A TOY ANALYTIC TRANSIT SURVEY

LGB, KM, JNW
Memo for internal use

ABSTRACT

We derive the signal to noise distribution of detected planets in a toy transit survey. We give corresponding expressions for the number of detected planets. We then discuss the effect of various errors that one makes by ignoring binarity when deriving occurrence rates.

Subject headings: planets and satellites: detection

STATEMENT OF PROBLEM

Ah, transiting planets! We learn so much by studying them. But how much can we actually learn, and how much is messed up by binarity?

Imagine the following transit survey:

- Take a magnitude-limited sample of "stars" of a single spectral class. For instance, G2V dwarfs. In other words, choose all "points" on-sky with $m_\lambda < m_{\lambda,\rm lim}$, or equivalently with $F_\lambda > F_{\lambda,\rm lim}$, and also with colors that make you think they are this spectral class.
- The true population of "points" (hereafter, systems) contains both single and double star systems. Single star systems have luminosity in the observed bandpass L_1, radii R_1, and effective temperature $T_{\rm eff,1}$. Double star systems have luminosity in the observed bandpass $L_d = (1+\gamma_R)L_1$, for $\gamma_R = L_2/L_1$ the ratio of the luminosity of the secondary to the primary. The ratio of the two number densities in a volume-limited sample is the binary fraction[1].
- The true population of planets around these stars is as follows:

  A fraction $\Gamma_{t,s}$ of stars in single star systems have a planet of radius R_p, with orbital period P.

  A fraction $\Gamma_{t,d}$ of each star in a double star system has a planet of radius R_p, with orbital period P. For instance, if $\Gamma_{t,s} = \Gamma_{t,d} = 0.1$, on average each double system contributes 0.2 planets, and each single system 0.1 planets. Any astrophysical difference in planet formation between singles and binaries is captured by these two terms.

A signal to noise limited transit survey is then conducted. All planets with S/N > $(\mathrm{S/N})_{\rm min}$ are detected. The rest are not. Note that while the S/N threshold does imply a minimum stellar flux about which planets of a given radius and period can be detected, this minimum flux is generally different from the minimum used to define the original magnitude cut for a sample [see Batalha et al, 2010].

[1] The binary fraction is equivalent to the multiplicity fraction if there are no triple, quadruple, ... systems.

Consider the following set of questions:

1. How many single and double star systems, respectively, are in the sample? Correspondingly, how many stars are in the sample?
2. How many planets are in the sample? (Orbiting single stars, and orbiting double stars respectively).
3. What is the true occurrence rate?
4. How many planets are detected?
5. What occurrence rate does astronomer A, who has never heard of binary star systems, derive for planets of radius R_p and period P?
6. What occurrence rate does astronomer B, who accounts for the "2 for 1" effect of binarity (*i.e.* that the sample actually has more stars than astronomer A thought) derive?
7. What about astronomer C, who accounts for "2 for 1" *and* misclassification due to diluted radii? In other words, astronomer C has lots of Keck time, and did high resolution imaging followup on every candidate, and correctly classifies the planetary radii.
8. What about astronomer D, who additionally notes the importance of completeness?

1. HOW MANY STARS ARE IN THE SAMPLE?

Let N_s be the number of single star systems, and N_d the number of double star systems. Then the total number of stars in the sample can be written

$$N_{\rm stars} = N_s + 2N_d. \qquad (1)$$

In a magnitude-limited sample in which stars are uniformly distributed in volume, the number of stars will be the number density times the volume. If the volume is taken to be a sphere over which the number density is uniform,

$$N_i = n_i \frac{4\pi}{3} d^3_{\max,i}, \qquad (2)$$

for $i \in \{\text{single}, \text{double}\} \equiv \{s, d\}$, and

$$\frac{n_d}{n_s} = \text{binary fraction} \equiv \text{BF} \qquad (3)$$

by definition. The absolute normalization of the number density is a measured quantity, as is the binary fraction.

For G2V dwarfs, the latter is ≈ 0.45 [Duchene & Kraus, 2013]. The former is given by [Bovy 2017].

$d_{\max,i}$ in Eq. 2 is the maximum distance corresponding to the given magnitude limit:

$$d_{\max,i} = \left(\frac{L_i}{4\pi F_{\rm lim}}\right)^{1/2}, \tag{4}$$

where the limiting flux in the bandpass $F_{\rm lim}$ can also be stated in terms of the limiting magnitude $m_{\rm lim}$,

$$m_{\rm lim} = m_0 - \frac{5}{2}\log_{10}\left(\frac{F_{\rm lim}}{F_0}\right), \tag{5}$$

for m_0 a zero-point magnitude and F_0 its corresponding flux (as always, everything is implicitly written in a defined bandpass).

In Eq. 4, again $i \in \{\text{single}, \text{double}\}$, and as a consequence the maximum distance to which binary stars will be selected is greater than that of single stars, simply as a consequence of imposing a magnitude cut. The ratio of double to single systems is

$$\frac{N_d}{N_s} = \frac{n_d}{n_s}\left(\frac{d_{\max,d}}{d_{\max,s}}\right)^3 \tag{6}$$

$$= \mathrm{BF} \times (1+\gamma_R)^{3/2}. \tag{7}$$

In the nominal case of twin binaries ($\gamma_R = 1$), with a binary fraction $\mathrm{BF} = 0.5$, there are $\sqrt{2}$ more binary systems than single systems in the sample. Correspondingly, there are $2\sqrt{2}$ more stars in binary systems than stars in single systems.

As a comment on Eq. 2, if we wished to write a stellar number density profile that accounted for the vertical structure of the Milky Way, we might choose a profile either $\propto \exp(-z/H)$, or $\propto \mathrm{sech}^2(z/H)$ for z the distance from the galactic midplane and H a scale-height. Both density profiles would lead closed form analytic solutions.

2. HOW MANY PLANETS ARE IN THE SAMPLE?

The number of planets in the sample is

$$N_{\rm planets} = N_{\text{planets in single star systems}} + N_{\text{planets in double star systems}} \tag{8}$$

$$= \Gamma_{t,s} N_s + 2\Gamma_{t,d} N_d. \tag{9}$$

The factor of 2 accounts for the fact that there are twice as many stars in double star systems.

3. WHAT IS THE TRUE OCCURRENCE RATE?

The "true occurrence rate" is the average number of planets per star. Thus

$$\Gamma_t = \frac{N_{\rm planets}}{N_{\rm stars}} \tag{10}$$

$$\Gamma_t = \frac{\Gamma_{t,s}N_s + 2\Gamma_{t,d}N_d}{N_s + 2N_d}. \tag{11}$$

4. HOW MANY PLANETS ARE DETECTED?

The total number of planet detections is the sum of the number of planets detected in single star systems $N_{{\rm det},s}$ and the number of planets detected in double star systems $N_{{\rm det},d}$. These can be expressed individually. The former is

$$N_{{\rm det},s} = N_s \Gamma_{t,s} f_{s,{\rm S/N}>({\rm S/N})_{\min}}, \tag{12}$$

where the product $N_s\Gamma_{t,s}$ is the number of planets in the single star systems of the sample, and $f_{s,({\rm S/N}>({\rm S/N})_{\min})}$ is the fraction of these planets that have signal to noise greater than the minimum detection threshold.

$$N_{{\rm det},d} = 2N_d \Gamma_{t,d} f_{d,{\rm S/N}>({\rm S/N})_{\min}}, \tag{13}$$

where now $2N_d\Gamma_{t,d}$ is the number of planets in the double star systems of the sample, and the completeness term must account for any differences in the signal to noise distribution that come from stellar binarity.

As a note of warning, Eqs. 12 and 13 ignore the geometric transit probability, $\approx R_\star/a$. This means we are assuming that every existing planet in this universe transits. This assumption can easily be rectified, if desired.

4.1. *Analytic completeness*

The only terms we have yet to compute are the completeness terms, $f_{i,{\rm S/N}>({\rm S/N})_{\min}}$ for $i \in \{\text{single}, \text{double}\}$. We proceed as follows.

The signal S for a box-car train transiting planet is

$$\mathrm{S} = \delta\mathcal{D} \tag{14}$$

$$= \left(\frac{R_p}{R_\star}\right)^2 \mathcal{D}, \tag{15}$$

for R_p the planet's radius, $R_\star$ that of its host star, and $\mathcal{D}$ the dilution parameter defined as

$$\mathcal{D} = \begin{cases} L_1/L_d, & \text{if binary and target primary} \\ \gamma_R L_1/L_d, & \text{if binary and target secondary} \\ 1, & \text{if single,} \end{cases} \tag{16}$$

where L_1, L_d, and γ_R were defined in the opening monograph.

Assuming the only source of noise is Poissonian counting noise, the noise N can be written

$$\mathrm{N} = \frac{1}{\sqrt{N_\gamma}}, \tag{17}$$

for N_γ the number of photons received by the detector. This noise model is a useful simplification – see [Howell 2006, pg 75] for the full CCD equation. The number of received photons can be written

$$N_\gamma = F_\gamma^{\rm N} A N_{\rm tra} T_{\rm dur}, \tag{18}$$

for $F_\gamma^{\rm N}$ the photon number flux from the system $[\mathrm{ph\,cm^{-2}\,s^{-1}}]$, A the detector area, $T_{\rm dur}$ the transit duration, and $N_{\rm tra}$ the number of transits observed, which is multiplied in assuming the transits are "phase-folded".

Thus the signal to noise ratio can be written

$$\mathrm{S/N} = \delta\mathcal{D}\sqrt{F_\gamma^{\rm N} A N_{\rm tra} T_{\rm dur}}. \tag{19}$$

In passing, given the parameters that define a survey and planet type, Eq. 19 would need to be re-expressed with $N_{\rm tra}$ roughly the ratio of the observing baseline to the planet period, and $T_{\rm dur}$ a function of $R_\star, P, a$, and impact parameter b, and then perhaps averaged over b. We leave them as-is for subsequent development.

The interesting term in Eq. 19 that changes between stars of the same binarity class in our idealized sample is the square root of $F^{\rm N}_{\gamma}$. This is the term that leads to a distribution of signal to noises for different stars. The completenesses $f_{i,{\rm S/N}>({\rm S/N})_{\rm min}}$ can be directly expressed in terms of those probability density functions:

$$f_{i,{\rm S/N}>({\rm S/N})_{\rm min}} = \int_{({\rm S/N})_{\rm min}}^{\infty} {\rm d}\left(\frac{\rm S}{\rm N}\right)_i \ {\rm prob}\left(\frac{\rm S}{\rm N}\right)_i . \quad (20)$$

We keep the subscript i because the signal to noise distributions are different for the cases of single star systems ($i = s$) and double star systems ($i = d$). Notably:

- The dilution differs (Eq. 16).
- The photon number flux from the system differs.

To simplify notation, we let $x_i \equiv ({\rm S/N})_i$, and rewrite Eq. 20 as

$$f_{i,x>x_{\rm min}} = \int_{\rm x_{\rm min}}^{\infty} {\rm d}x_i \ {\rm prob}(x_i). \quad (21)$$

4.2. *Deriving* prob(x_i)

We want expressions for the probability density function of the observed signal to noise ratio, prob(x_i), for both the single and binary system case.

First, note that a star placed uniformly in the volume of the search space will have a probability density function for its distance r from the origin of

$${\rm prob}(r) = \frac{3r^2}{d^3_{\rm max}}, \quad (22)$$

where the appropriate maximum distances should be substituted per Eq. 4. Noting the transformation rule for probability density functions, we can evaluate the probability of a star having a observed number flux $F^{\rm N}_{i,\gamma}$ in the bandpass,

$$\begin{aligned}{\rm prob}(F^{\rm N}_{i,\gamma}) &= {\rm prob}(r(F^{\rm N}_{i,\gamma})) \left|\frac{{\rm d}r}{{\rm d}F^{\rm N}_{i,\gamma}}\right| \quad (23)\\ &= \frac{3}{2d^3_{\rm max}} c_i^{3/2} (F^{\rm N}_{i,\gamma})^{-5/2}, \quad (24)\end{aligned}$$

where in the latter equality we have written a "number luminosity" c_i (units of inverse time) defined for $i \in \{{\rm single}, {\rm double}\}$ as

$$c_i = \begin{cases} R_1^2 F^{\rm N}_{s1,\gamma}, & \text{if single} \\ R_1^2 F^{\rm N}_{s1,\gamma} + R_2^2 F^{\rm N}_{s2,\gamma} & \text{if double.}\end{cases} \quad (25)$$

In Eq. 25, $F^{\rm N}_{s1,\gamma}$ and $F^{\rm N}_{s2,\gamma}$ are the photon number fluxes at the surfaces of the stars. To derive Eq. 24, we simply scaled these by the distance:

$$F^{\rm N}_{i,\gamma} = \frac{c_i}{r^2}. \quad (26)$$

The surface photon number fluxes $F^{\rm N}_{si,\gamma}$ in Eq. 25 are usually evaluated numerically, by convolving the wavelength-specific photon flux density of a star with the dimensionless spectral response function of the instrument. In other words,

$$F^{\rm N}_{s,\gamma} = \int F_\lambda T_\lambda \, {\rm d}\lambda. \quad (27)$$

The wavelength-specific photon flux density F_λ [${\rm ph\,cm^{-2}\,s^{-1}\,\AA^{-1}}$] could be from Pickles' library, or could be a blackbody function. The transmission function is, up to factor of order unity, a step function between two wavelengths $\lambda_{\rm min}$ and $\lambda_{\rm max}$. If we assume a blackbody source, Eq. 27 becomes

$$F^{\rm N}_{s,\gamma} = 8\pi c\left(\frac{kT}{hc}\right)^3 \int_{hc/(\lambda_{\rm max}kT)}^{hc/(\lambda_{\rm min}kT)} \frac{u^2}{e^u - 1}\,{\rm d}u, \quad (28)$$

which can be evaluated numerically[2].

The importance of the functional form of $F^{\rm N}_{s,\gamma}$ is that it is to first order only a function of the blackbody temperature and the bandpass wavelength bounds. Thus in the most general case $c_i(R_1, R_2, T_1, T_2, \lambda_{\rm min}, \lambda_{\rm max})$ and nothing else. The only random variable involved in the flux being received at the detector is r, so we can indeed write the flux received at the detector as in Eq. 26.

We can finally write out the probability density functions for the signal to noise ratios in the single and double-star cases by using the transformation rule for pdfs, and applying Eq. 19. For single stars,

$${\rm prob}(x_s) = \frac{3}{d^3_{\rm max,s}} c_s^{3/2} \delta^3 \left(AT_{\rm dur}N_{\rm tra}\right)^{3/2} x_s^{-4}. \quad (29)$$

Analogously for double stars,

$${\rm prob}(x_d) = \frac{3}{d^3_{\rm max,d}} c_d^{3/2} (\mathcal{D}\delta)^3 \left(AT_{\rm dur}N_{\rm tra}\right)^{3/2} x_d^{-4}. \quad (30)$$

4.3. *Number of detected planets*

Performing the integrals of Eq. 21, we get:

$$\begin{aligned}N_{{\rm det},s} &= N_s \Gamma_{t,s} f_{s,{\rm S/N}>({\rm S/N})_{\rm min}} \quad (31)\\ &= N_s \Gamma_{t,s} \frac{1}{d^3_{\rm max,s}} c_s^{3/2} \delta^3 \left(AT_{\rm dur}N_{\rm tra}\right)^{3/2} x_{\rm min}^{-3}, \quad (32)\end{aligned}$$

and

$$\begin{aligned}N_{{\rm det},d} &= 2N_d \Gamma_{t,d} f_{d,{\rm S/N}>({\rm S/N})_{\rm min}} \quad (33)\\ &= 2N_d \Gamma_{t,d} \frac{1}{d^3_{\rm max,d}} c_d^{3/2} (\mathcal{D}\delta)^3 \left(AT_{\rm dur}N_{\rm tra}\right)^{3/2} x_{\rm min}^{-3}. \quad (34)\end{aligned}$$

Formally, the $f_{s,{\rm S/N}>({\rm S/N})_{\rm min}}$ term should really be written as $\min(1, \ldots)$, where $(\ldots)$ is the given expression. This ensures that the fraction of planets above the signal to noise threshold is less than 1.

[2] It may help in the numerics to note that infinite series representations of this type of dimensionless integral exist and converge quickly. For instance, one can show that

$$\int_0^a \frac{u^3}{e^u - 1}\,{\rm d}u = \sum_{n=1}^{\infty} \frac{6}{n^4} - \frac{e^{-an}}{n^4}(6 + 6an + 3(an)^2 + (an)^3).$$

A similar expression exists for the similar integral in the text. [Michels 1968] explains an analogous derivation.

The number of detected planets $N_{\rm det}$ is the sum of the two previously written equations, and can be written

$$N_{\rm det} = \left(\frac{\delta}{x_{\rm min}}\right)^3 (AT_{\rm dur}N_{\rm tra})^{3/2} \times \left[N_s\Gamma_{t,s}\frac{c_s^{3/2}}{d^3_{\rm max,s}} + 2N_d\Gamma_{t,d}\frac{c_d^{3/2}}{d^3_{\rm max,d}}\right]. \quad (35)$$

All terms can be input to a computer, and checked against a Monte Carlo simulation if desired.

5. ASTRONOMER A IGNORES BINARITY

Astronomer A has never heard of binary star systems. What occurrence rate does he derive for planets of radius R_p and period P?

The *total* occurrence rate (number of detections divided number of "stars") for Astronomer A would be $N_{\rm det}/(N_s + N_d)$. However, even though Astronomer A does not know about binaries, the radii he derives for any planets in binary systems are too small, by a factor $\sqrt{\mathcal{D}}$. The question is what occurrence rate is derived for planets of radius R_p and period P. The answer is thus

$$\Gamma_{\rm A,\ planets\ of\ R_p} = \frac{N_{\rm det,s}}{N_s + N_d}. \quad (36)$$

This astronomer will also think there is a second population of planets, with radius $R_p\sqrt{\mathcal{D}}$, and will thus rush to Nature claiming to also have derived a second occurrence rate,

$$\Gamma_{\rm A,\ planets\ of\ R_p\sqrt{\mathcal{D}}} = \frac{N_{\rm det,d}}{N_s + N_d}. \quad (37)$$

6. ASTRONOMER B COUNTS HOST STARS CORRECTLY

Astronomer B can somehow account correctly for the "2 for 1" effect of binarity, *i.e.* that the sample actually has more stars than astronomer A thought.

By the same token as above,

$$\Gamma_{\rm B,\ planets\ of\ R_p} = \frac{N_{\rm det,s}}{N_s + 2N_d}, \quad (38)$$

and

$$\Gamma_{\rm B,\ planets\ of\ R_p\sqrt{\mathcal{D}}} = \frac{N_{\rm det,d}}{N_s + 2N_d}. \quad (39)$$

7. ASTRONOMER C COUNTS HOST STARS CORRECTLY AND FIGURES OUT DILUTED RADII

Astronomer C did high resolution imaging followup on every candidate, and correctly classifies the planetary radii. Thus, she also knows which planets are in binary systems, and which are in single star systems.

She knows that the purported population of planets with radii $R_p\sqrt{\mathcal{D}}$ does not exist. All detected planets from this survey have radii R_p. She computes an occurrence rate

$$\Gamma_{\rm C,\ planets\ of\ R_p} = \frac{N_{\rm det,s} + N_{\rm det,d}}{N_s + 2N_d} \quad (40)$$

$$= \frac{N_{\rm det}}{N_s + 2N_d}, \quad (41)$$

the closest yet to the true rate (Sec. 3).

8. ASTRONOMER D COUNTS HOST STARS CORRECTLY, FIGURES OUT DILUTED RADII, AND ACCOUNTS FOR COMPLETENESS

Astronomer D, just like Astronomer C, knows which detections were around binaries and the associated radius correction.

Astronomer D notes that the number of detected planets is not equal to the actual number of planets in the sample. They must account for completeness. They do injection recovery, and derive independent estimates for their completeness functions about single stars, $f_{s,x>x_{\rm min}} \equiv f_s$ and about double star systems, $f_{d,x>x_{\rm min}} \equiv f_d$. With this knowledge in hand, they compute the respective single and binary occurrence rates

$$\Gamma_{t,s} = \frac{N_{\rm det,s}}{N_s f_s}, \quad (42)$$

$$\Gamma_{t,d} = \frac{N_{\rm det,d}}{2N_d f_d}. \quad (43)$$

With these in hand, they derive the overall occurrence rate

$$\Gamma_{\rm D,\ planets\ of\ R_p} = \frac{N_{\rm det,s}/f_s + N_{\rm det,d}/f_d}{N_s + 2N_d} \quad (44)$$

$$= \frac{\Gamma_{t,s}N_s + 2\Gamma_{t,d}N_d}{N_s + 2N_d}. \quad (45)$$

Although we admit it's been a bit of a slog, it happens that Astronomer D's occurrence rate is the true occurrence rate (cf. Eq. 11).

All it takes is $\approx$ a full semester at Keck, a system-by-system analysis, and perfect understanding of the completeness of the detection efficiency for single and double star systems.

9. NUMERICAL VERIFICATION

As a check on the preceding analytic development, we implemented a Monte Carlo simulation of this idealized transit survey. To run the survey, we defined the instrument specifications (detector area and transmission function), the stellar population (binary fraction, total number density of a given stellar class, binary light ratio, fixed stellar properties), the planet population (fixed planet radius, period, and occurrence rate about single and binary stars), and finally the survey parameters (observing baseline, minimum SNR for "detection"). We then randomly drew star positions, randomly assigned planets to stars in single and binary systems, and computed the resulting signal to noise (Eq. 19) with which the transits would be observed. As in the preceding analytics[3], we assumed "twin" binaries (same stellar radii, same effective temperature, and dilution does not depend on which stellar binary is the "target").

The results are shown in Fig. 1, and indicate that the analytic probability distribution functions Eqs. 29, 30, and the number of detections (Eq. 35) are correct.

A point evident in Fig. 1 is that, for fixed planet parameters, and fixed stellar parameters ($R_\star, L_\star$, and distance r) the SNR distribution for planets in binaries is poorer than that of planets in single star systems. We can see

[3] TODO could be to generalize away from this

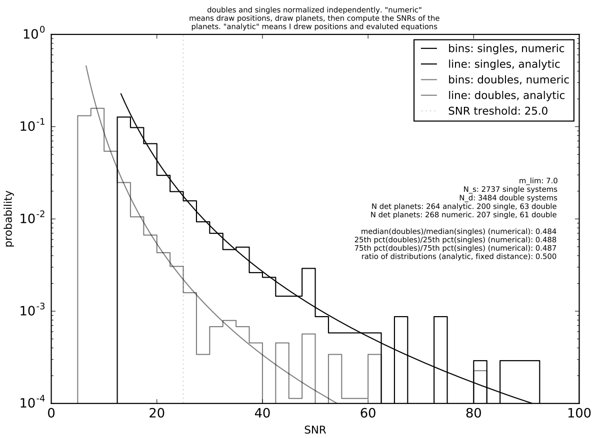

FIG. 1.— Comparison of analytic and numeric probability density functions of the SNR in an idealized transit survey. The analytic lines are Eqs. 29 and 30 for the planet populations orbiting single and binary stars. The underlying stepped histogram is output from Monte Carlo simulations. Poisson noise leads to a small deviation at the faint and bright limits, but the numerics and analytics otherwise agree.

analytically that this can be written only as function of the binary light ratio:

$$\frac{\mathrm{prob}(x_d)}{\mathrm{prob}(x_s)} = (1+\gamma_R)^{-1}. \tag{46}$$

Deriving this simple form requires noting that the ratios of the bandpass-specific number luminosities (Eq. 25) is equal to the ratio of the bandpass-specific energy luminosities (otherwise a term with c_s/c_d must be included).

10. REPRESENTATIVE NUMBERS FOR A FEW CASES

10.1. *Twin binaries: if we ignore binarity, for what fraction of detections do we misclassify the radii?*

Ignoring binarity, we will detect $N_{\mathrm{det,s}}$ planets around single stars, and $N_{\mathrm{det,d}}$ planets around double stars. The latter set will be assumed to have radii $R_p\sqrt{\mathcal{D}}$. The fraction of detections with misclassified radii can then be written

$$\frac{N_{\mathrm{det,d}}}{N_{\mathrm{det,s}}+N_{\mathrm{det,d}}} = \frac{1}{1+\alpha}, \tag{47}$$

for

$$\alpha \equiv \frac{1}{2(\mathrm{BF})}(1+\gamma_R)^{3/2}\frac{\Gamma_{t,s}}{\Gamma_{t,d}}. \tag{48}$$

For the nominal G2V dwarf case of BF = 0.45, twin binaries with equal occurrence rates this produces a misclassification rate of 24%, in agreement with Fig. 1.

10.2. *Twin binaries: if we ignore binarity, how wrong is our occurrence rate for planets of radius R_p?*

This is almost simply asking "what is the relative difference between the occurrence rates derived by Astronomers D and A for planets of radius R_p?" However, in the more realistic case, Astronomer A also has derived a completeness, which we assume is the same as for Astronomer D in the single star case. So Astronomer A now misclassifies planetary radii, and miscounts the total number of stars, but knows his completeness for single stars. Astronomer D corrects all these errors.

For brevity, write $\Gamma_{\mathrm{A,\ planets\ of\ R_p}} = \Gamma_{\mathrm{A,R_p}}$, and similarly for D. Then the relative difference between the two occurrence rates is

$$\left|\frac{\Gamma_{\mathrm{A,R_p}}-\Gamma_{\mathrm{D,R_p}}}{\Gamma_{\mathrm{A,R_p}}}\right| = \left|1-\frac{\Gamma_{\mathrm{D,R_p}}}{\Gamma_{\mathrm{A,R_p}}}\right| \tag{49}$$

$$= \left|1-\left(\frac{\Gamma_{t,s}N_s+2\Gamma_{t,d}N_d}{N_s+2N_d}\cdot\frac{N_s+N_d}{\Gamma_{t,s}N_s}\right)\right| \tag{50}$$

$$= \left|1-\frac{(1+2\beta\Gamma_{t,d}/\Gamma_{t,s})(1+\beta)}{(1+2\beta)}\right|, \tag{51}$$

for

$$\beta \equiv N_d/N_s = \mathrm{BF}\times(1+\gamma_R)^{3/2}. \tag{52}$$

For the nominal G2V dwarf case of BF = 0.45 with twin binaries ($\gamma_R = 1$) and $\Gamma_{t,d} = \Gamma_{t,s}$ this gives a relative error of 127%. For instance, in the numerical simulation corresponding to Fig. 1, Astronomer A finds $\Gamma_{\mathrm{A,R_p}} = 0.22$, while Astronomer D derives the true (input) occurrence rate of $\Gamma_{\mathrm{D,R_p}} = 0.5$.