (KIC).
(9) Stellar radius, according to the KIC.
(10) Likelihood × prior for the blended binary scenario.
(11) Likelihood × prior for the eclipsing hierarchical triple scenario.
(12) Likelihood × prior for the transiting planet.
(13) False positive probability = $1 - L_{\rm pl}/(L_{\rm pl} + L_{\rm BB} + L_{\rm HT})$.

(This table is available in its entirety in machine-readable form in the online journal. A portion is shown here for guidance regarding its form and content.)

the multiple star period–eccentricity distribution as illustrated in Tokovinin 1999) and use the actual detectability threshold for each KOI (based on calculating what depth corresponds to S/N = 3 according to the data in Borucki et al. 2011) to count "appropriate" eclipsing stellar systems, instead of the generic 10^{-4} value. We then calculate the FPP using the actual area of the photometric aperture, as determined from the publicly available pixel data for each KOI and the transit depth as given in Borucki et al. (2011). The distribution of FPPs is illustrated in Figure 9.

In Table 2 we list the KOI parameters relevant to the FPP calculation, the calculated FPPs, and the values of the intermediate factors in the calculation, which we summarize as $L_{\rm pl}$, $L_{\rm BB}$, and $L_{\rm HT}$, where

$$L_{\rm pl} = \mathcal{L}_{\rm pl}(\delta)\pi_{\rm pl} = f_{\rm pl} \cdot \Pr(\text{Transit}) \cdot \Phi_{\rm pl}(\log\delta), \qquad (20)$$

where $f_{\rm pl}$ is the overall planet occurrence frequency, Pr(Transit) is the geometric transit probability, and $\Phi_{\rm pl} = dN/d\log\delta$ is the probability density function for $\log\delta$. Thus

$$\text{FPP} = 1 - \frac{\rm L_{pl}}{\rm L_{pl} + L_{BB} + L_{HT}}, \qquad (21)$$

where $L_{\rm BB}$ and $L_{\rm HT}$ are the corresponding terms for the two false positive scenarios.

We list these individual components in the table primarily because the FPP calculation fundamentally depends on assumptions of the planetary occurrence rate and radius distribution,

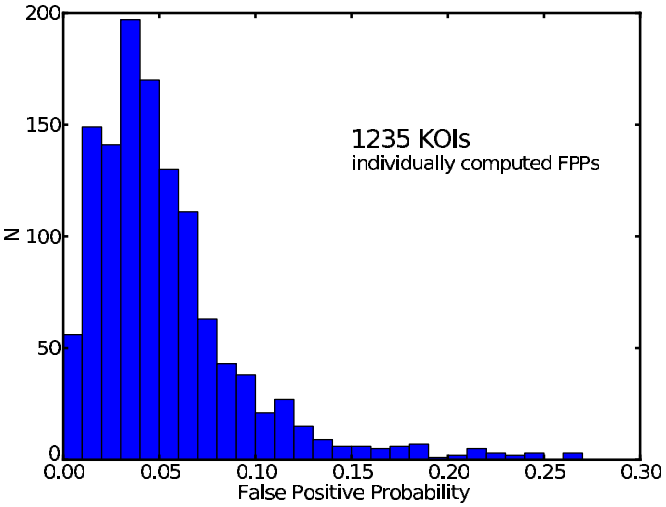

Figure 9. Distribution of false positive probabilities (FPPs) among the 1235 *Kepler* planet candidates announced in Borucki et al. (2011). FPP for each candidate is calculated individually, taking into account the apparent *Kepler* magnitude, Galactic latitude, mass and radius of the host star, the depth and period of the signal, the number of pixels contained the optimal aperture used for *Kepler* photometry, and the detection threshold of each KOI time series. Nearly 90% (1098) have FPPs less than 10%, and over half (713) have FPPs less than 5%. The mean FPP of the sample is about 6%, indicating that we expect there to be fewer than ∼75 false positives among the candidate sample. An important caveat here is that these calculations assume that all candidates have passed the false positive-vetting tests that are possible using *Kepler* photometry.

(A color version of this figure is available in the online journal.)

and different assumptions will result in different FPPs. Though we show in Figure 4 that these assumptions are unlikely to dramatically affect the final FPP numbers, one could in principle calculate $L_{\rm pl}$ based on different assumptions and recalculate FPP, given all the components.

5. DISCUSSION: RELATIONSHIP TO "BLENDER"

The FPP analysis we present in this paper is not the first false positive analysis that has been done regarding *Kepler* candidates. In fact, the *Kepler* team has statistically "validated" three planets: Kepler-9d (Torres et al. 2011), Kepler-11g (Lissauer et al. 2011), and Kepler-10c (Fressin et al. 2011) by demonstrating that the chance of any of those signals being due to a false positive is low enough to consider the candidate a *bona fide* planet. This has been done using the procedure the *Kepler* team has named BLENDER (Torres et al. 2004, 2011).

BLENDER attempts to directly model the candidate light curve using every conceivable false positive scenario, informed by high-resolution imaging follow-up observations. The goodness of fit of the false positive models is then compared to the best-fit planetary model. The false positive scenarios that cannot fit the light curve as well as a transiting planet model are rejected. The a priori likelihood of the remaining scenarios (those false positive scenarios that provide comparable-quality fits to the light curve) is then assessed relative to the likelihood of a *bona fide* transiting planet, and if the planetary explanation is much more likely, then the planet is considered validated.

As can be inferred from the fact that the *Kepler* team has published only three validated planets to date out of over 1200 planet candidates that have been made public, BLENDER is a very time-consuming procedure, being both computationally expensive and labor-intensive. Relying on extensive modeling of individual light curves and requiring a suite of follow-up observations to be most effective, it can only be applied to single KOIs on an individual basis.

If BLENDER may be characterized as a "deep and narrow" false positive analysis tool, the FPP analysis we present in this paper might be described as "shallow and wide." It takes only 15 s per candidate[9] for us to generate the δ distributions required to calculate the individualized FPP numbers listed in Table 2, which makes our analysis easily and immediately applicable to all the KOIs, whereas BLENDER takes weeks of computation and analysis per candidate. On the other hand, BLENDER takes into account all possible information about each KOI (detailed light curve shape, adaptive optics (AO) imaging, multiwavelength transit information, etc.), whereas we only consider the depth and period of the transit signal and the properties of the target star.

Another way to think of the relationship between our FPP analysis and BLENDER is that if BLENDER is an N-step procedure, our analysis is step N. We ignore most of the detail of the light curve and make no use of any follow-up observations, but go straight to the a priori likelihood calculation and do that step as carefully as possible. What is remarkably encouraging for the *Kepler* mission is that even this "shallow," single-step analysis is enough to determine that the FPP for almost every KOI is less than 10%, and for over half the KOIs is less than 5%.

If our analysis is step N of the BLENDER process, how would the first $N-1$ steps be incorporated into the present analysis to improve upon the FPPs published here? First, consider that if $x = L_{\rm pl}$ and $y = L_{\rm BB} + L_{\rm HT}$, then the probability $p_{\rm pl}$ that a signal is a planet is the following:

$$p_{\rm pl} = \frac{x}{x+y}. \tag{22}$$

This may be rewritten as

$$p_{\rm pl} = \frac{1}{1+y/x}. \tag{23}$$

If $y \ll x$ (as we have shown it typically is) then

$$p_{\rm pl} \approx 1 - \frac{y}{x}, \tag{24}$$

or

$$\mathrm{FPP} \approx \frac{y}{x}. \tag{25}$$

The typical role of BLENDER in this context can then be thought of as multiplying y by a factor we call $f_{\rm BLENDER}$ ($0 < f_{\rm BLENDER} < 1$) that represents the fraction of the potential false positive scenarios (weighted by their intrinsic likelihoods) that produce acceptable fits to the light curve. Thus if BLENDER were to rule out 90% of the false positive scenarios considered in our analysis ($f_{\rm BLENDER} = 0.1$) for a particular system, then that would decrease the FPP for that system by a factor of 10—such analysis would be enough to make FPP < 0.01 for almost every KOI.

In a similar spirit, for those KOIs whose FPP is dominated by the blended binary scenarios (mostly the shallowest signals), y can also be significantly decreased simply if deep high-resolution imaging shows no potentially blending companions. This effect is illustrated in Figure 8, where dotted FPP contours are drawn illustrating the effect of restricting the "blend radius" to $2''$. Decreasing this even further to $1''$ or smaller would give another factor of four or more reduction in FPP.

In some cases of course, follow-up imaging observations will identify the presence of nearby stars within the "blend radius" inside of which astrometric offset methods were previously unable to identify displaced blends. In these cases, the analysis presented in this paper must be superceded by a more specifically tailored analysis such as BLENDER. In general, a detected nearby blend will cause the preliminary FPP to substantially increase, as the Pr(Blend) factor that we found to be of order $\sim$0.10 (Section 2.1.1) is then divided out from the $L_{\rm BB}$ term, making it more comparable to the $L_{\rm pl}$ term. In these cases a full suite of follow-up observations and the more detailed approach that BLENDER uses will become necessary to validate candidates.

6. CAVEATS AND CONCLUSIONS

We present both a framework to analyze the a priori FPP of *Kepler* planet candidates and preliminary FPPs for the entire sample of 1235 released candidates, finding that FPP $<$ 10% for almost all the KOIs and $<$5% for over half the KOIs. The philosophy we adopt in this work is to calculate conservative upper limits to these FPPs; further analysis may well demonstrate them to be lower, but we do not expect them to be higher. Our analysis indicates that fewer than $\sim$75 are likely to turn out to be false positives.

However, the details of these conclusions are based on several assumptions (apart from those regarding the accuracy of the TRILEGAL simulations and the properties of binary and triple systems described in Section 2) that come with some caveats.

[9] Non-parallelized computation on an iMac with a 2.66 GHz quad-core i5 processor.

1. We assume that all candidates have passed all preliminary false-positive-vetting procedures that are possible using *Kepler* photometry and astrometry alone. In particular, we assume that the transits are not obviously V-shaped, there is no detectable secondary eclipse, and that careful centroid analysis has not revealed the presence of a displaced blended binary. If photometry or astrometry for a candidate actually does turn out to indicate a possible false positive, then the FPPs calculated in this paper for that KOI are not accurate for that system.
2. We assume host star stellar parameters according to the KIC. If stellar radii or stellar types are found to be significantly different from the KIC estimates, then that could change the interpretation of transit signals (e.g., turning a Jupiter-sized planet into an M dwarf).
3. We assume a planet radius function that increases toward smaller planets. There are many reasons, both theoretical and observational, to assume this is correct, but if it is not, then the false positive numbers for the smallest candidates would be a factor of two or so higher.

We emphasize that in this analysis the only information about the transit that we have considered in detail is the depth. Certainly more information could be used, if one is willing to make more assumptions. For example, one might include a more detailed period dependence in the likelihood function by comparing the period distribution of known binary and triple systems to an assumed period distribution of planets, the way we have done with δ. However, we do not believe that the true period distribution of exoplanets is known well enough yet for such an analysis to be useful.

We also emphasize that the intention of this paper is not to encourage other analyses to completely ignore the possibility that some *Kepler* candidates might be false positives. Rather, we suggest that in statistical analyses using the ensemble of KOIs to investigate the distribution of planet properties, the FPPs in this paper (or based on the calculations in this paper; e.g., with different assumptions of the planet occurrence rate or radius function) be used to count "fractional planets;" i.e., for a KOI with FPP $= 0.05$ to count as 95% of a planet.

Finally, we provide several suggestions to guide and optimize *Kepler* follow-up efforts, based on the results of our analysis.

1. For the shallowest candidates, or those for which a blended binary is the most likely false positive scenario, we recommend deep high-resolution imaging (with a target contrast ratio corresponding to the depth of the signal: $\Delta m_K = -2.5 \log \delta$), as excluding the presence of potentially blending stars at close separation will be the quickest path toward validation of such systems. Contrast ratios up to 10 mag as close as $1''$ have long proven to be technically feasible with existing AO instruments (e.g., Luhman & Jayawardhana 2002; Biller 2007).
2. For deeper candidate signals for which an HT is the most likely false positive scenario we recommend follow-up efforts targeted toward the identification of physically bound companions to the KOI. High-resolution imaging is one useful tool here (though not necessarily as deep as those observations targeting projected binaries) to target wide-separation physically bound companions, but high S/N spectroscopy (both optical and infrared) may be even more important, in order to spectroscopically identify or constrain the presence of low-mass stellar companions.
3. For all the candidates we recommend spectroscopic follow-up to improve our knowledge of the physical parameters of the candidate host stars, in order to rule out the possibility of an eclipsing binary being misclassified as a transiting planet due to an incorrect assumed radius.

In summary, the exquisite photometric and astrometric precision of the *Kepler* instrument enables many of the false positives that have traditionally plagued transit surveys to be identified prior to follow-up observations. The result is that the majority of the candidates announced by Borucki et al. (2011) are likely to be bona fide planets. Thus, having surveyed the landscape of false positives in the *Kepler* field, we conclude that the outlook is bright for statistical analyses of exoplanet occurrence and properties based on the data made public by the *Kepler* team.

We gratefully acknowledge Ed Turner and Scott Gaudi for their thoughtful and detailed comments on early drafts of this paper. We thank Geoff Marcy, Steve Bryson, and Guillermo Torres for their useful discussions related to the *Kepler* mission at the 2011 January AAS meeting and during the referee process, as well as François Fressin and other *Kepler* team members for enlightening discussions during and following the 2011 May Boston AAS meeting. We thank Jack Lissauer for pointing out the importance of considering the eccentricities of binary systems. We thank Jessica Lu for bringing TRILEGAL to our attention. Finally, we acknowledge the dedication and hard work of the *Kepler* team for opening up this amazing new frontier in exoplanetary science.

APPENDIX

BLENDED PLANETS

In the present work, we consider as false positives only astrophysical configurations that do not involve any planets but still mimic the signal of a transiting planet. However, there are various other scenarios involving "blended planets" that, while not strictly false positives (i.e., a transiting planet is still involved), may contribute significantly to uncertainty in the planet parameters derived from the transit signal. A "blended planet" for our purposes is a transit signal that appears to be a planet of a particular size transiting the target star but is actually a larger planet transiting a fainter blended star. As before, these scenarios can be divided into chance-alignment systems or physically associated hierarchical systems.

We have calculated that chance-alignment blended planets are significantly less likely to occur than their blended stellar binary cousins; this can be heuristically understood from the following considerations.

1. Because the deepest intrinsic planetary transits have depths of only $\sim$0.02 and the diluted signal has to be detectable (we have adopted $\delta \gtrsim 10^{-4}$ as a threshold), then the maximum contrast between the target star and the blending star is $\Delta m_K = 5.75$, which is significantly less than the $\Delta m_K = 10$ we adopted for blended binaries in Section 2.1.1. The sky density of stars available for the chance-alignment blended planet scenario is thus about 5.5 times lower than that for the blended binary scenario, according to the TRILEGAL simulations.
2. Our assumed planet frequency ($\sim$20%) is lower than our assumed binary fraction ($\sim$40%).
3. The largest planets, while the most amenable to causing the blended planet scenario because of their larger intrinsic transit depth, are the least common—only $\sim$1% of solar-type stars host close-in giant planets, and this occurrence

rate is even lower for lower-mass stars (Endl et al. 2003; Johnson et al. 2010), which are the most common blending stars.

Physically associated hierarchical planets, on the other hand, might well be relatively common compared to the stellar false positive scenarios or chance-alignment blended planets. Another way of saying this is that binary stellar systems are relatively common, and so it seems likely that a substantial fraction of *Kepler* targets (and therefore candidates) are in fact binaries of unknown architecture. The net effect of this on the interpretation of the sample of planet candidates will be additional uncertainty in the derived planet properties due to both diluting light from a binary companion and from possible stellar misclassification by the KIC, which assumes each star is single. We note that the *Kepler* team does include blended planets in the BLENDER procedure, and in fact that such scenarios are often the most difficult to rule out (*Kepler* team 2011, private communication).

In summary, while the analysis presented in this paper may provide confidence that "classic false positive" stellar systems are not often masquerading as *Kepler* transiting planet candidates, we do caution that uncertainties regarding candidate host systems (including whether or not they are binary) must be considered in any statistical analysis of the whole candidate sample.

REFERENCES

Almenara, J. M., Deeg, H. J., Aigrain, S., et al. 2009, A&A, 506, 337
Baglin, A. 2003, Adv. Space Res., 31, 345
Basri, G., Walkowicz, L. M., Batalha, N., et al. 2010, ApJ, 713, L155
Batalha, N. M., Borucki, W. J., Bryson, S. T., et al. 2011, ApJ, 729, 27
Batalha, N. M., Borucki, W. J., Koch, D. G., et al. 2010a, ApJ, 713, L109
Batalha, N. M., Rowe, J. F., Gilliland, R. L., et al. 2010b, ApJ, 713, L103
Beatty, T. G., & Gaudi, B. S. 2008, ApJ, 686, 1302
Biller, B. A. 2007, PhD thesis, Univ. Arizona
Borucki, W., Koch, D., Basri, G., et al. 2008, in IAU Symp. 249, Finding Earth-size Planets in the Habitable Zone: The Kepler Mission, ed. Y.-S. Sun, S. Ferraz-Mello, & J.-L. Zhou (Cambridge: Cambridge Univ. Press), 17
Borucki, W. J., Koch, D. G., Basri, G., et al. 2011, ApJ, 736, 19
Brown, T. M. 2003, ApJ, 593, L125
Chabrier, G. 2001, ApJ, 554, 1274
Chabrier, G., Baraffe, I., Allard, F., & Hauschildt, P. 2000, ApJ, 542, 464
Cumming, A., Butler, R. P., Marcy, G. W., et al. 2008, PASP, 120, 531
Endl, M., Cochran, W. D., Tull, R. G., & MacQueen, P. J. 2003, AJ, 126, 3099
Evans, T. M., & Sackett, P. D. 2010, ApJ, 712, 38
Fortney, J. J., Marley, M. S., & Barnes, J. W. 2007, ApJ, 659, 1661
Fressin, F., Torres, G., Desert, J.-M., et al. 2011, arXiv:1105.4647
Gaudi, B. S. 2005, ApJ, 628, L73
Gaudi, B. S., Seager, S., & Mallen-Ornelas, G. 2005, ApJ, 623, 472
Girardi, L., Bertelli, G., Bressan, A., et al. 2002, A&A, 391, 195
Girardi, L., Groenewegen, M. A. T., Hatziminaoglou, E., & da Costa, L. 2005, A&A, 436, 895
Hartman, J. D., Bakos, G., Stanek, K. Z., & Noyes, R. W. 2004, AJ, 128, 1761
Ho, S., & Turner, E. L. 2010, arXiv:1007.0245
Howard, A. W., Marcy, G. W., Johnson, J. A., et al. 2010, Science, 330, 653
Jenkins, J. M., Borucki, W. J., Koch, D. G., et al. 2010a, ApJ, 724, 1108
Jenkins, J. M., Caldwell, D. A., Chandrasekaran, H., et al. 2010b, ApJ, 713, L120
Johnson, J. A., Aller, K. M., Howard, A. W., & Crepp, J. R. 2010, PASP, 122, 905
Koch, D. G., Borucki, W., Webster, L., et al. 1998, Proc. SPIE, 3356, 599
Konacki, M., Torres, G., Sasselov, D. D., & Jha, S. 2003, ApJ, 597, 1076
Latham, D. W., Bakos, G. Á., Torres, G., et al. 2009, ApJ, 704, 1107
Latham, D. W., Brown, T. M., Monet, D. G., et al. 2005, BAAS, 37, 1340
Lissauer, J. J., Fabrycky, D. C., Ford, E. B., et al. 2011, Nature, 470, 53
Luhman, K. L., & Jayawardhana, R. 2002, ApJ, 566, 1132
O'Donovan, F. T., Charbonneau, D., Torres, G., et al. 2006, ApJ, 644, 1237
Poleski, R., McCullough, P. R., Valenti, J. A., et al. 2010, ApJS, 189, 134
Raghavan, D., McAlister, H. A., Henry, T. J., et al. 2010, ApJS, 190, 1
Rogers, L. A., Bodenheimer, P., Lissauer, J. J., & Seager, S. 2011, arXiv:1106.2807
Seager, S., Kuchner, M., Hier-Majumder, C. A., & Militzer, B. 2007, ApJ, 669, 1279
Tokovinin, A. A. 1999, VizieR Online Data Catalog, 412, 40075
Torres, G., Fressin, F., Batalha, N. M., et al. 2011, ApJ, 727, 24
Torres, G., Konacki, M., Sasselov, D. D., & Jha, S. 2004, ApJ, 614, 979
Winn, J. N. 2010, in Exoplanets, ed. S. Seager (Tucson, AZ: Univ. Arizona Press), 55