Draft version October 24, 2019
Preprint typeset using LaTeX style emulateapj v. 12/16/11

EVOLUTION OF THE OCCURRENCE RATE OF SMALL CLOSE-IN PLANETS AROUND LOW MASS STARS FROM *KEPLER* AND *K2*

Ryan Cloutier[1,2] and Kristen Menou[3,2] et al.

Draft version October 24, 2019

ABSTRACT

Recent observational studies have revealed a prominent gap in the occurrence rate of close-in planet radii around Sun-like stars. Resolving the so-called radius valley around low mass stars can provide valuable constraints on the physical mechanisms that sculpt the valley and so far have been largely limited by relatively poor counting statistics. Here we calculate the occurrence rate of small close-in planets around mid-K to mid-M dwarfs using the known planet populations from the *Kepler* and *K2* missions while exploiting the *Gaia* DR2 data to refine the stellar parameters. The application of appropriate completeness corrections clearly reveals the radius valley in the maximum a-posteriori occurrence rates as a function of orbital period and planet radius. **The slope of the valley with orbital separation is shown to differ in sign from that measured around Sun-like stars which suggests that thermally driven mass loss may not dominate the formation and evolution of planets in the low stellar mass regime.** The prominence and location of radius valley features are also shown to evolve with stellar mass as the relative occurrence of terrestrial to gaseous planets increases by an order of magnitude from mid-K to mid-M dwarfs and each valley feature is shifted to smaller planet sizes with decreasing stellar mass in agreement with physical models of photoevaporation, core-powered mass loss, and gas-poor formation. Although current measurements are insufficient to robustly identify the dominant formation pathway of the radius valley, we argue that such inferences may be obtained by *TESS* with $\mathcal{O}(85,000)$ mid-to-late M dwarfs observed with 2 minute cadence.

1. INTRODUCTION

NASA's *Kepler* space telescope has discovered thousands of exoplanets over its lifetime and consequently enabled robust investigations of the occurrence rate of planets within our galaxy. One striking outcome of such studies was that the so-called super-Earths and sub-Neptunes—whose radii span sizes intermediate between those of the Earth and Neptune—represent the most common type of planet around Sun-like stars and early M dwarfs alike (e.g. Youdin 2011; Howard et al. 2012; Dressing & Charbonneau 2013; Fressin et al. 2013; Petigura et al. 2013; Morton & Swift 2014; Dressing & Charbonneau 2015; Mulders et al. 2015; Gaidos et al. 2016; Fulton et al. 2017; Hardegree-Ullman et al. 2019). Furthermore, mass measurements of many of these transiting planets via transit-timing variations or precision radial velocity measurements revealed that the majority of planets smaller than ~ 1.6 $R_\oplus$ are consistent with having bulk terrestrial compositions (e.g. Weiss & Marcy 2014; Dressing et al. 2015; Rogers 2015).

Early studies of the *Kepler* planet population hinted that planets at small orbital separations exhibited a bimodal radius distribution (e.g. Owen & Wu 2013)—commonly referred to as the radius valley—that is thought to be representative of a population of small, predominantly rocky planets plus a population of inflated gaseous planets that have retained significant H/He envelopes. Consequently, numerous studies of planet formation and evolution sought to explain the apparent bimodality. One such proposed mechanism is that of photoevaporation wherein the gaseous envelopes of small close-in planets may be stripped by X-ray and extreme ultraviolet (XUV) radiation from their host stars during the first ~ 100 Myrs of the planet's lifetime (Jackson et al. 2012; Owen & Wu 2013; Jin et al. 2014; Lopez & Fortney 2014; Chen & Rogers 2016; Owen & Wu 2017; Jin & Mordasini 2018; Lopez & Rice 2018). Another possible explanation invokes gas-poor formation wherein gas accretion is delayed by dynamical friction whilst the planetary core is still embedded within the protoplanetary disk until a point at which the gaseous disk has almost completely dissipated after just a few Myrs (Lee et al. 2014; Lee & Chiang 2016; Lopez & Rice 2018). More recently, the radius valley may also be explained by core-powered mass loss wherein the luminosity from a planetary core's primordial energy reservoir from formation drives atmospheric escape over Gyr timescales (Ginzburg et al. 2018; Gupta & Schlichting 2019a,b).

Observational tests of the aforementioned theoretical frameworks have become feasible in recent years due to the precise refinement of measured planet radii following improved stellar host characterization via spectroscopy, asteroseismology, and *Gaia* parallaxes (e.g. Fulton et al. 2017; Berger et al. 2018; Fulton & Petigura 2018; Van Eylen et al. 2018; Martinez et al. 2019). Each of these independent studies clearly resolved the radius valley among small close-in planets orbiting Sun-like stars. A variety of trends were also observed in either the raw or in the completeness-corrected (i.e. the occurrence rate) distributions of close-in planets. Firstly, the location of

ryan.cloutier@cfa.harvard.edu
[1] Center for Astrophysics | Harvard & Smithsonian, 60 Garden Street, Cambridge, MA, 02138, USA
[2] Dept. of Astronomy & Astrophysics, University of Toronto, 50 St. George Street, Toronto, ON, M5S 3H4, Canada
[3] Physics & Astrophysics Group, Dept. of Physical & Environmental Sciences, University of Toronto Scarborough, 1265 Military Trail, Toronto, ON, M1C 1A4, Canada

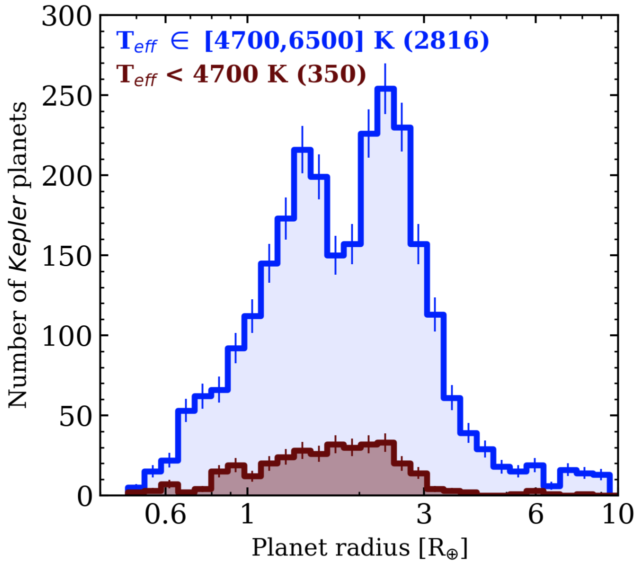

FIG. 1.— Empirical distributions of *Kepler* planet radii. Histograms of *Kepler* planet radii from Berger et al. (2018) for planets with host stellar effective temperatures $T_{\rm eff} \in [4700, 6500]$ K (*blue*) and $T_{\rm eff} < 4700$ K (*red*). The former subset of 2816 planets corresponds to the effective temperature range considered in the California Kepler Survey (CKS; Fulton et al. 2017) wherein the radius valley is clearly resolved in the empirical distribution even without completeness corrections. A similar bimodal structure is not resolved in the empirical distribution of the latter subset around low mass stars due in-part to the relatively poor counting statistics with just 350 planets.

the radius valley around FGK stars is period-dependent with slope $\mathrm{d}\log r_p/\mathrm{d}\log P \sim -0.1$ (Van Eylen et al. 2018; Martinez et al. 2019), a result that is consistent with both photoevaporation and core-powered mass loss models but is inconsistent with the late formation of terrestrial planets in a gas-poor environment. Secondly, the feature locations (i.e. the weighted average radius of the peaks and valley) appear to exist at smaller planet radii with decreasing stellar mass (Fulton & Petigura 2018; Wu 2019).

In this study, we extend the investigation of the occurrence rate of small close-in planets to systems hosted by low mass dwarf stars later than mid-K dwarfs. The empirical population of known planets in this stellar mass regime features nearly an order of magnitude fewer planets than around Sun-like stars, thus making the detection of the radius valley around low mass stars more difficult and at a lower signal-to-noise. This fact is clearly evidenced in the empirical *Kepler* planet population for which the radius valley around Sun-like stars ($T_{\rm eff} \in [4700, 6500]$ K) is clearly exhibited whereas a similar feature around low mass stars ($T_{\rm eff} < 4700$ K) is not easily discernible by-eye (Fig. 1 based on the data from Berger et al. 2018). This study leverages the precise stellar parallaxes from the *Gaia* DR2 for low mass stars observed by *Kepler* and *K2* to refine the stellar parameters and compute precise occurrence rates of close-in planets with the goal of resolving the radius valley and accurately measuring the location of the radius valley features and their uncertainties. Although it is unlikely that a single physical mechanism is responsible for sculpting the radius valley, investigation the evolution of the valley features with stellar mass can allude to which process—if any—dominates the evolution of close-in planets.

In Sects. 2 and 3 we define our stellar sample from *Kepler* and *K2* and compile our sample of confirmed planets from each mission. In Sect. 4 we derive the transiting planet detection completeness and use those results to calculate the occurrence rate of small close-in planets in which the structure of the radius valley around low mass stars is resolved (Sect. 5). Our results as a function stellar mass are compared to model predictions in Sect. 6. Sect. 7 presents a discussion of our results and its implications with a summary of our main findings being presented in Sect. 8.

2. LOW MASS DWARF STELLAR SAMPLE

The goal of this study is to extend measurements of the occurrence rate of close-in planets to planetary systems hosted by low mass dwarf stars with effective temperatures $T_{\rm eff} < 4700$ K: the lower limit of $T_{\rm eff}$ considered in the California Kepler Survey (CKS; Fulton et al. 2017). This adopted temperature threshold approximately corresponds to spectral types later than K3.5V (Pecaut & Mamajek 2013). In the following subsections we define our stellar sample from both *Kepler* or *K2*.

2.1. *Kepler Stellar Sample*

Following the release of *Gaia* DR2 (Lindegren et al. 2018), Berger et al. (2018) cross-matched *Kepler* target stars with DR2 and compiled a catalog of stellar parallaxes ϖ, 2MASS K_s-band magnitudes, and spectroscopic measurements of $T_{\rm eff}$, $\log g$, and [Fe/H] for $\sim 178,000$ stars observed as part of the primary *Kepler* mission. Spectroscopic measurements were obtained from either the Data Release 25 (DR25) Kepler Stellar Properties Catalog (KSPC; Mathur et al. 2017), the California Kepler Survey (CKS; Petigura et al. 2017) where available, and $T_{\rm eff}$ values for stars with $T_{\rm eff} < 4000$ K were compiled from Gaidos et al. (2016). The full set of available stellar parameters were used as input within the spectral classification code `isoclassify` (Huber et al. 2017) to calculate stellar luminosities. The resulting luminosity values were consequently combined with $T_{\rm eff}$ measurements to refine the stellar radii using the Stefan-Boltzmann law for the majority of *Kepler* FGK stars. However, bolometric corrections for *Kepler* M dwarfs with $T_{\rm eff} < 4100$ K and absolute K_s-band magnitudes $M_{K_s} > 3$ are known to suffer significant inaccuracies owing to incomplete molecular line lists. For these stars, Berger et al. (2018) instead adopted the empirically-derived M dwarf radius-luminosity relation from Mann et al. (2015) to refine the M dwarf stellar radii. Berger et al. (2018) also combined the $T_{\rm eff}$ luminosity measurements to derive stellar evolutionary flags aimed at classifying stars as either a dwarf, a subgiant, or a red giant.

Stellar masses M_s are not reported by Berger et al. (2018). In order to study the *Kepler* planet population as a function of M_s, we derive M_s values given the measured stellar radii R_s using the mass-radius relation from Boyajian et al. (2012) which is applicable to both K and M dwarfs. Boyajian et al. (2012) acquired interferometric measurements with the *CHARA* array of 21 nearby K and M dwarfs to measure the angular size of each stellar disk at the level of $\lesssim 5\%$. Their stellar sample was supplemented by 12 literature measurements of R_s from interferometry. Mass measurements were then derived using the K_s-band mass-luminosity relation from Henry

& McCarthy (1993) which was valid for their full stellar sample spanning 0.13-0.90 $R_\odot$. Boyajian et al. (2012) parameterized the stellar mass-radius relationship as a quadratic in M_s and reported values and uncertainties for each polynomial coefficient. Here, we assume independent Gaussian probability density functions (PDF) for each coefficient and sample their values along with each star's R_s from their respective measurement uncertainties to derive the M_s PDF for all of the low mass dwarfs in our preliminary *Kepler* sample.

We define our final *Kepler* stellar sample by focusing on stars that satisfy the following criteria:

1. *Kepler* magnitude $K_p < 16$,
2. $T_{\rm eff} - \sigma_{T_{\rm eff}} \leq 4700$ K,
3. $R_s - \sigma_{R_s} \leq 0.8$ $R_\odot$,
4. $M_s - \sigma_{M_s} \leq 0.8$ $M_\odot$, and
5. and an evolutionary flag corresponding to a dwarf star.

We also only consider *Kepler* stars for which reliable completeness products from DR25 are available (see Sect. 4.1). Based on these criteria, we retrieve 3965 low mass *Kepler* stars whose stellar parameters are depicted in Fig. 2. In our *Kepler* sample, the *Kepler* magnitudes span $K_p \in [10.35, 16.00]$ with a median value of 15.16, effective temperatures span $T_{\rm eff} \in [3154, 4870]$ K with a median value of 4394 K, stellar radii span $R_s \in [0.17, 0.87]$ $R_\odot$ with a median value of 0.68 $R_\odot$, and stellar masses span $M_s \in [0.13, 0.88]$ $M_\odot$ with a median value of 0.70 $M_\odot$. Our final *Kepler* sample boasts a median fractional R_s uncertainty of $\sim 6.7\%$ which is $\sim 4-5$ times smaller than the typical R_s uncertainty reported in the KSPC. The median fractional uncertainty on M_s is $\sim 5.5\%$.

2.2. *K2 Stellar Sample*

We first retrieved the list of probable low mass dwarf stars observed in any *K2* campaign by querying MAST[4]. Our initial search was restricted to *K2* stars with $T_{\rm eff}$ < 4900 K, $\log g > 4$, and $R_s < 1$ $R_\odot$. Note that these criteria are not intended to represent the parameter ranges for low mass dwarf stars but are intended as conservative conditions to encapsulate all such stars prior to their refinement using the *Gaia* DR2 data. From MAST we retrieve each star's Ecliptic Plane Input Catalog (EPIC) numerical identifier, stellar photometry in the *Kepler* bandpass K_p and 2MASS bands JHK_s, along with measured values of $T_{\rm eff}$, $\log g$, [Fe/H], and R_s.

We proceed with refining the stellar parameters by cross-matching our initial *K2* sample with *Gaia* DR2 using the *Gaia*-*K2* data products from Megan Bedell[5]. Where available, we retrieve reach star's celestial coordinates, stellar parallaxes ϖ, and *Gaia* photometry. Measurements of R_s then follow from the methodology of Berger et al. (2018) and outlined as follows. The formalism of Bailer-Jones et al. (2018) is used to transform

[4] Mikulski Archive for Space Telescopes, https://archive.stsci.edu/k2/.
[5] https://gaia-kepler.fun/

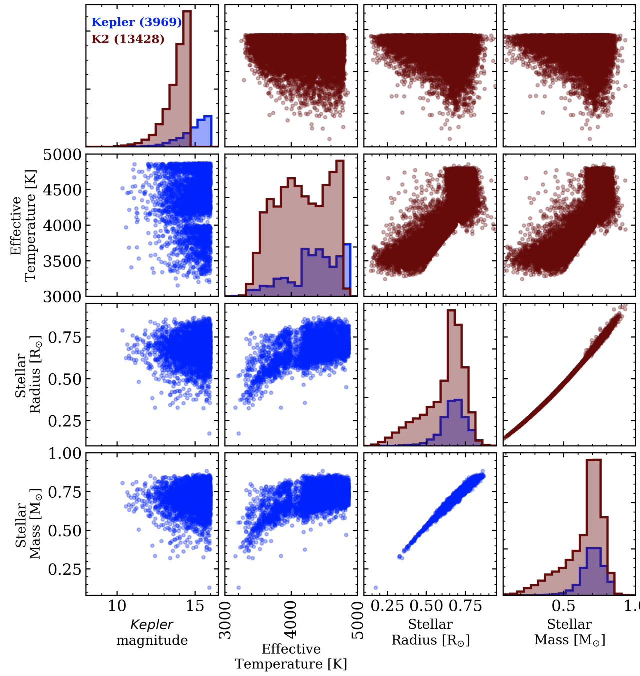

Fig. 2.— Low mass dwarf stellar samples from *Kepler* and *K2*. Distributions of *Kepler* magnitudes, effective temperatures, stellar radii, and stellar masses for stars in our final stellar sample from either *Kepler* (*blue histogram and markers*) or *K2* (*red histogram and markers*).

the assumed Gaussian-distributed ϖ PDFs into stellar distance PDFs which need not remain Gaussian. Using the measured distances d and celestial coordinates, we interpolate over the E_{B-V} extinction maps using the `mwdust` software (Bovy et al. 2016) to derive both the V and K_s-band extinction coefficients A_V and A_{K_s}. We then calculate each star's absolute K_s-band magnitude $M_{K_s} = K_s - \mu - A_{K_s}$ where the distance modulus is $\mu = 5\log_{10}(d/10~{\rm pc})$.

For the earliest stars in our sample ($M_{K_s} \leq 4.6$) , for which the bolometric corrections are still reliable, we interpolate the MIST bolometric correction grids (Choi et al. 2016) over $T_{\rm eff}$, $\log g$, [Fe/H], and A_V to derive the K_s-band bolometric corrections BC_{K_s}. We then compute the absolute bolometric magnitudes $M_{\rm bol} = M_{K_s} + BC_{K_s}$ and consequently the bolometric stellar luminosities as

$$L_{\rm bol} = L_0 \cdot 10^{-0.4 M_{\rm bol}}, \quad (1)$$

where $L_0 = 3.0128 \times 10^{28}$ W (Mamajek et al. 2015). The refined R_s values are then calculated using the Stefan-Boltzmann law given $L_{\rm bol}$ and $T_{\rm eff}$ with measurement uncertainties propagated throughout.

For the remaining late type stars with $M_{K_s} > 4.6$, we revert to the empirically-derived radius-luminosity relation from Mann et al. (2015) to calculate the M dwarf stellar radii. Mann et al. (2015) fit a second-order polynomial to R_s as a function of M_{K_s} which has a characteristic dispersion in the fractional radius uncertainty of 2.89%. To quantify the final R_s uncertainty we sample M_{K_s} from its posterior PDF and transform each M_{K_s} draw to an R_s value using the aforementioned radius-luminosity relation. To each star's derived R_s PDF, we add an additional dispersion term, in quadrature, whose fractional uncertainty is 2.89%. Stellar masses within our *K2* sample are derived identically to the method applied to the *Kepler* sample using the Boyajian et al. (2012) stellar mass-radius relation (see Sect. 2.1).

We define our final *K2* stellar sample of low mass dwarf stars similarly to our definition of the *Kepler* sample. Explicitly, we focus on stars that obey the following criteria:

1. $K_p < 14.7$,
2. $T_{\rm eff} - \sigma_{T_{\rm eff}} \leq 4700$ K,
3. $R_s - \sigma_{R_s} \leq 0.8$ R$_\odot$,
4. $M_s - \sigma_{M_s} \leq 0.8$ M$_\odot$, and
5. $R_s < R_{s,\max}$.

Because our *K2* sample lacks any evolutionary flags, we adopt the following ad hoc upper limit on R_s from Fulton et al. (2017) that aims to reject evolved stars:

$$R_{s,\max} = \mathrm{R}_\odot \cdot 10^{0.00025(T_{\rm eff}/\mathrm{K}-5500)+0.2}. \tag{2}$$

Based on these criteria, we retrieve 13428 low mass *K2* stars whose stellar parameters are also depicted in Fig. 2. In our *K2* sample, the *Kepler* magnitudes span $K_p \in [8.47, 14.68]$ with a median value of 14.04, effective temperatures span $T_{\rm eff} \in [3246, 4856]$ K with a median value of 4017 K, stellar radii span $R_s \in [0.14, 0.94]$ R$_\odot$ with a median value of 0.70 R$_\odot$, and stellar masses span $M_s \in [0.09, 0.93]$ M$_\odot$ with a median value of 0.69 M$_\odot$. The stars in this sample exhibit a median fractional R_s uncertainty of $\sim 3.5\%$ which is ~ 2 times smaller than the typical R_s uncertainty obtained for stars in our *Kepler* sample. The median fractional uncertainty on M_s is $\sim 3.9\%$.

Our complete stellar sample contains 17393 stars. Each of the *Kepler* and *K2* stellar samples are dominated by mid-to-late K dwarfs with temperatures and radii $\gtrsim 3800$ K and $\gtrsim 0.6$ R$_\odot$ respectively. This fact will have important implications on our ability to precisely measure the planet occurrence rate around the lowest mass stars in our sample.

3. POPULATION OF SMALL CLOSE-IN PLANETS AROUND LOW MASS DWARF STARS

Here we define the population of small close-in planets orbiting stars contained in our stellar sample. Our initial sample of transiting planets from either *Kepler* or *K2* were retrieved from the NASA Exoplanet Archive (Akeson et al. 2013) on June 15, 2019. Only confirmed planets—based on their Exoplanet Archive dispositions—with orbital periods $P \in [0.5, 100]$ days are included. By considering confirmed planets only we naturally focus on a subset of the true empirical population of small close-in planets without being contaminated by various astrophysical false positive scenarios that may plague the planet candidates excluded from our initial sample.

The refined stellar radii derived in Sect. 2 enable us to derive more precise planetary radii. We refine the planetary radii r_p by retrieving point estimates of each planet's scaled planetary radius r_p/R_s which often includes a median value accompanied by the 16$^{\rm th}$ and 84$^{\rm th}$ percentiles. In cases for which the r_p/R_s uncertainties are symmetric, we assume that the r_p/R_s posterior PDF is Gaussian. For planets with asymmetric reported uncertainties, we fit the r_p/R_s percentiles with a skew-normal distribution using the `scipy.skewnorm python` class. We fit for the location, scale, and shape parameters of the distribution such that its resulting percentiles are consistent with the r_p/R_s point estimates reported for each planet. The refined planetary radii are then derived by sampling the fitted r_p/R_s and R_s distributions. We then update our planet sample by only considering planets whose radii are consistent with $r_p = 0.5 - 4$ R$_\oplus$.

From the distributions of R_s, $T_{\rm eff}$, M_s, and P or each planet and host star, we derive the planets' semimajor axes a and insolations F via

$$\frac{F}{F_\oplus} = \left(\frac{R_s}{R_\odot}\right)^2 \left(\frac{T_{\rm eff}}{5777\text{ K}}\right)^4 \left(\frac{a}{1\text{ AU}}\right)^{-2}. \tag{3}$$

Our final sample of confirmed small close-in planets contains 275 *Kepler* and 53 *K2* planets respectively. Their respective median fractional radius uncertainties are 7.1% and 9.0%. Properties of the 328 confirmed planets in our sample are reported in Tables 1 and 2. Our planet sample is depicted in Fig. 3 as two-dimensional maps of the number of planet detections in the period-radius and insolation-radius spaces. The two-dimensional histogram maps are computed by Monte-Carlo sampling planets from their F and r_p measurement uncertainties and with a fractional precision on P inflated to 20%.

The empirical planet popluation in Fig. 3 highlights many of the known features in the distribution of planets orbiting low mass stars (e.g. Morton & Swift 2014; Dressing & Charbonneau 2015; Gaidos et al. 2016). Namely, the dearth of planets with $r_p \gtrsim 2$ R$_\oplus$ at short orbital periods known as the Neptunian desert (Lundkvist et al. 2016; Mazeh et al. 2016), the prominence of super-Earth and sub-Neptune-sized planets with orbital periods of a few to tens of days, and the lack of small planets at long orbital periods ($P \gtrsim 40$ days) due to the poor transit detection completeness in this region. Any features resembling the radius valley are not prominent in the empirical planet distribution. Assuming that the observed radius valley around Sun-like stars persists in some form around the low mass stars in our sample, the fact that a distinct valley is not visible highlights the importance of measuring valley features from the completeness-corrected planet distribution. Alternatively, the valley—close to the expected rocky-to-gaseous transition of $\sim 1.5 - 1.8$ R$_\oplus$ (Weiss & Marcy 2014)—may not be entirely void of planets. Indeed there exists a signficant subset of planets between 1.5-1.8 R$_\oplus$ with periods out to ~ 12 days indicating that the mechanism for scultping the radius valley might not be as efficient as it is when operating on planetary systems around Sun-like stars.

4. TRANSITING PLANET DETECTION COMPLETENESS

Derivation of the planet occurrence rate requires the empirical distribution of planet detections to be corrected for imperfect survey completeness. The completeness correction is treated separately for each subset of planets from *Kepler* or *K2* in the following subsections. Each set of corrections is designed to account for detection biases

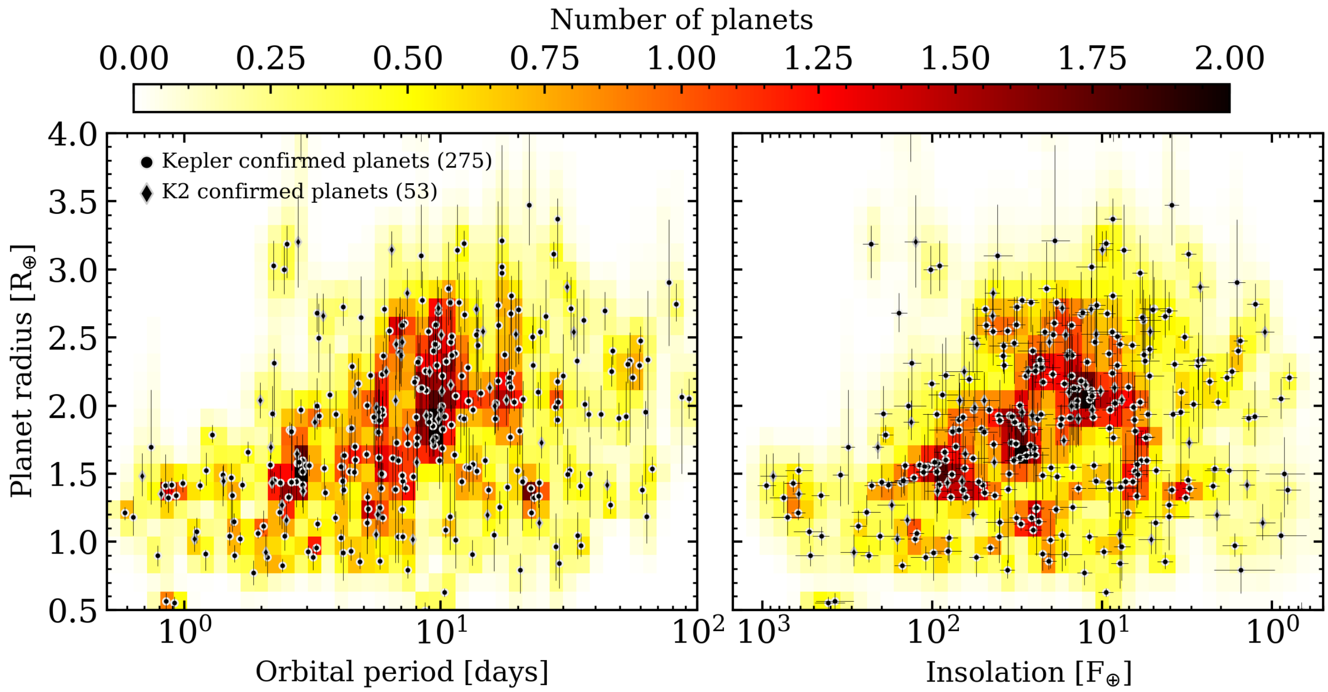

FIG. 3.— Empirical population of confirmed close-in planets around low mass stars. The distribution of 275 and 53 confirmed planets from *Kepler* and *K2* respectively as a function of orbital period, insolation, and planet radius. The two-dimensional maps are Monte-Carlo sampled from the measurement uncertainties on the planetary radii and insolations while the fractional uncertainties on the orbital periods are inflated to 20%.

TABLE 1
Kepler CONFIRMED PLANET PARAMETERS

KIC	Planet name	P [days]	F [F$_\oplus$]	F upper limit [F$_\oplus$]	F lower limit [F$_\oplus$]	r_p [R$_\oplus$]	r_p upper limit [R$_\oplus$]	r_p lower limit [R$_\oplus$]
1873513	Kepler-1624 b	3.29030	36.3	5.9	5.5	6.53	0.19	0.20
2556650	Kepler-1124 b	2.85235	47.4	5.0	4.6	1.97	0.08	0.10
2715135	Kepler-753 b	5.74771	41.3	5.1	4.8	1.89	0.30	0.12
3234598	Kepler-383 b	12.90468	20.6	3.2	2.5	1.54	0.30	0.17
3234598	Kepler-383 c	31.20122	6.4	0.9	0.7	1.49	0.34	0.22

TABLE 2
K2 CONFIRMED PLANET PARAMETERS

EPIC	Planet name	P [days]	F [F$_\oplus$]	F upper limit [F$_\oplus$]	F lower limit [F$_\oplus$]	r_p [R$_\oplus$]	r_p upper limit [R$_\oplus$]	r_p lower limit [R$_\oplus$]
201110617	K2-156 b	0.81315	1146.7	5038.8	734.4	1.35	0.12	0.10
201155177	K2-42 b	6.68796	93.9	353.9	58.8	2.45	0.27	0.25
201205469	K2-43 b	3.47114	83.9	316.0	52.2	4.09	0.28	0.25
201208431	K2-4 b	10.00440	34.2	167.1	23.5	2.52	0.34	0.31
201338508	K2-5 b	5.73597	47.3	209.2	29.5	1.95	0.17	0.18

arising from the imperfect transit detection sensitivity and for the geometric probability of a planetary transit to occur.

4.1. *Kepler Sensitivity*

The derivation of the *Kepler* planet detection sensitivity follows from the methodology outlined in Christiansen et al. (2016) and used by Fulton et al. (2017) to resolve the radius valley around FGK stars. Per-target *Kepler* completeness products for DR25 and the SOC 9.3 version of the *Kepler* pipeline (Jenkins et al. 2010) are available for all of the stars in our *Kepler* sample (Burke et al. 2015; Burke & Catanzarite 2017). Detection sensitivities (or efficiencies) were calculated via transiting planetary signal injections at the pixel level which are subsequently processed by the *Kepler* pipeline Transiting Planet Search (TPS) module from which the detection sensitivity is computed as the fraction of injected signals that are successfully recovered by the pipeline as a function of the Multi-event statistic (MES; Christiansen et al. 2015, 2017).

The MES represents the level of significance of a repeating transit signal at a specified transit duration ranging from 1.5-15 hours. Following Petigura et al. (2018), we adopt an alternative diagnostic for the transit signal significance in the form of the transit signal-to-noise ratio

$$\mathrm{S/N} = \frac{Z}{\mathrm{CDPP}_D}\sqrt{n_{\mathrm{transits}}(\mathbf{t}, P, T_0)} \tag{4}$$

where $Z=(r_p/R_s)^2$ is the transit depth assuming a non-grazing transit (i.e. $b \lesssim 0.9$), CDPP_D is the Combined Differential Photometric Precision on the timescale of the transit duration D (Koch et al. 2010), and n_{transits} is the number of observed transits given the target's data span and duty cycle of the observations $\mathbf{t}$, the planet's orbital period P, and its time of mid-transit T_0.

To compute the *Kepler* detection sensitivity as a function of S/N, we first derive the mapping between the MES and the transit S/N using the data from Christiansen et al. (2015) who derived the detection sensitivity of the *Kepler* pipeline from one year of data. The parameters of the injected planets are provided along with their corresponding MES and CDPP at each value of D considered. For each injected planet we interpolate its MES and CDPP values to D and calculate the transit S/N using Eq. 4. The mapping between MES and S/N is shown in Fig. 4 for the full set of injected planets whose transit S/N values span 2.7-4843. Given the large number of injected planetary signals ($> 10^4$), we fit the number-weighted S/N to MES mapping using the `scipy.curve_fit` non-linear least squares algorithm with a powerlaw function of the form $\mathrm{MES} = A \cdot \mathrm{S/N}^{\alpha}$. We find a best-fit amplitude and powerlaw index of $A = 0.977$ and $\alpha = 0.967$ respectively with negligible uncertainties. This relation is used to map the transit S/N to MES which is then mapped to the detection sensitivity. The average *Kepler* detection sensitivity curve as a function of transit S/N, along with the 16^{th} and 84^{th} percentiles for the stars in our *Kepler* sample are shown in Fig. 5.

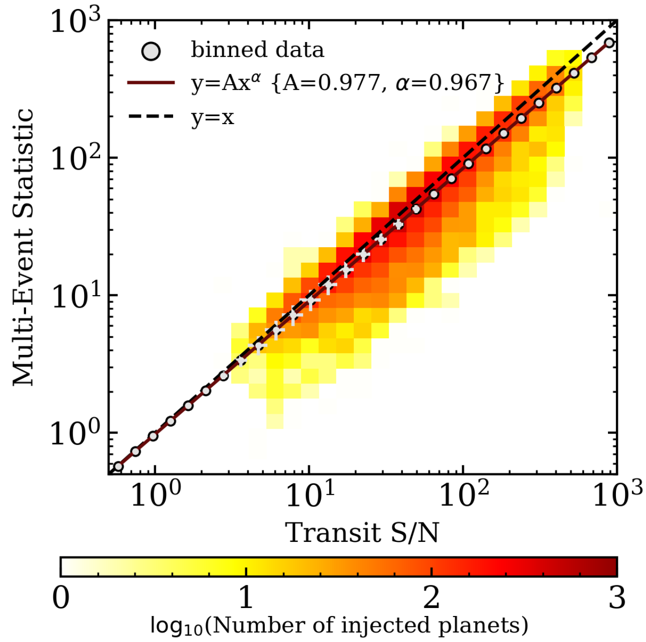

FIG. 4.— Correlation between the *Kepler* multi-event statistic and transit S/N. The mapping between the MES and S/N based on the synthetic planetary signals injected into the *Kepler* pipeline (Christiansen et al. 2015). The number-weighted powerlaw fit (*solid line*) to the correlation differs slightly from a one-to-one relation (*dashed line*) with a marginally lower amplitude and shallower slope.

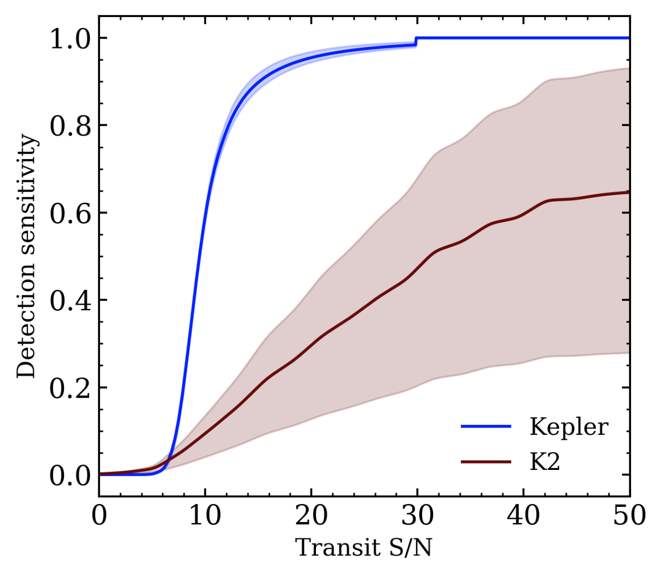

FIG. 5.— Average detection sensitivity for *Kepler* and *K2*. The *solid curves* represent the average transiting planet detection sensitivity for the *Kepler* and *K2* stars in our sample as a function of the transit S/N (Eq. 4). The shaded regions mark the 16^{th} and 84^{th} percentiles of the measured detection sensitivities.

4.2. *K2 Sensitivity*

Unlike the primary *Kepler* mission, the *K2* data products do not feature detailed completeness and reliability products. To derive the detection sensitivity among the *K2* stars in our sample we employ the transit detection pipeline `ORION` (Cloutier 2019).

The failure of the second reaction wheel on board the *Kepler* spacecraft in 2013 prevented the observatory from maintaining the fine pointing accuracy required to continue to obtain ultra precise photometry. The repurposed *K2* mission exploited the solar wind pressure by enabling the observatory to continue pointing along the ecliptic plane with realignments via thruster firings (Howell et al. 2014). `ORION` does not feature a specialized module to correct for the temporally correlated pointing corrections. This requires that pointing-corrected light curves be used as input. We adopt the `EVEREST`-reduced *K2* light curves which use a pixel level decorrelation to remove systematics from the spacecraft's variable pointing (Luger et al. 2016, 2018). We favor the `EVEREST` *K2* light curves over light curves produced by analogous pipelines (e.g. `K2SFF`; Vanderburg & Johnson 2014, `K2SC`; Aigrain et al. 2015, 2016) due to its demonstrated performance in obtaining improved photometric precision by a factor of $\sim 20-50\%$ (Luger et al. 2016).

We quantify the *K2* detection sensitivity using `ORION` by first retrieving the `EVEREST` light curve from MAST for each star in our sample. We only consider light curves from individual campaigns. As `ORION` input we supply the time sampling $\mathbf{t}$ in BJD, the corrected flux, and flux uncertainties in e^-/second, from the `EVEREST` keywords

TIME, FCOR, and FRAW_ERR. The duty cycle is derived by restricting to light curve measurements for which the QUALITY flag is zero. In light curves with known signals from planets or planet candidates, those signals are modele and removed from the light curve based on their reported transit parameters and using the batman (Kreidberg 2015) implementation of the Mandel & Agol (2002) transit model. We then inject transiting planetary signals directly into the light curve by sampling planets from the linear transit S/N grid $\mathcal{U}(0, 50)$. The per-system multiplicity is drawn from the cumulative occurrence rate of small planets out to 200 days around mid-K to early M dwarfs from *Kepler* (2.5 ± 0.2; Dressing & Charbonneau 2015). Each planet's time of mid-transit T_0 is drawn from $\mathcal{U}(\min(\mathbf{t}), \max(\mathbf{t}))$. In a given light curve, with fixed $\mathbf{t}$ and CDPP$_D$, for a star whose R_s and M_s values are fixed to their maximum likelihood values, we draw each planet's logarithmic orbital period from $\mathcal{U}(\log_{10}(0.5 \text{ days}), \log_{10}(200 \text{ days}))$ which allows us to compute the number of transits that occur within $\mathbf{t}$. Note that some injected planets will exhibit $n_{\text{transits}} = 0$ due to the limited *K2* baselines of typically ~ 80 days. The drawn orbital period also uniquely determines the planet's radius corresponding to its drawn value of the S/N. To ensure dynamical stability in multi-planet systems, we compute the maximum likelihood planet mass from the probabilistic mass-radius relation forecaster (Chen & Kipping 2017) and analytically assess the Lagrange stability of each neighboring planet pair assuming circular orbits (Barnes & Greenberg 2006). Each planet's scaled semimajor axis a/R_s and scaled radius r_p/R_s follow from their sampled radius r_p and the stellar parameters R_s and M_s. We sample impact parameters b from $\mathcal{U}(0, 0.9)$ to compute the orbital inclinations. Furthermore, we adopt fixed quadratic limb darkening coefficients by interpolating the *Kepler* bandpass coefficient grid along T_{eff}, $\log g$, and [Fe/H], assuming solar metallicity when [Fe/H] measurements are absent (Claret et al. 2012). These parameters are used to compute transit models in the absence of any transit timing variations. Transit signals are then injected into the cleaned *K2* light curves and fed to ORION to conduct a blind search for transiting signals.

The detection sensitivity as a function of S/N for each *K2* star is computed by considering 10^2 injected planetary systems per star and computing the recovery fraction of injected planets with $P \leq 100$ days. The average *K2* detection sensitivity curve, along with the 16$^{\text{th}}$ and 84$^{\text{th}}$ percentiles, are also included in Fig. 5. The quality of the pointing corrections within the EVEREST light curves can vary widely within our sample such that there is considerably more variance in the *K2* detection sensitivity relative to *Kepler*. Furthermore, the average detection sensitivity is significantly reduced compared to *Kepler*. The reduced sensitivity is due in-part to the imperfect corrections of the reduced pointing accuracy and to the limited time baseline of ~ 80 days in a typical *K2* light curve compared to *Kepler*. Furthermore, we have not attempted to optimize the performance of ORION on *K2* light curves beyond slight modifications to the algorithm's performance hyperparameters that were made to ensure the detection of 52/53 confirmed *K2* planets. The planet K2-21c (EPIC 206011691.02, $P = 15.5$ days) remains undetected by ORION because of the algorithm's requirement to discard putative signals that are commensurate with other high S/N signals in the light curve. The presence of K2-21b at $P = 9.32$ days is within 1% of a 5:3 period ratio with K2-21c and thus prohibits the identification of the 15.5 day signal as being independent and planetary.

4.3. *Two-dimensional sensitivity maps*

The sensitivity curves depicted in Fig. 5 enable us to extend the visualization of the detection sensitivity to two dimensions. Explicitly, we consider the detection sensitivity s_{nij} for each star (indexed by n) and as a function of P and r_p which are indexed by i and j respectively. Consideration of the sensitivity in $P - r_p$ space is needed to evaluate the occurrence rates in that parameter space and ultimately for understanding the structure of the radius valley around low mass stars due to the dependence of the efficiency of atmospheric loss on both planet size and separation, regardless of the physical mechanism involved.

We consider orbital periods $P \in [0.5, 100]$ days and planet radii $r_p \in [0.5, 4]$ R$_\oplus$. At each grid cell nij we compute the average S/N within the cell and map that value to the detection sensitivity using the data in Fig. 5. The detection sensitivity maps for *Kepler* and *K2*, averaged over the index n, are shown in Fig. 6.

4.4. *Survey Completeness*

Only transiting planets are detectable in transit surveys. To correct for the non-detection of otherwise detectable but non-transiting planets we compute the geometric transit probability for each star n and at each grid cell ij in the $P - r_p$ space to be

$$p_{t,nij} = \frac{R_{s,n} + r_{p,j}}{a_{ni}}. \tag{5}$$

Note that we are only interested in the relative planet occurrence rate and therefore do not consider constant scalar modifications to $p_{t,nij}$ from effects such as grazing transits or non-zero eccentricities (Barnes 2007).

The product of each star's detection sensitivity with its geometric transit probability yield completeness maps as a function of P and r_p. The average completeness maps for our *Kepler* and *K2* stars are shown in Fig. 7.

5. THE OCCURRENCE RATE OF SMALL CLOSE-IN PLANETS AROUND LOW MASS DWARF STARS

5.1. *Occurrence rates versus orbital period and planet radius*

The detection and validation of planets from the *Kepler* and *K2* missions enables the measurement of the occurrence rate of planets given the completeness corrections derived in Sect. 4. For the index i representing a planet's orbital period and j representing the planetary radius, the probability of detecting an integer number of planets within that grid cell (k_{ij}) around N_s stars is given by the binomial likelihood function

$$\mathcal{L}_{nij}(k_{ij}|N_s, P_{nij}) = \binom{N_s}{k_{ij}} \prod_{n=1}^{N_s} P_{nij}^{k_{ij}} (1 - P_{nij})^{N_s - k_{ij}} \tag{6}$$

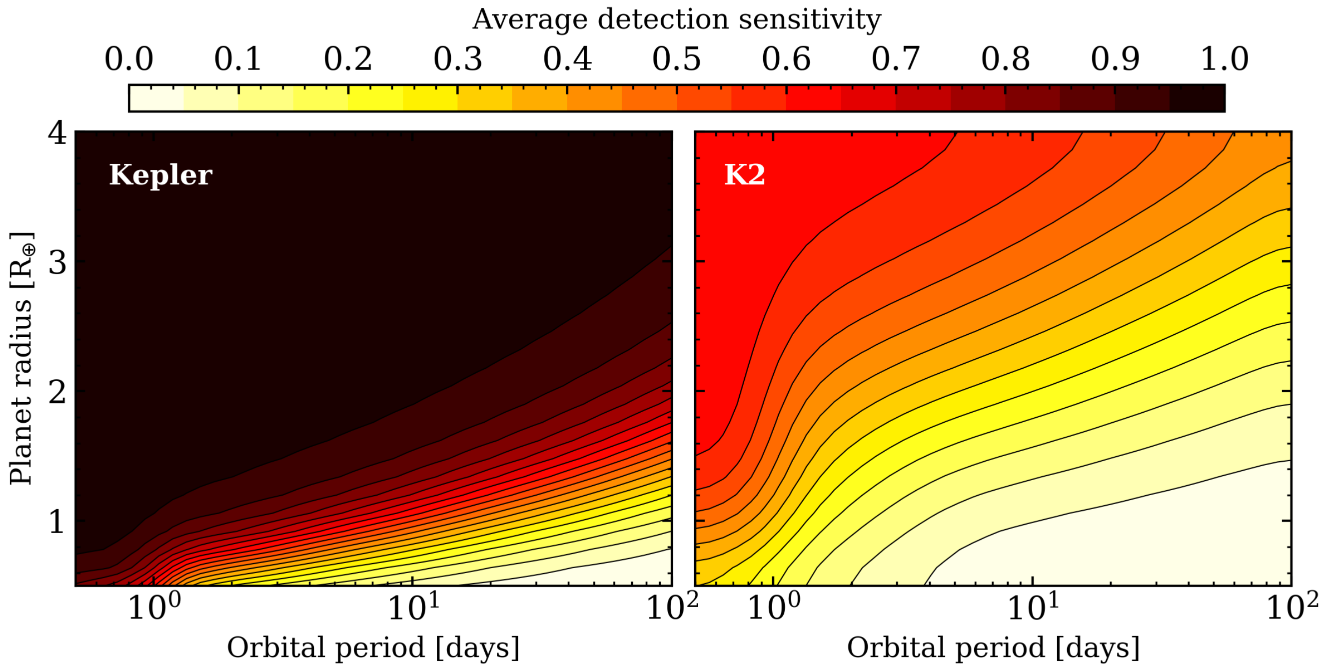

FIG. 6.— Average detection sensitivity versus orbital period and planetary radius. The detection sensitivity maps averaged over *Kepler* stars (*left panel*) and over *K2* stars (*right panel*) from our sample of low mass dwarf stars.

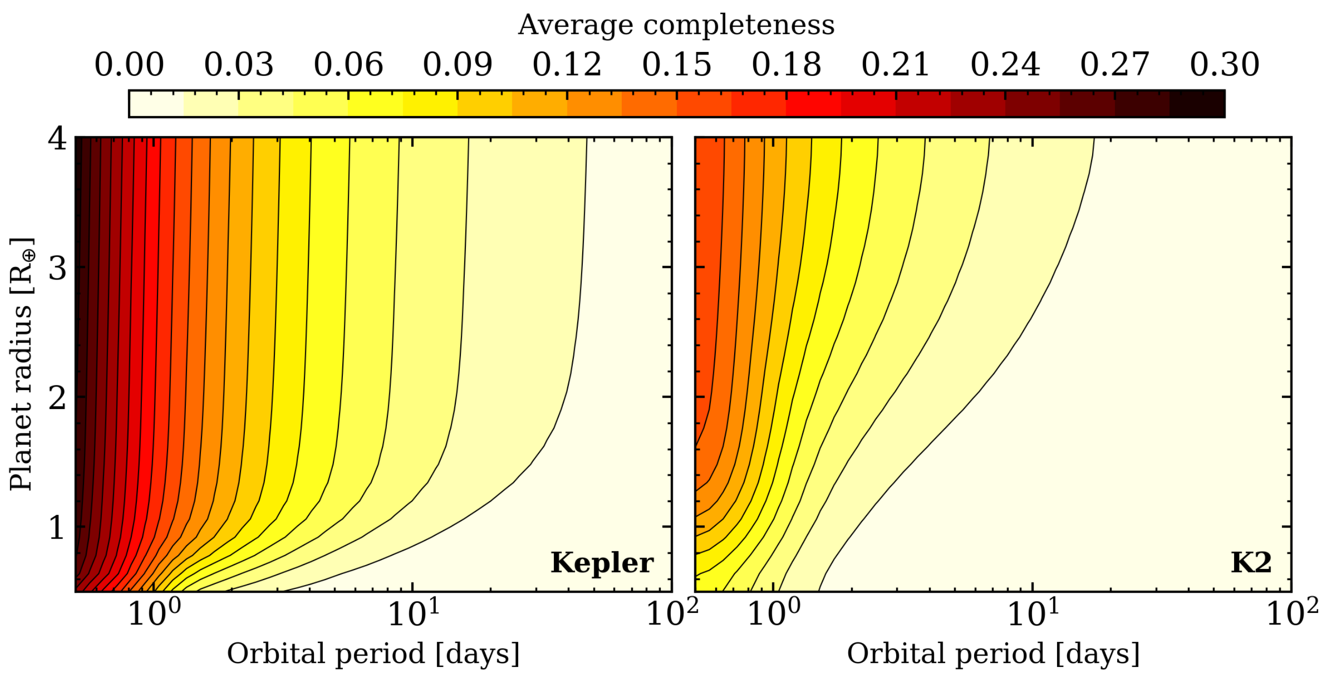

FIG. 7.— Average completeness versus orbital period and planetary radius. Maps of the product of the detection sensitivity and geometric transit probability averaged over *Kepler* stars (*left panel*) and over *K2* stars (*right panel*) from our sample of low mass dwarf stars.

where

$$P_{nij} = s_{nij} \cdot p_{t,nij} \cdot f_{ij}, \tag{7}$$

is the probability of detecting a planet in the ij grid cell around the n^{th} star. This quantity is dependent on the detection sensitivity s_{nij}, the transit probability $p_{t,nij}$, and the intrinsic occurrence rate of planets in the grid cell ij f_{ij} which is assumed to be common to all N_s stars. The number of planet detections k_{ij} was depicted in Fig. 3. Calculations of s_{nij} and $p_{t,nij}$ produced the completeness maps shown in Fig. 7. Taken together, and noting from Bayes theorem that the posterior probability of f_{ij} is

$$p(f_{ij}|N_s, s_{nij}, p_{t,nij}, k_{ij}) \propto \mathcal{L}_{nij}(k_{ij}|N_s, s_{nij}, p_{t,nij}, f_{ij}), \tag{8}$$

modulo the coefficient of proportionality which we set to unity, we are able to compute the maximum a-posteriori (MAP) occurrence rate and uncertainty maps according to Eq. 8.

Before proceeding, first recall that our planet sample contains ~ 5 times more confirmed planets from *Kepler* than from *K2* (see Fig. 3) despite our stellar sample including ~ 3.5 times more *K2* stars (see Fig. 2). These factors compound to produce a lower planet occurrence rate measured with *K2* than with *Kepler* as the reduced *K2* detection completeness (see Fig. 7) is insufficient to account for the lower measured planet occurrence rates. The discrepancy instead arises from the disparate resources that have been dedicated to the confirmation of planets from *Kepler* and *K2*. The result being that the number of confirmed planets existing within the set of *K2* planet candidates is underestimated by the number of planet candidates that have been reported as validated to date. We address this discrepancy by scaling the cumulative occurrence rate measured by *K2* to that of *Kepler*. In this way, we are assuming that the planet populations studied by each mission are inherently identical despite existing within distinct stellar populations within the galaxy.

The MAP f_{ij} map is depicted in Fig. 8. Here the existence of the radius valley around low mass stars is clearly evident. Distinct peaks in the planet frequency are separated along the planetary radius axis and span $\sim 0.9-1.4$ R$_\oplus$ and $\sim 1.9-2.3$ R$_\oplus$ respectively. Note however that the lower limit on the former peak approaches the region in which the *Kepler* sensitivity falls below 10% and the f values become unreliable. The occurrence rates also highlight the relative dearth of planets larger than ~ 3 R$_\oplus$ including the Neptunian desert at short orbital periods (Lundkvist et al. 2016; Mazeh et al. 2016). The large scale structure of the measured occurrence rates are also broadly consistent with previous investigations of the planet population around low mass *Kepler* stars (Morton & Swift 2014; Dressing & Charbonneau 2015; Gaidos et al. 2016) such as the prominence of planets $\lesssim 2$ R$_\oplus$ with $P \sim 10-60$ days and a measured cumulative occurrence rate of 2.83 ± 0.36 planets per star.

The location and slope of the radius valley (i.e. $\mathrm{d}\log r_p/\mathrm{d}\log P$) appear broadly consistent with the valley structure measured from the empirical planet population of FGK stars characterized via asteroseismology (Van Eylen et al. 2018). Wu (2019) also provided a visual approximation to the location of the radius valley around stars with $M_s \in [0.5, 0.76]$ M$_\odot$ in their *Gaia-Kepler* sample. However we find the location of the terrestrial-sized planet peak to exist at longer $P \sim 30$ days compared its location at ~ 5 days from Wu (2019) (c.f. Fig. 2). The discrepancy likely originates from differences in the method of correcting for survey incompleteness. Recall that in this study the detection sensitivity for *Kepler* stars is computed on a per star basis given the unique completeness products from the *Kepler* pipeline whereas Wu (2019) adopt the piecewise completeness levels of 10, 50, or 90% complete as a function of P and r_p from Zhu et al. (2018).

Also included in Fig. 8 are planets with $\geq 3\sigma$ bulk density measurements from either precision radial velocities or transit timing variations. Planet parameters are retrieved from the NASA Exoplanet Archive for planets orbiting stars with $T_{\rm eff} \in [2800, 4700]$ K, whose orbital periods and radii span the domain considered in Fig. 8, and whose masses are inconsistent with zero (i.e. no mass upper limits). The properties of the resulting 20 planets are listed in Table 3. Based on the planetary bulk densities ρ_p we define the following composition dispositions: terrestrial dispositions have $\rho_p + \sigma_{\rho_p} \geq 5.5$ g cm^{-3} and $\rho_p - \sigma_{\rho_p} \geq 3.5$ g cm^{-3}, gaseous dispositions have $\rho_p - \sigma_{\rho_p} \leq 2$ g cm^{-3} and $\rho_p + \sigma_{\rho_p} \leq 3.5$ g cm^{-3}, with all remaining planets being flagged as having an ambiguous bulk composition (i.e. not clearly terrestrial-like or hosting a gaseous envelope).

The vast majority of the retrieved planets in Fig. 8 demonstrate clear composition clustering with planet radius. With the exception of the probable terrestrial planet GJ 143b ($P = 35.6$ days, $r_p = 2.6$ R$_\oplus$), all terrestrial planets appear to be smaller than 1.8 R$_\oplus$, independently of orbital period. Similarly, all four gaseous planets are larger than 2.6 R$_\oplus$ while the three remaining planets with intermediate radii correspond to those with ambiguous bulk compositions. Thus we are justified in classifying the occurrence rate peak spanning $\sim 0.9-1.4$ R$_\oplus$ as representing terrestrial planets. The second peak between $\sim 1.9-2.3$ R$_\oplus$ hosts planets with ambiguous bulk compositions but are clearly inconsistent with being terrstrial such that we will refer to this feature as the gaseous peak in what follows.

Lastly, we note that the radius valley as a function of P is not completely void of planets. This may present evidence of the efficiency of the gap clearing mechanism around low mass stars and is discussed further in Sect. 7.3.

5.2. *Occurrence rates versus planet radius*

Next we marginalize over P and compute the one-dimensional occurrence rate of small, close-in planets as a function of r_p. The resulting occurrence rates are shown in Fig. 9 in which the bimodal distribution of planet sizes is again clearly discernible in the MAP occurrence rates. The uncertainties on each f_j bin are computed from the 16$^{\rm th}$ and 84$^{\rm th}$ percentiles of the f_j posterior. In Fig. 9 we ignore the measured occurrence rate in bins with $r_p \lesssim 1$ R$_\oplus$ where our detection sensitivity is poor.

From the bimodal distribution we highlight the approximate radii likely corresponding to planets with terrestrial bulk compositions ($r_p \lesssim 1.55$ R$_\oplus$) versus planets with significant size fractions in a gaseous envelope ($r_p \gtrsim 1.55$ R$_\oplus$) around low mass stars. Also depicted in Fig. 9 is f_j with a binning scale twice that of the primary f_j depiction (i.e. 0.06 R$_\oplus$ compared to 0.13 R$_\oplus$). With finer binning the fractional uncertainties on f_j are sufficiently large to eliminate the significance of the distinct bimodal peaks. Despite this, the bimodality in the MAP occurrence rate continues to persist with the location of the valley features only being marginally affected. We interpret this as further evidence for the existence of the radius valley in the close-in planet population around low mass stars.

5.3. *Inclusion of supplemental K2 planet candidates*

In an attempt to improve the counting statistics in the occurrence rate calculations, we consider an enlarged planet sample. This population is the union of our exist-

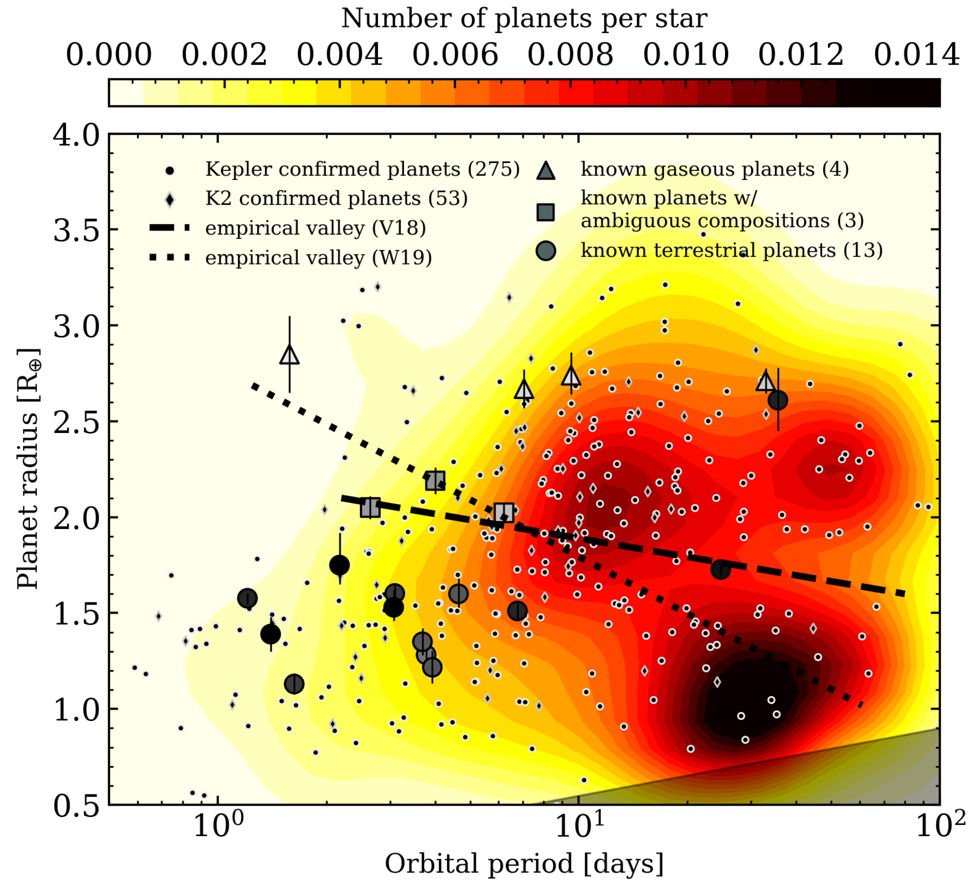

FIG. 8.— Planet occurrence rate versus orbital period and planetary radius. The maximum a-posteriori occurrence rate map calculated from the population of confirmed planets from *Kepler* (*circles*) and *K2* (*diamonds*) around low mass dwarf stars. Overplotted are the empirical locations of the radius valley around FGK stars characterized via asteroseismology (*dashed line*, Van Eylen et al. 2018) and the approximate radius valley around early-M to mid-K dwarfs (*dotted line*, Wu 2019). Also overplotted are planets with $\geq 3\sigma$ bulk density measurements from the literature that are classified as having either a terrestrial (*circles*), a gaseous (*triangles*), or an ambiguous (*squares*) bulk composition.

ing sample of confirmed planets with a set of additional planet candidates (PCs) from the *K2* mission. Specifically, we consider the set of PCs reported by Kruse et al. (2019) from *K2* campaigns 0-8 that includes 126 PCs not already included in our sample of confirmed planets and orbiting stars contained within our stellar sample.

By definition, we cannot identify which PCs are true planets of interest for this study and which PCs are instead produced by an astrophysical false positive. The inclusion of *K2* PCs therefore requires that we account for sample contamination by false positives probabilistically. We do so by considering a number of studies from the literature that perform a transiting planet search in *K2*, from any subset of its campaigns, and attempt to validate their uncovered PCs statistically using on follow-up observations (Montet et al. 2015; Crossfield et al. 2016; Dressing et al. 2017; Hirano et al. 2018; Livingston et al. 2018; Mayo et al. 2018). Each of these studies utilized some combination of ground-based photometry to validate planet ephemerides, reconnaissance spectroscopy to identify spectroscopic binaries, and speckle or AO-assisted imaging to search for nearby stellar companions to their PC host stars. Each of the aforementioned studies used their respective set of follow-up observations together with the statistical validation tool `vespa` (Morton 2012, 2015) to statistically classify their PCs as either a validated planet (VP)[6], a false positive (FP), or some other inconclusive disposition (e.g. remains a PC). The FP rate around cool stars ($T_{\rm eff} < 4700$ K) from each study is estimated by calculating the ratio of the number of reported FPs to the total number of FPs and VPs. Notably, Crossfield et al. (2016) showed that the FP rate in their *K2* sample is dependent on the measured planet radius as giant PCs have a larger likelihood of being a FP. Hence, we focus on PCs with $r_p < 4$ R$_\oplus$ when deriving FP rates.

The resulting FP rates are reported in Table 4. Half of the studies do not find any probable FP signals among the small PCs orbiting cool stars in their samples. In such cases, only upper limits on the FP rate can be derived which all agree that the FP rate is $\lesssim 20\%$ at 95%. The remaining studies each detect at least one FP such that a non-zero maximum likelihood FP rate is measured. Their average FP rate is 5.7% which is also in agreement with the derived upper limits from the aforementioned studies. We proceed by constructing 10^3 realizations of the planet population that includes all confirmed planets from both *Kepler* and *K2* plus a subset of the 126 *K2* PCs from Kruse et al. (2019). The subset of included

[6] We treat validated and confirmed planets as equivalent dispositions.

TABLE 3
Planets with $\geq 3\sigma$ bulk density measurements around low mass stars

Planet name	P [days]	F [$F_\oplus$]	r_p [$R_\oplus$]	m_p [$M_\oplus$]	ρ_p [g cm^{-3}]	Method	Composition disposition	Refs.
GJ 143b	35.61253	5.5	2.61	22.70	7.09	RV	Ter	1
GJ 357b	3.93072	13.2	1.22	1.84	5.67	RV	Ter	2
GJ 1132b	1.62892	19.4	1.13	1.66	6.39	RV	Ter	3,4
GJ 1214b	1.58040	22.2	2.85	6.26	1.50	RV	Gas	5
GJ 4332b	1.40150	144.2	1.39	4.60	9.51	RV	Ter	6
GJ 9827b	1.20898	326.9	1.58	4.91	6.95	RV	Ter	7
GJ 9827d	6.20147	37.0	2.02	4.04	2.71	RV	Am	7
HD 219134b	3.09293	176.7	1.60	4.74	6.40	RV	Ter	8
HD 219134c	6.76458	62.3	1.51	4.36	7.01	RV	Ter	8
K2-18b	32.93962	1.2	2.71	8.63	2.40	RV	Gas	9,10
K2-146b	2.64460	19.2	2.05	5.77	3.72	TTV	Am	11
K2-146c	4.00498	11.0	2.19	7.49	3.96	TTV	Am	11
K2-216b	2.17480	1837.4	1.75	8.00	8.28	RV	Ter	12
Kepler-80b	7.05246	41.8	2.67	6.93	2.02	TTV	Gas	13
Kepler-80c	9.52355	28.0	2.74	6.74	1.82	TTV	Gas	13
Kepler-80d	3.07222	126.5	1.53	6.75	10.46	TTV	Ter	13
Kepler-80e	4.64489	72.9	1.60	4.13	5.60	TTV	Ter	13
L 98-59c	3.69040	11.9	1.35	2.46	5.55	RV	Ter	14,15
LHS 1140b	24.73696	0.5	1.73	6.98	7.52	RV	Ter	16
LHS 1140c	3.77793	6.1	1.28	1.81	4.77	RV	Ter	16

Note. — References: 1) Dragomir et al. 2019 2) Luque et al. 2019 3) Dittmann et al. 2017 4) Bonfils et al. 2018 5) Harpsøe et al. 2013 6) Astudillo-Defru et al. in prep. 7) Rice et al. 2019 8) Gillon et al. 2017 9) Benneke et al. 2017 10) Cloutier et al. 2019b 11) Hamann et al. 2019 12) Persson et al. 2018 13) MacDonald et al. 2016 14) Kostov et al. 2019 15) Cloutier et al. 2019a 16) Ment et al. 2019.

TABLE 4
K2 false positive rates for small planets around cool stars

Reference	$N_{\rm FP}$	$N_{\rm VP}$	FP rate [%]
Montet et al. (2015)[a]	0	8	< 30.7
Crossfield et al. (2016)	2	39	$4.9^{+6.0}_{-1.4}$
Dressing et al. (2017)	2	34	$5.6^{+6.4}_{-2.0}$
Hirano et al. (2018)[a]	0	16	< 19.5
Livingston et al. (2018)[a]	0	14	< 21.0
Mayo et al. (2018)[b]	1	14	$6.7^{+12.4}_{-2.0}$

Note. — Within each study we only consider PCs with $r_p < 4$ $R_\oplus$ and orbiting cool stars with $T_{\rm eff}$ < 4700 K. FP: false positive. VP: validated planet.
[a] These studies do not detect any FPs such that the reported FP rate upper limit is represented by its 95% confidence interval.
[b] Mayo et al. (2018) did not explicitly classify their non-validated planets as FPs so we define FPs within their sample as any PC whose false positive probability exceeds 10%.

PCs are randomly sampled from the full set of PCs according to the adopted FP rate such that each realization contains $0.943 \cdot 126 \approx 119$ PCs.

The effect of including PCs on the derived occurrence rates is assessed by comparing the f_j distributions measured with and without the inclusion of PCs. The results are depicted in Fig. 10. Again we scale the *K2* occurrence rates to those from *Kepler* such that the cumulative occurrence of close-in planets with $r_p \leq 4$ $R_\oplus$ is identical between the two planet populations as their contributions from *Kepler* planets are identical. The radius valley continues to be resolved in the MAP occurrence rates. Furthermore, the addition of PCs reduces the median f_j uncertainty among planets with $r_p > 1$ $R_\oplus$ from 0.0216 to 0.0186 planets per star (i.e. $\sim 15\%$ improvement). However, the partial filling of the gap is more substantial as the contrast between the maximum f_j of the terrestrial planet peak ($r_p \sim 1.3$ $R_\oplus$) and the minimum f_j of the valley ($r_p \sim 1.6$ $R_\oplus$) decreases from 0.070 to 0.054 $R_\oplus$ (i.e. $3.2\sigma \to 2.9\sigma$).

6. EVOLUTION OF THE RADIUS VALLEY AROUND LOW MASS STARS

6.1. *Slope of the radius valley*

Fig. 11 shows the two-dimensional occurrence rate of planets in the $F - r_p$ space for our planet sample as well as for the close-in *Kepler* planets around Sun-like stars from Martinez et al. (2019). In this parameter space, we calculate the slope of the radius valley with F and compare the measured value to model predictions of the transition from terrestrial to gaseous planets versus insolation. We measure the slope by resampling 10^3 planet populations from the inverse occurrence rates (and their fractional uncertainties at each point ij) over the domains $F \in [1, 30]$ $F_\oplus$ and $r_p \in [1, 2.5]$ $R_\oplus$. Each domain is defined to neglect regions far from the radius valley which would otherwise dominate the inverse occurrence rates. The number of sampled planets in the $\log F - r_p$ space is then fit with a linear function as depicted in Fig. 11. Over the 10^3 realizations of the resampled planet population, we measure an average slope and standard deviation of $\mathrm{d}r_p/\mathrm{d}\log F = -0.123 \pm 0.054$. Similarly, repeating this exercise in the $P - r_p$ space yields $\mathrm{d}r_p/\mathrm{d}\log P = 0.085 \pm 0.074$.

The negative slope of $\mathrm{d}r_p/\mathrm{d}\log F = -0.123$ indicates that the location of the radius valley drops to smaller planet radii with increasing insolation (i.e. at smaller orbital separations). **This behavior is broadly consistent with models of the formation small terrestrial planets in a gas-poor environment (Lee**

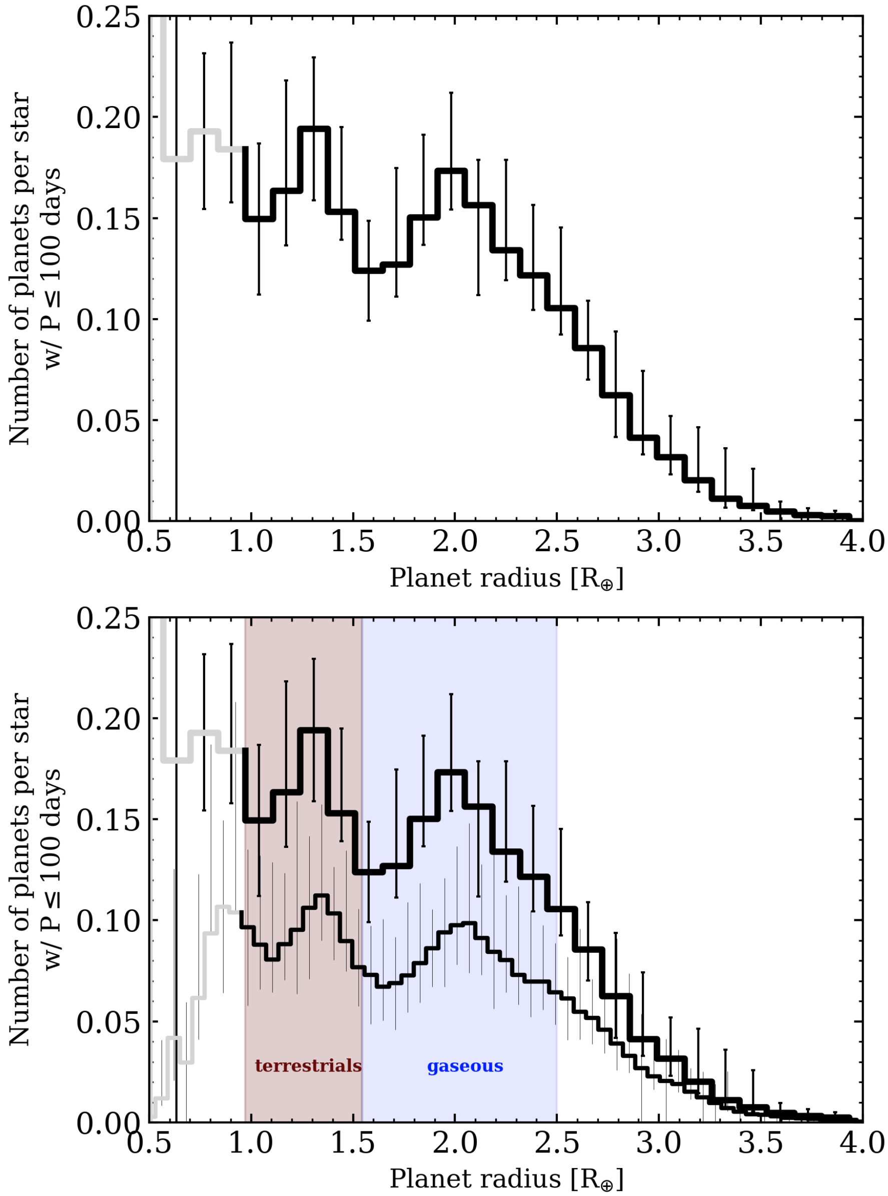

FIG. 9.— Occurrence rate of planets as a function of size. *Upper panel*: histogram depicting the relative occurrence rate of close-in planets with orbital periods < 100 days derived from the population of confirmed *Kepler* and *K2* planets around low mass stars. The bimodal distribution of planet radii peaking at occurrence rate-weighted radii of 1.12 and 2.07 $R_\oplus$, highlights the presence of the radius valley around low mass stars centered at 1.54 $R_\oplus$. Uncertainties in the planet occurrences follow from binomial statistics and are limited by the relatively small number of confirmed planets around low mass stars from *Kepler* and *K2*. The measured occurrence rates below ~ 1 $R_\oplus$ (shown in *grey*) should be ignored due to poor detection sensitivity. *Lower panel*: identical occurrence rates as in the upper panel accompanied by the same occurrence rates with finer radius bins. The corresponding occurrence rate uncertainties are too large to robustly infer the presence of features but the bimodal structure continues to be exhibited in the maximum likelihood occurrence rates. The shaded regions highlight our approximate definitions of terrestrial planets ($r_p \in [0.97, 1.54]$ $R_\oplus$), down to reasonable sensitivity limits, and gaseous planets ($r_p > 1.57$ $R_\oplus$) around low mass stars. Note the 2.5 $R_\oplus$ outer limit of the shaded region is arbitrary.

et al. 2014; Lee & Chiang 2016; Lopez & Rice 2018) leading to the transition from terrestrial to gaseous planets emerging due to the superposition of the two planet populations whose formation timescales differ. In this scenario, the transition radius as a function of orbital separation is set by the maximum mass of a bare rocky core which itself is set by the amount of available solid material for the proto-planet to form out of via collisional growth. According to the minimum-mass extrasolar nebula (Chiang & Laughlin 2013), the solid surface density radial profile is $\sigma_{\mathrm{solid}} \propto a^{-1.6}$, where a is the semimajor axis. The amount of solid material accreted by a proto-planet is proportional to its Hill radius $r_H = a(m_p/3M_s)^{1/3}$ such that integrating over the disk surface results in the maximum mass of a bare rocky core $m_{p,\mathrm{max}} \propto a^{0.6} M_s^{-0.5}$, or $r_{p,\mathrm{valley}} \propto a^{0.16} M_s^{-0.14}$ after applying the rocky planet mass-radius relation (Zeng et al. 2016). Thus the transition radius is predicted to occur at larger planet radii with increasing a for a given host spectral

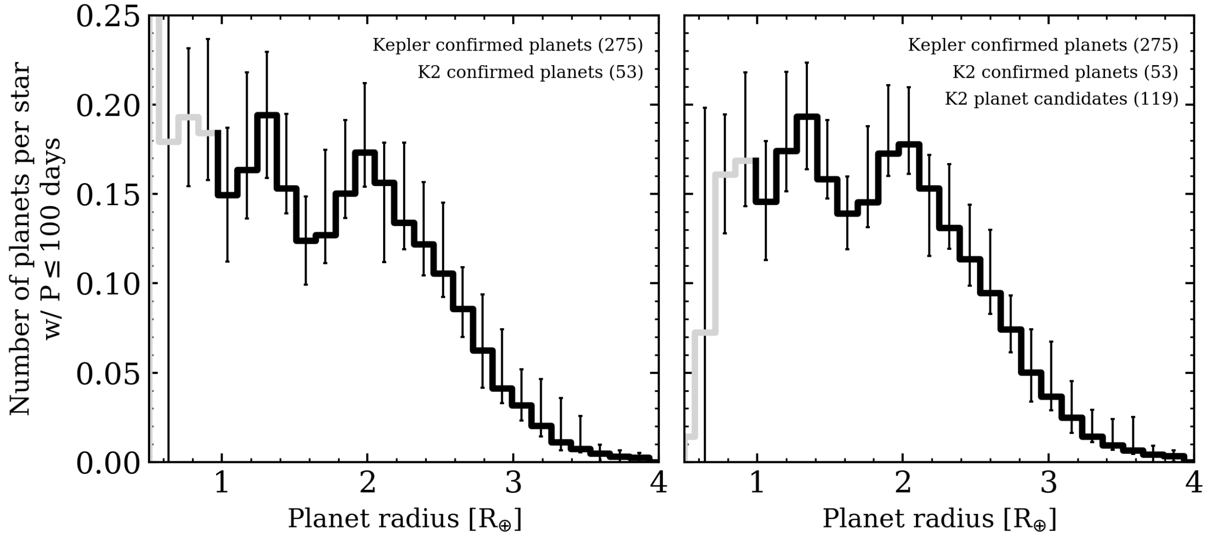

FIG. 10.— Comparison of occurrence rates with and without planet candidates included. *Left panel*: same as Fig. 9. *Right panel*: histogram depicting the relative occurrence rate of close-in planets with orbital periods < 100 days derived from the population of confirmed planets from *Kepler* and *K2* and supplemented by 119 PCs around low mass stars from Kruse et al. (2019). The radius valley continues to be resolved with the inclusion of PCs which improve the median uncertainty on the occurrence rate bins although the gap becomes less prominent with numerous PCs partially filling the valley.

type (Lopez & Rice 2018). The theoretical scaling of the transition radius with insolation for a given spectral type is $r_{p,\mathrm{valley}} \propto F^{-0.08}$ (Lopez & Rice 2018) which is consistent with our measured scaling between $r_{p,\mathrm{valley}}$ and $\log F$ of -0.123 ± 0.054. **Furthermore, our slope measurements are inconsistent with models of thermally-driven atmospheric escape from photoevaporation or core-powered mass loss that predict an increasing transition radius with increasing insolation ($r_{p,\mathrm{valley}} \propto F^{0.11}$; Lopez & Rice 2018, $r_{p,valley} \propto F^{0.10}$; Gupta & Schlichting 2019b respectively).** The negative slope in the radius valley around low mass stars differs from the trend seen around Sun-like stars (Fulton et al. 2017; Van Eylen et al. 2018; Martinez et al. 2019). We interpret these differing observational signatures as an elucidation to possible changes in the dominant physical mechanisms that sculpt the radius valley around different host spectral types.

The inclusion of planets with $\geq 3\sigma$ bulk density measurements (see Table 3) in Fig. 11 reveals that all planets that are inconsistent with having bulk terrestrial compositions lie above the transition radius predictions from all physical models considered. However, the temperate terrestrial planet LHS 1140b ($F = 0.5$ F$_\oplus$, $r_p = 1.7$ R$_\oplus$) sits in the $F - r_p$ parameter space above the predicted transition radius from photoevaporation and core-powered mass loss but below the predicted transition from gas-poor formation. Although LHS 1140b is the only instance of a planet existing between the radius valley predictions from gas-poor formation and thermally-driven mass loss in Fig. 11, its location and composition provide supporting evidence for the applicability of gas-poor terrestrial planet formation to the emergence of the radius valley around low mass stars.

6.2. *Planet populations versus stellar mass*

In addition to calculating the occurrence rates f_{ij} among our full stellar sample, we also consider the evolution of the planet population in unique host stellar mass bins. Fig. 12 shows the MAP f_{ij} maps in $P - r_p$ space, and the marginalized f_j distributions, in four stellar mass bins representing our full stellar sample ($M_s \in [0.08, 0.93]$ M$_\odot$), the massive half of the sample ($M_s > 0.65$ M$_\odot$), the low mass half of the sample ($M_s < 0.65$ M$_\odot$), and a subset of the latter focusing on increasingly lower mass stars ($M_s < 0.42$ M$_\odot$). The statistically significant resolution of the radius valley in the f_j occurrence rates is only accomplished with the full stellar sample. The reduction of the sample size in the three remaining mass bins inflates the f_j uncertainties such that the valley is observed at $< 1\sigma$ and hence not significant. However, the characteristic bimodality is exhibited in the MAP f_{ij} of the full and massive half stellar samples. Furthermore, their f_{ij} structures are similar as the majority of our full planet sample orbit stars more massive than the median stellar mass of 0.65 M$_\odot$ (i.e. $\sim 62\%$ of our confirmed planet sample).

In considering stars less massive than 0.65 M$_\odot$, the gaseous planet peak begins to diminish relative to the terrestrial-sized planets. As evidenced in the MAP f_j distribution around stars with $M_s \in [0.08, 0.65]$ M$_\odot$, the radius valley might persist around 1.6 R$_\oplus$ but the gaseous planet peak does not appear distinct from the terrestrial planet peak in the MAP f_{ij} map. That is that the relative frequency of terrestrial to gaseous planets appears to increase significantly around M dwarfs compared to the more massive K dwarfs. This feature is further accentuated around the lowest mass stars in our sample (< 0.42 M$_\odot$) for which terrestrial-sized planets clearly dominate the distribution of close-in planets. The relative frequency of terrestrial to gaseous planets in each stellar mass bin are reported in Table 5 for fixed definitions of $r_p \in [1, 1.6]$ R$_\oplus$ and $r_p \in [1.6, 2.5]$ R$_\oplus$ respectively. The inner limit of 1 R$_\oplus$ restricts this analysis to

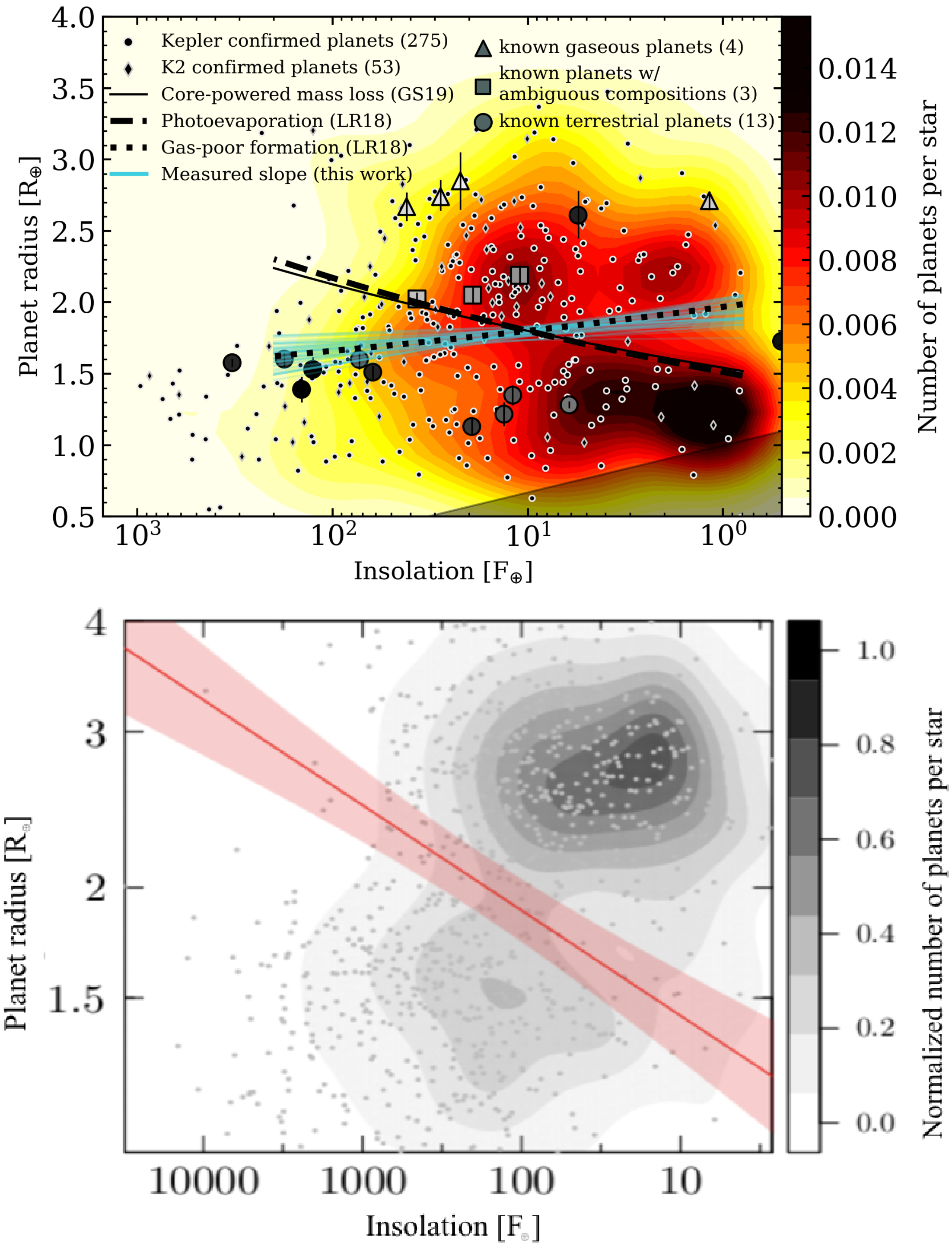

FIG. 11.— Planet occurrence rates versus insolation and planet radius around low mass and Sun-like stars. *Upper panel*: the maximum a-posteriori occurrence rate map calculated from the population of confirmed planets from *Kepler* (*circles*) and *K2* (*diamonds*) around low mass dwarf stars. Overplotted in black are model predictions of the transition from terrestrial to gaseous planets in the following scenarios: core-powered mass loss (Gupta & Schlichting 2019b), photoevaporation (Lopez & Rice 2018), and gas-poor formation Lopez & Rice (2018). We measure the slope of the radius valley to be $\mathrm{d}r_p/\mathrm{d}\log F = -0.123 \pm 0.054$ which is consistent with predictions from gas-poor formation of terrestrial planets. Also overplotted are planets with $\geq 3\sigma$ bulk density measurements from the literature that are classified as having either a terrestrial (*circles*), a gaseous (*triangles*), or an ambiguous (*squares*) bulk composition. *Lower panel*: the occurrence rate map of close-in *Kepler* planets around Sun-like stars (Martinez et al. 2019). The measured slopes for the planet populations around Sun-like and low mass stars have opposing signs. Note the unique F and r_p scales compared to the top panel.

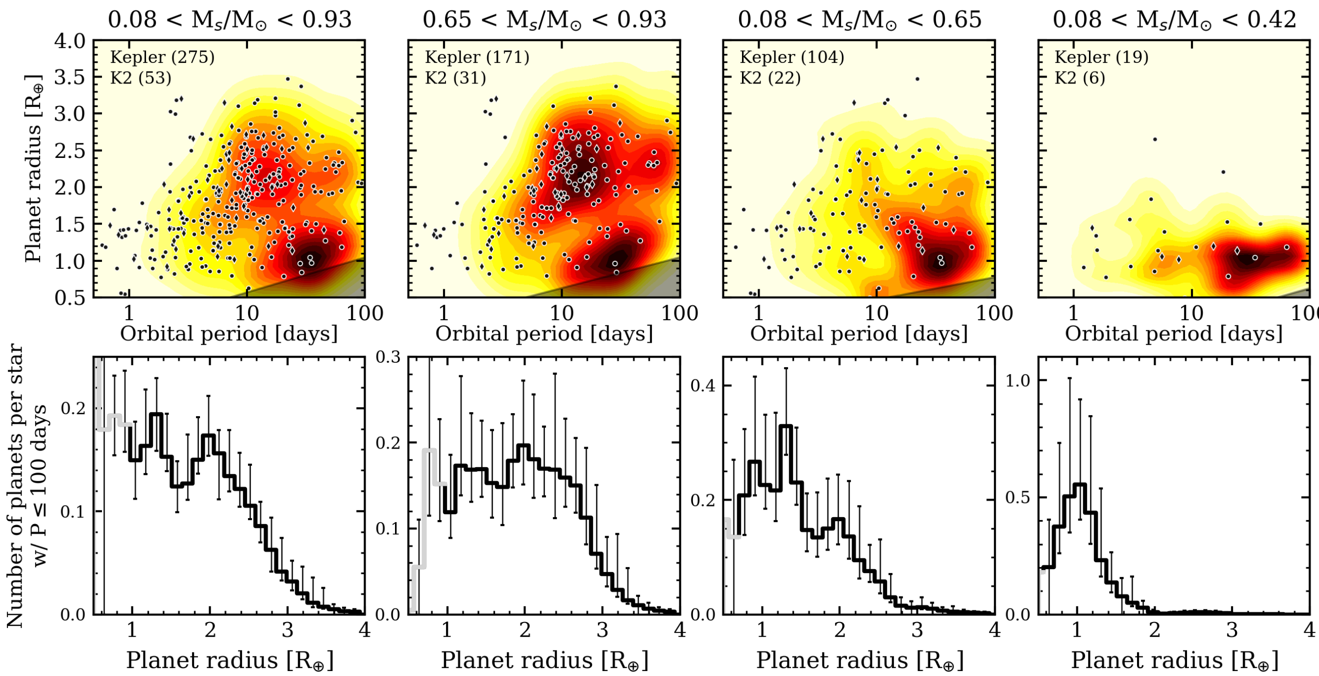

FIG. 12.— 2D and 1D planet occurrence rates in various stellar mass bins. *Top panels*: planet occurrence rate maps as a function of orbital period and planet radius. *Bottom panels*: distributions of the relative planet occurrence rate as a function of planet size. Note the differing occurrence rate scales. Each column corresponds to a unique cut in stellar masses which represent the full stellar sample ($M_s \in [0.08, 0.93]$ M$_\odot$), the early half of the stellar sample ($M_s \in [0.65, 0.93]$ M$_\odot$), the late half of the stellar sample ($M_s \in [0.08, 0.65]$ M$_\odot$), and the low mass bin ($M_s \in [0.08, 0.42]$ M$_\odot$) depicting a subset of the late half of the stellar sample. The relative occurrence of terrestrial to gaseous planets appears to increase around lower mass stars.

TABLE 5
Relative occurrence rates of close-in terrestrial and gaseous planets around low mass stars

Stellar mass range [M$_\odot$]	$f_{\rm terr}$ $r_p \in [1, 1.6]$	$f_{\rm gas}$ $r_p \in [1.6, 2.5]$	$f_{\rm terr}/f_{\rm gas}$
$[0.08, 0.90]$	0.68 ± 0.07	1.02 ± 0.08	0.66 ± 0.09
$[0.63, 0.90]$	0.69 ± 0.11	1.28 ± 0.16	0.54 ± 0.11
$[0.08, 0.63]$	1.10 ± 0.16	1.02 ± 0.16	1.08 ± 0.23
$[0.08, 0.42]$	1.64 ± 0.43	0.19 ± 0.09	8.46 ± 4.62

where the detection sensitivity is still informative. The outer limit of 2.5 R$_\oplus$ is chosen such that the full width at half maximum of the gaseous planet peak in the f_j distribution from the full stellar sample is approximately identical for each peak (Fig. 9).

The values in Table 5 indicate the significant increase in the relative occurrence of terrestrial planets with decreasing stellar mass that is illustrated in Fig. 12. Our measurements show that gaseous planets are nearly twice as common as terrestrial planets around mid to late K dwarfs ($M_s \in [0.65, 0.93]$ M$_\odot$) while the relative frequency approaches unity around the full suite of M dwarfs ($M_s \in [0.08, 0.65]$ M$_\odot$). Focusing on mid-to-late M dwarfs only in the lowest stellar mass bin considered, terrestrial planets become much more prominent as they outnumber gaseous planets by a factor of 8.46 ± 4.62. This result is broadly consistent the calculations of Hardegree-Ullman et al. (2019) who find that terrestrial-sized planets ($r_p \in [0.5, 1.5]$ R$_\oplus$) are about $4-5$ times as common as gaseous planets ($r_p \in [1.5, 2.5]$ R$_\oplus$) around M3-5.5 dwarfs ($M_s \in [0.12, 0.38]$ M$_\odot$). Our calculations provide supporting evidence for an increase in the frequency of close-in terrestrial planets around increasingly lower mass stars even with the small number of confirmed transiting planets in that mass regime. More robust statements regarding the absolute occurrence rate of terrestrial planets around mid-to-late M dwarfs will require a larger stellar sample in transit surveys with sensitivity to wider separations out to hundreds of days where giant planets begin to emerge around these stars (Bonfils et al. 2013; Morales et al. 2019).

6.3. *Dependence of radius valley features on stellar mass*

Here we measure the locations and uncertainties of features in the radius valley in each of the stellar mass bins considered in Sect. 6.2. For each stellar mass bin we measure the frequency-weighted central radius of the terrestrial planet peak, the gaseous planet peak (where applicable), and the radius valley. The uncertainties in the feature locations are largely determined by uncertainties in the measured occurrence rates but are also directly affected by the following hyperparameters: the f_{ij} smoothing parameter, the minimum detection sensitivity, the P bin width, the r_p bin width, and the imposed upper and lower P and r_p limits on each peak. The upper and lower r_p limits are defined based on the visual inspection of the f_{ij} maps in Fig. 12 and are used to demarcate the boundaries of each peak, and by extension, the valley separating the peaks. As an example, if the prescribed boundaries on the terrestrial peak are set to 1-50 days and 0.8-1.4 R$_\oplus$, then only the occurrence rates over that subset of the $P - r_p$ parameter space are used to calculate the f_j-weighted terrestrial peak radius. The range of boundary values for each peak are listed in Table 6. In practice, we derive 10^3 realizations of each f_{ij} map with each realization having a unique set of the aforementioned hyperparameters. The resulting f_{ij} maps are marginalized over P and the f_j-weighted radius of each peak is computed over the domain bounded by the relevant hyperparameters (see Table 6). The same is done for the radius valley using the inverse occurrence rates.

The resulting locations of each radius peak and valley are depicted in Fig. 13 as a function of stellar mass. The locations and uncertainties are also given explicitly in Table 7. The depicted M_s values are represented by the median stellar mass within the bin and whose uncertainties are derived from the 16$^{\rm th}$ and 84$^{\rm th}$ percentiles. In computing the feature locations we assume that the bimodal r_p distribution is present in all stellar mass bins aside from the lowest mass bin (see Fig. 12). In the lowest stellar mass bin we only measure the central radius of the terrestrial planet peak and its edge which marks the transition from terrestrial to gaseous planets despite the latter being inherently rare around these types of stars.

The measured feature radii are compared to those measured in Fulton & Petigura (2018) around Sun-like stars with $M_s < 0.97$, $\in [0.97, 1.11]$, and > 1.11 M$_\odot$. Most notably, the location of each feature measured from our full stellar sample continues the trend of monotonically decreasing towards smaller r_p with decreasing M_s. The slopes of this decrease for the terrestrial and gaseous planet peaks measured with the three points from Fulton & Petigura (2018) and the measurement from our full stellar sample are ${\rm d}r_{p,{\rm terr}}/{\rm d}M_s = 0.40$ and ${\rm d}r_{p,{\rm gas}}/{\rm d}M_s = 0.97$ respectively. The relative slopes indicate that the most common size of gaseous planet decreases more steeply with M_s than the typical size of terrestrial planets. This trend is indicative of the effective disappearance of gaseous planets around increasingly lower mass stars (see Table 5) while terrestrial-sized planets appear to persist. Furthermore, the reduced slope of the terrestrial peak may be evidence of a characteristic planetary core size of ≈ 1 R$_\oplus$ although its exact location is largely uncertain due to the limited detection sensitivity to sub-Earth-sized planets.

Models of the formation of the radius valley based upon photoevaporation (Wu 2019), gas-poor formation (Lopez & Rice 2018), and core-powered mass loss (Gupta & Schlichting 2019b) all make explicit predictions for the evolution of the radius valley location with stellar mass. Predictions from the core-powered mass loss scenario are dependent on the stellar mass-luminosity relation (MLR) $L_s \propto M_s^\alpha$. In Fig. 13, we consider cases with a constant MLR with $\alpha = 5$ (Gupta & Schlichting 2019b) and with the empirically-derived piecewise MLR from Eker et al. (2018). All models predict a decreasing radius valley with decreasing stellar mass but differ in their slopes. At the median stellar mass of our full stellar sample (0.65 M$_\odot$), the measured location of the radius valley is 1.54 ± 0.16 R$_\oplus$. This value favors a steep ${\rm d}r_{p,{\rm valley}}/{\rm d}M_s$ slope although we are unable to distinguish between competing physical models given the measurement uncertainties. Fortunately, the model predictions continue to diverge with decreasing stellar mass such that measurements of the valley location in decreasing M_s bins may be used to rule out the operation of certain physical mecha-

TABLE 6
ASSUMED BOUNDARY RANGES ON THE LOCATIONS OF RADIUS VALLEY FEATURES

Stellar mass range [$M_\odot$]	log P lower boundary [days]	log P upper boundary [days]	Terrestrial peak lower r_p boundary [$R_\oplus$]	Terrestrial peak upper r_p boundary [$R_\oplus$]	Gaseous peak lower r_p boundary [$R_\oplus$]	Gaseous peak upper r_p boundary [$R_\oplus$]
[0.08, 0.90]	$\mathcal{U}(\log 0.5, \log 2)$	$\mathcal{U}(\log 50, \log 100)$	$\mathcal{U}(0.8, 1)$	$\mathcal{U}(1.2, 1.5)$	$\mathcal{U}(1.6, 1.9)$	$\mathcal{U}(2.3, 2.5)$
[0.63, 0.90]	$\mathcal{U}(\log 0.5, \log 2)$	$\mathcal{U}(\log 50, \log 100)$	$\mathcal{U}(0.8, 1)$	$\mathcal{U}(1.3, 1.5)$	$\mathcal{U}(1.8, 2)$	$\mathcal{U}(2.4, 2.7)$
[0.08, 0.63]	$\mathcal{U}(\log 0.5, \log 2)$	$\mathcal{U}(\log 50, \log 100)$	$\mathcal{U}(0.6, 0.9)$	$\mathcal{U}(1.2, 1.4)$	$\mathcal{U}(1.8, 2)$	$\mathcal{U}(2.1, 2.3)$
[0.08, 0.42]	$\mathcal{U}(\log 0.5, \log 2)$	$\mathcal{U}(\log 50, \log 100)$	$\mathcal{U}(0.5, 0.7)$	$\mathcal{U}(1.3, 1.4)$	$\mathcal{U}(1.7, 1.8)$	$\mathcal{U}(1.8, 2)$

NOTE. — The r_p boundaries on the radius valley are given implicitly by the upper r_p limit on the terrestrial peak and the lower r_p limit on the gaseous peak.

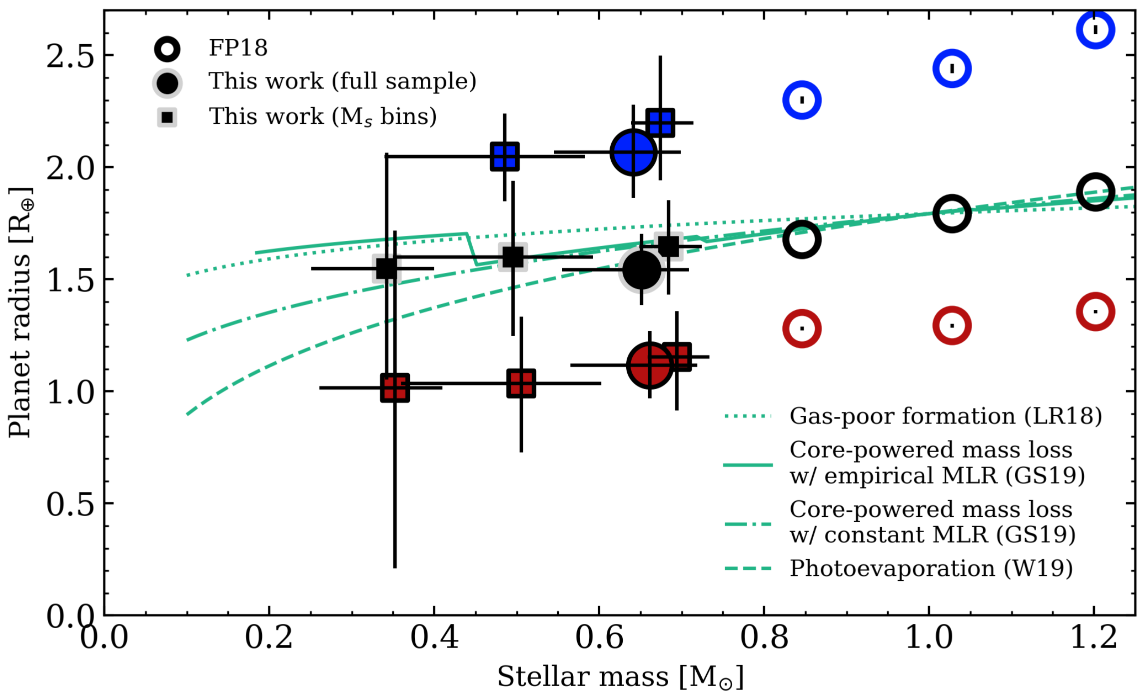

FIG. 13.— Evolution of the radius valley features with stellar mass. *Solid markers*: the occurrence rate-weighted locations of the terrestrial planet peak (*red markers*), the radius valley (*black markers*) and the gaseous planet peak (*blue markers*) as a function of host stellar mass. Measurements around Sun-like stars with $M_s > 0.8$ $M_\odot$ are retrieved from Fulton & Petigura (2018) (*open circles*). Feature radii from our full sample with a median value of $M_s = 0.651^{+0.058}_{-0.096}$ $M_\odot$ are depicted as *filled circles*. *Filled squares* depict the feature radii from partitioning our stellar sample into three M_s bins: $M_s \in [0.65, 0.93]$ $M_\odot$, $M_s \in [0.08, 0.65]$ $M_\odot$, and $M_s \in [0.08, 0.42]$ $M_\odot$. Markers in each stellar mass bin are slightly offset along the M_s axis to assist in visualizing the errorbars. Uncertainties on the peak and valley locations are derived by sampling the measured occurrence rates and their uncertainties along with samples of the hyperparameters controlling map smoothing, minimum detection sensitivity, planet parameter binning, and the assumed feature ranges in P and r_p. The green curves represent theoretical predictions for the evolution of the radius valley with stellar mass based on physical models of gas-poor terrestrial planet formation (*dotted*; Lopez & Rice 2018), core-powered atmospheric mass loss with an empirical mass-luminosity relation (*solid*; Gupta & Schlichting 2019b), a constant mass-luminosity relation (*dash-dotted*; Gupta & Schlichting 2019b), and photoevaporation (*dashed*; Wu 2019). The models only predict scaling relations with M_s and as such are anchored to the measured valley location at $M_s \sim M_\odot$.

TABLE 7
RADIUS VALLEY FEATURES VERSUS STELLAR MASS

Stellar mass [$M_\odot$]	Terrestrial peak [$R_\oplus$]	Radius valley [$R_\oplus$]	Gaseous peak [$R_\oplus$]
$0.651^{+0.058}_{-0.096}$	$1.118^{+0.151}_{-0.148}$	$1.543^{+0.160}_{-0.160}$	$2.068^{+0.211}_{-0.205}$
$0.684^{+0.040}_{-0.035}$	$1.154^{+0.205}_{-0.239}$	$1.647^{+0.207}_{-0.215}$	$2.197^{+0.301}_{-0.256}$
$0.500^{+0.097}_{-0.146}$	$1.036^{+0.297}_{-0.308}$	$1.599^{+0.340}_{-0.352}$	$2.048^{+0.191}_{-0.199}$
$0.343^{+0.057}_{-0.092}$	$1.017^{+0.700}_{-0.807}$	$1.548^{+0.515}_{-0.496}$	-

NOTE. — As depicted in Fig. 13.

nisms in the low stellar mass regime. Athough the trend of decreasing median feature radii with decreasing stellar mass appears to be upheld, the poor counting statistics in the reduced M_s bins prevent any significant inference regarding the relative strength of the competing physical mechanisms. This problem can only be addressed by increasing the number of mid-to-late M dwarfs in transit surveys and by maximizing the detection sensitivity to planets spanning the radius valley.

7. DISCUSSION

7.1. *Improving constraints on the occurrence rate of small close-in planets orbiting mid-M dwarfs*

The issue of having insufficient information to distinguish between photoevaporation, core-powered mass loss, and gas-poor formation around low mass stars can be addressed with two steps. Firstly, by expanding the low mass stellar sample in transiting planet searches and secondly, by quantifying the detection sensitivity in those searches. NASA's Transiting Exoplanet Survey Satellite (*TESS*; Ricker et al. 2015) is expected to provide hundreds of new transiting planet discoveries in the vicinity of the radius valley (Barclay et al. 2018). *TESS* is particularly well-suited to the discovery of close-in planets around low mass stars down to M5V ($M_s \sim 0.16$ M$_\odot$) due to its red bandpass (600-1000 nm) and its high cadence (2 minute) observations of 200,000-400,000 stars over $\sim 94\%$ of the sky by the completion of its recently approved extended mission.

The *TESS* primary mission, lasting one year, has been ongoing since July 2018. Based on the photometric performance of the mission and consequently on the success of planet searches by the Science Processing Operations Center (SPOC; Jenkins et al. 2016; Twicken et al. 2018; Li et al. 2018) at the time of writing, we can estimate the number of low mass stars required to be observed by *TESS* to enable robust conclusions regarding the nature of the radius valley's formation. These calculations proceed by noting that the measurement uncertainty on the feature locations from binomial statistics scales as $\sqrt{N_s P(1-P)}$ where N_s is the number of observed stars and P is the probability of detecting a planet close to the radius valley given the detection sensitvity, the transit proability, and their inherent rate of occurrence (see Eq. 7). Through sectors 1-14, *TESS* has observed $N_{s,\mathrm{TESS}} = 23051$ stars less massive than 0.4 M$_\odot$ with 2 minute cadence from its Candidate Target List (CTL; Stassun et al. 2019). Among these stars, the SPOC has reported three objects of interest close to the radius valley between $1.4 - 1.6$ R$_\oplus$. Assuming a 0% false positive rate among these planet candidates, and the same MAP occurrence rate measured with *Kepler* ($f_{\mathrm{valley}} \approx 0.19$ planets per star), we find the probability of *TESS* to detect a transiting planet spanning the radius valley around a star with $M_s < 0.4$ M$_\odot$ to be $P_{\mathrm{valley,TESS}} = 1.30 \times 10^{-4}$. We can compare these to the *Kepler* values of $N_{s,\mathrm{Kep}} = 33$ and $P_{\mathrm{valley,Kep}} = 8.56 \times 10^{-3}$ to scale the uncertainty on f_{valley}, and hence on the radius valley location, as an increasing number of mid-to-late M dwarfs are observed with 2 minute cadence with *TESS*.

The resulting improvement in the measurement precision of the radius valley with observations of additional mis-to-late M dwarfs is shown in Fig. 14. The *TESS* curve reveals how precisely the location of the radius valley can be measured given *TESS*'s approximate detection sensivity to planets spanning the radius valley and as the number of low mass stars observed with 2 minute cadence is increased. Note that the improvement allotted by *TESS* should only be interpretted as an approximation given that its detection sensitivity has not been adequately characterized. In our calculations, the *TESS* detection sensitivity is estimated as a constant value as described in the preceding paragraph.

We define a target measurement precision as that which is required to distinguish between predictions from photoevaporation and core-powered mass loss (assuming an empirical mass-luminosity relation) at 3σ around low mass stars with a median stellar mass of 0.35 M$_\odot$. Based on the model curves in Fig. 13, this required precision corresponds to a radius vally uncertainty of ~ 0.12 R$_\oplus$. A very similar level of precision would be required to distinguish between photoevaporation and gas-poor formation as well. The approximate *TESS* detection sensitivity implies that *TESS* will be required to observe $\sim 85,000$ mid-to-late M dwarfs to distinguish between model predictions of photoevaporation and core-powered mass low or gas-poor formation at 3σ. At the time of writing, only 23051 such stars have been targeted with 2 minute cadence. Extrapolating to end of *TESS*'s primary mission, we expect a total of $\sim 42,000$ such stars to be observed with 2 minute cadence. If the *TESS* detection sensitivity is well-characterized by that time and is roughly consistent with the approximate value assumed here then *TESS* could achieve a radius valley uncertainty of ~ 0.17 R$_\oplus$ by the end of its prime mission. This would still be useful for constraining radius valley formation models as predictions from photoevaporation and core-powered mass loss, or gas-poor formation, could be distingiushed at $\sim 2.1\sigma$ with this level of precision. Note that these calculations do not include non-CTL stars that may be targeted in the 30 minute *TESS* Full Frame Images and could also contribute to occurrence rate measurements. Albeit with a reduced detection sensitivity.

Also included in Fig. 14 is the curve for a hypothetical continuation of the primary *Kepler* mission. The calculation reveals that had *Kepler* been able to continue its prime mission and had access to thousands of additional mid-M dwarfs than were targeted in the primary *Kepler* field, then the location of the radius valley could have been precisely measured with ~ 1200 stars observed.

The stellar input catalog for the up-coming ESA *PLATO* mission (Rauer et al. 2014) has yet to be defined. The primary goal of *PLATO* is to detect and characterize transiting habitable zone planets around bright FGK stars. Despite this, according to the mission's Defintion Study Report[7], a subset of the PLATO Input Catalog (PIC) known as sample P4 will target ≥ 5000 M dwarfs brighter than $V = 16$ as part of the mission's Long-Duration Observing Phase (LOP) lasting a minimum of two years. Furthermore, the expected random noise in P4 is 800 ppm on one hour timescales. To compute the probability of detecting a radius valley planet around a mid-to-late M dwarf targeted by *PLATO*, we first assume that for a given transit S/N, *PLATO*'s detection sensivity will be equivalent to that of *Kepler*. We fix the transit S/N (Eq. 4) of a radius valley planet orbiting a mid-M dwarf using the values $r_p = 1.5$ R$_\oplus$, $R_s = 0.35$ R$_\odot$, $M_s = 0.35$ M$_\odot$, CDPP$_{1\ \mathrm{hr}} = 800$ ppm, and $n_{\mathrm{transits}} = 73$ for stars in the LOP observing phase. We note the inexact nature of this calculation which neglects the observing cadence and variations in the transit depth and photometric precision with each P4 star. Nevertheless, assuming the *Kepler* occurrence rate we estimate that $P_{\mathrm{valley,PLATO}} = 2.35 \times 10^{-3}$. This probability is ~ 18 times the estimated value for *TESS* but is about one quarter that of *Kepler*. The expected radius valley measurement precision with *PLATO* is also depicted in Fig. 14.

Although the exact P4 M dwarf sample is not yet de-

[7] https://sci.esa.int/web/plato/-/59252-plato-definition-study-report-red-book

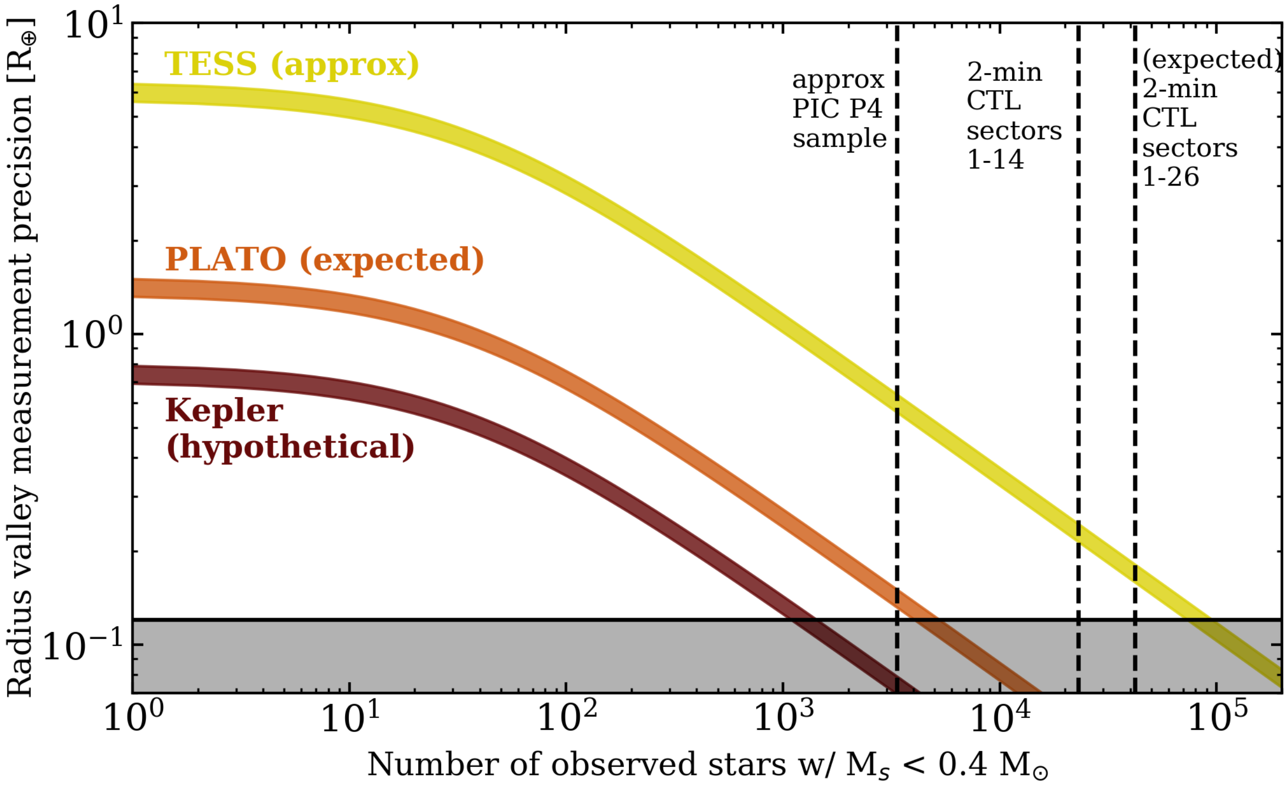

FIG. 14.— Expected improvement in the measurement precision of the radius valley location with the number of observed low mass stars. Measurement precision curves of the upper and lower uncertainties on the location of the radius valley are derived from binomial statistics. These calculations use approximate and expected detection sensitivity values for the *TESS* (*yellow*) and *PLATO* (*orange*) missions. The curve representing a hypothetical continuation of the primary *Kepler* mission is depicted in *red*. The target precision of 0.12 $R_\oplus$ (*shaded region*) would enable models of photoevaporation to be distinguished from core-powered mass loss or gas-poor formation at 3σ at $M_s = 0.35$ $M_\odot$. Measurement precision curves are compared to the sample sizes of mid-to-late M dwarfs from the *TESS* Candidate Target List for sectors 1-14 (i.e. to the time of writing) and 1-26 (i.e. to the end of the *TESS* primary mission). Also included is the approximate number of mid-to-late M dwarfs included in the P4 sample of the PLATO Input Catalog.

fined, recent developments at the September 2019 PIC Workshop in Italy[8] concluded that the properties of the sample will be consistent with M dwarf stars in the solar vicinity. Although this statement if very rough and not binding to the final definition of the PIC, we combine this expectation with knowledge of M dwarfs in the solar neightbourhood to estimate the number of mid-to-late M dwarfs in the P4 sample. We do so by retrieving the M dwarf sample within 25 pc from Winters et al. (2019). Noting that this volume-limited sample is $\sim 33\%$ complete (J. Winters private communication), we identify ~ 2829 M dwarfs within 25 pc and with $V < 16$. We then scale up the M dwarf population beyond 25 pc until 5000 stars with $V < 16$ are included assuming a homogenous M dwarf population beyond 25 pc. Of those stars which represent probable targets within the P4 sample, 3358 have masses < 0.4 $M_\odot$. With this many mid-to-late M dwarfs targeted by *PLATO*, we expect the radius valley uncertainty to reach ~ 0.14 $R_\oplus$ which would enable models of photoevaporation is be distrignshed from models of core-powered mass low or gas-poor formation at $\sim 2.6\sigma$.

7.2. *Implications for RV planet searches around low mass stars*

Many existing and up-coming RV spectrographs will be partially focused on characterizing the masses of planets spanning the radius valley in order to improve our physical understanding of the nature of those planets. In particular, a subset of those spectrographs operating in the near-IR will focus heavily on M dwarf planetary systems (e.g. CARMENES; Quirrenbach et al. 2014, HPF; Mahadevan et al. 2012, IRD; Kotani et al. 2014, NIRPS; Bouchy et al. 2017, SPIRou; Donati et al. 2018). In defining target samples that are equally complete on either side of the radius valley, it is critically important that the location of the transition between terrestrial and gaseous planets is known. In our full stellar sample which includes mid-to-late K dwarfs, the measured radius valley location is 1.54 ± 0.16 $R_\oplus$. Although we remind the reader that the exact value is dependent on the planet's separation (see Fig. 11). A consistent value of $1.55^{+0.52}_{-0.50}$ is also recovered, albeit with reduced signficance, around stars later than about M2.5V. This value is slightly lower than the valley locations measured around Sun-like stars of ~ 1.9 $R_\oplus$ for $M_s \sim 1.2$ $M_\odot$ and ~ 1.7 $R_\oplus$ for $M_s \sim 0.85$ $M_\odot$ (Fulton & Petigura 2018).

7.3. *Imperfect clearing of the radius valley*

should I even talk about this? not sure what to say other than the gap is not fully cleared. As evidenced in Fig. 8, the radius valley around low mass stars is partially filled by planets. This feature

[8] https://indico.ict.inaf.it/event/806/

This work elucidates the location of the radius valley around M dwarf host stars and guides observers to the planetary radii from transit surveys that are of interest for fully characterizing the radius valley in terms of planetary bulk densities.

mass dependence of the gap:

The weighted feature radii are also effected by planetary magnetic fields which directly impact the efficiency of atmospheric stripping in the photoevaporation scenario (Owen & Adams 2019). The persistence of a planetary magnetic field acts to shield the planet's atmosphere from XUV stellar photons thus enhancing the retention of the atmosphere and shifting the location of the radius valley to larger radii.

valley filling increases with deacreasing stellar mass

In the photoevaporation scenario, the partial filling of the gap around low mass stars may be explained by their lower XUV luminosities relative to Sun-like such as those included in the CKS stellar sample ().

This explanation seems to be supported by the stellar mass dependent gap measurements from Fulton & Petigura (2018).

summary of McDonald+2019 (https://ui-adsabs-harvard-edu.ezp-prod1.hul.harvard.edu/abs/2019ApJ...876...22M/abstract): X-rays only since XUV observations are difficult for non-Sun-like stars and X-rays are the dominant driver of atmospheric loss by photoevaporation. Jackson+12 & Shkolnik+14 derived scalings from data for the LX/Lbol evolution over time for 0.3 - 1.3 solar mass stars on the MS, low mass stars ($\lesssim 0.8$ M$_\odot$) exhibit a LX/Lbol that is typically a few to ten times greater than around Sun-like stars ($0.8 - 1.12$ M$_\odot$) (fig 1 in McDonald+2019). scaling these values by the typical bolometric luminosities of stars in the various mass bins reveals that Sun-like stars having higher absolute X-ray luminosities which contributes to more efficient clearing of the gap by photoevaporation.

8. SUMMARY OF PRIMARY FINDINGS

This study presented calculations of the occurrence rate of small close-in planets orbiting low mass stars using data from the *Kepler* and *K2* transit surveys. Our main findings are summarized below.

- The radius valley structure in the occurrence rate of small close-in planets, previously resolved around Sun-like stars, is demonstrated to persist around low mass stars (i.e. mid-K to mid-M dwarfs).

- The radius valley around low mass stars exhibits a negative slope with insolation unlike around Sun-like stars whose measured slope is positive. The observed slope supports models of gas-poor terrestrial planet formation without invoking atmospheric escape from photoevaporation or core-powered mass loss.

- The gaseous planet peak in the bimodal occurrence rates, centered at ~ 2 R$_\oplus$, effectively vanishes around mid-M dwarfs as terrestrial planets ($\lesssim 1.6$ R$_\oplus$) increasingly dominate the close-in planet population towards later spectral types. The relative fraction of terrestrial to gaseous planets increases from $\sim 0.5 \pm 0.1 \to 8.5 \pm 4.6$ from mid-K to mid-M dwarfs.

- The occurrence rate-weighted central radius of each planet peak and the radius valley decrease to smaller sizes with decreasing stellar mass. The slope of the gaseous peak's radius with stellar mass is shown to be twice that of the terrestrial peak radius indicating that the complete planet population in each stellar mass bin tends to converge towards rocky planet cores of ~ 1 R$_\oplus$ around later spectral types.

- Robust inferences to distinguish between various proposed physical mechanisms for the formation of the radius valley are expoected to require $\mathcal{O}(85,000)$ mid-M dwarfs observed with 2 minute cadence with *TESS* or $\mathcal{O}(4700)$ mid-M dwarfs observed with *PLATO* based on its expected performance and observing strategy.

We thank Martin Paegert for his efforts in contributing relevent data to and for his assistance in querying the *TESS* CTL database. We also thank Jennifer Winters and Jonathan Irwin for discussions regarding the population of nearby M dwarfs. We thank the Canadian Institute for Theoretical Astrophysics for use of the Sunnyvale computing cluster throughout the early stages of this work. RC was supported by the Natural Sciences and Engineering Research Council of Canada and the NASA ??.

REFERENCES

Aigrain, S., Hodgkin, S. T., Irwin, M. J., Lewis, J. R., & Roberts, S. J. 2015, MNRAS, 447, 2880
Aigrain, S., Parviainen, H., & Pope, B. J. S. 2016, MNRAS, 459, 2408
Akeson, R. L., Chen, X., Ciardi, D., et al. 2013, PASP, 125, 989
Bailer-Jones, C. A. L., Rybizki, J., Fouesneau, M., Mantelet, G., & Andrae, R. 2018, AJ, 156, 58
Barclay, T., Pepper, J., & Quintana, E. V. 2018, ArXiv e-prints, arXiv:1804.05050
Barnes, J. W. 2007, PASP, 119, 986
Barnes, R., & Greenberg, R. 2006, ApJL, 647, L163
Benneke, B., Werner, M., Petigura, E., et al. 2017, ApJ, 834, 187
Berger, T. A., Huber, D., Gaidos, E., & van Saders, J. L. 2018, ApJ, 866, 99
Bonfils, X., Delfosse, X., Udry, S., et al. 2013, A&A, 549, A109
Bonfils, X., Almenara, J.-M., Cloutier, R., et al. 2018, A&A, 618, A142
Bouchy, F., Doyon, R., Artigau, É., et al. 2017, The Messenger, 169, 21
Bovy, J., Rix, H.-W., Green, G. M., Schlafly, E. F., & Finkbeiner, D. P. 2016, ApJ, 818, 130
Boyajian, T. S., von Braun, K., van Belle, G., et al. 2012, ApJ, 757, 112
Burke, C. J., & Catanzarite, J. 2017, Planet Detection Metrics: Per-Target Flux-Level Transit Injection Tests of TPS for Data Release 25, Tech. rep.
Burke, C. J., Christiansen, J. L., Mullally, F., et al. 2015, ApJ, 809, 8
Chen, H., & Rogers, L. A. 2016, ApJ, 831, 180
Chen, J., & Kipping, D. 2017, ApJ, 834, 17

Chiang, E., & Laughlin, G. 2013, MNRAS, 431, 3444
Choi, J., Dotter, A., Conroy, C., et al. 2016, ApJ, 823, 102
Christiansen, J. L., Clarke, B. D., Burke, C. J., et al. 2015, ApJ, 810, 95
—. 2016, ApJ, 828, 99
Christiansen, J. L., Vanderburg, A., Burt, J., et al. 2017, AJ, 154, 122
Claret, A., Hauschildt, P. H., & Witte, S. 2012, A&A, 546, A14
Cloutier, R. 2019, AJ, 158, 81
Cloutier, R., Astudillo-Defru, N., Bonfils, X., et al. 2019a, arXiv e-prints, arXiv:1905.10669
Cloutier, R., Astudillo-Defru, N., Doyon, R., et al. 2019b, A&A, 621, A49
Crossfield, I. J. M., Ciardi, D. R., Petigura, E. A., et al. 2016, ApJS, 226, 7
Dittmann, J. A., Irwin, J. M., Charbonneau, D., Berta-Thompson, Z. K., & Newton, E. R. 2017, AJ, 154, 142
Donati, J.-F., Kouach, D., Lacombe, M., et al. 2018, SPIRou: A NIR Spectropolarimeter/High-Precision Velocimeter for the CFHT, 107
Dragomir, D., Teske, J., Günther, M. N., et al. 2019, ApJ, 875, L7
Dressing, C. D., & Charbonneau, D. 2013, ApJ, 767, 95
—. 2015, ApJ, 807, 45
Dressing, C. D., Charbonneau, D., Dumusque, X., et al. 2015, ApJ, 800, 135
Dressing, C. D., Vanderburg, A., Schlieder, J. E., et al. 2017, AJ, 154, 207
Eker, Z., Bakış, V., Bilir, S., et al. 2018, MNRAS, 479, 5491
Fressin, F., Torres, G., Charbonneau, D., et al. 2013, ApJ, 766, 81
Fulton, B. J., & Petigura, E. A. 2018, AJ, 156, 264
Fulton, B. J., Petigura, E. A., Howard, A. W., et al. 2017, AJ, 154, 109
Gaidos, E., Mann, A. W., Kraus, A. L., & Ireland, M. 2016, MNRAS, 457, 2877
Gillon, M., Triaud, A. H. M. J., Demory, B.-O., et al. 2017, Nature, 542, 456
Ginzburg, S., Schlichting, H. E., & Sari, R. 2018, MNRAS, 476, 759
Gupta, A., & Schlichting, H. E. 2019a, MNRAS, 487, 24
—. 2019b, arXiv e-prints, arXiv:1907.03732
Hamann, A., Montet, B. T., Fabrycky, D. C., Agol, E., & Kruse, E. 2019, AJ, 158, 133
Hardegree-Ullman, K. K., Cushing, M. C., Muirhead, P. S., & Christiansen, J. L. 2019, AJ, 158, 75
Harpsøe, K. B. W., Hardis, S., Hinse, T. C., et al. 2013, A&A, 549, A10
Henry, T. J., & McCarthy, Jr., D. W. 1993, AJ, 106, 773
Hirano, T., Dai, F., Gandolfi, D., et al. 2018, AJ, 155, 127
Howard, A. W., Marcy, G. W., Bryson, S. T., et al. 2012, ApJS, 201, 15
Howell, S. B., Sobeck, C., Haas, M., et al. 2014, PASP, 126, 398
Huber, D., Zinn, J., Bojsen-Hansen, M., et al. 2017, ApJ, 844, 102
Jackson, A. P., Davis, T. A., & Wheatley, P. J. 2012, MNRAS, 422, 2024
Jenkins, J. M., Chandrasekaran, H., McCauliff, S. D., et al. 2010, in SPIE, Vol. 7740, Software and Cyberinfrastructure for Astronomy, 77400D
Jenkins, J. M., Twicken, J. D., McCauliff, S., et al. 2016, in SPIE, Vol. 9913, Software and Cyberinfrastructure for Astronomy IV, 99133E
Jin, S., & Mordasini, C. 2018, ApJ, 853, 163
Jin, S., Mordasini, C., Parmentier, V., et al. 2014, ApJ, 795, 65
Koch, D. G., Borucki, W. J., Basri, G., et al. 2010, ApJ, 713, L79
Kostov, V. B., Schlieder, J. E., Barclay, T., et al. 2019, AJ, 158, 32
Kotani, T., Tamura, M., Suto, H., et al. 2014, in SPIE, Vol. 9147, Ground-based and Airborne Instrumentation for Astronomy V, 914714
Kreidberg, L. 2015, PASP, 127, 1161
Kruse, E., Agol, E., Luger, R., & Foreman-Mackey, D. 2019, ApJS, 244, 11
Lee, E. J., & Chiang, E. 2016, ApJ, 817, 90
Lee, E. J., Chiang, E., & Ormel, C. W. 2014, ApJ, 797, 95
Li, J., Tenenbaum, P., Twicken, J. D., et al. 2018, arXiv e-prints, arXiv:1812.00103
Lindegren, L., Hernández, J., Bombrun, A., et al. 2018, A&A, 616, A2
Livingston, J. H., Crossfield, I. J. M., Petigura, E. A., et al. 2018, AJ, 156, 277
Lopez, E. D., & Fortney, J. J. 2014, ApJ, 792, 1
Lopez, E. D., & Rice, K. 2018, MNRAS, 479, 5303
Luger, R., Agol, E., Kruse, E., et al. 2016, AJ, 152, 100
Luger, R., Kruse, E., Foreman-Mackey, D., Agol, E., & Saunders, N. 2018, AJ, 156, 99
Lundkvist, M. S., Kjeldsen, H., Albrecht, S., et al. 2016, Nature Communications, 7, 11201
Luque, R., Pallé, E., Kossakowski, D., et al. 2019, A&A, 628, A39
MacDonald, M. G., Ragozzine, D., Fabrycky, D. C., et al. 2016, AJ, 152, 105
Mahadevan, S., Ramsey, L., Bender, C., et al. 2012, in SPIE, Vol. 8446, Ground-based and Airborne Instrumentation for Astronomy IV, 84461S
Mamajek, E. E., Torres, G., Prsa, A., et al. 2015, arXiv e-prints, arXiv:1510.06262
Mandel, K., & Agol, E. 2002, ApJL, 580, L171
Mann, A. W., Feiden, G. A., Gaidos, E., Boyajian, T., & von Braun, K. 2015, ApJ, 804, 64
Martinez, C. F., Cunha, K., Ghezzi, L., & Smith, V. V. 2019, ApJ, 875, 29
Mathur, S., Huber, D., Batalha, N. M., et al. 2017, ApJS, 229, 30
Mayo, A. W., Vanderburg, A., Latham, D. W., et al. 2018, AJ, 155, 136
Mazeh, T., Holczer, T., & Faigler, S. 2016, A&A, 589, A75
Ment, K., Dittmann, J. A., Astudillo-Defru, N., et al. 2019, AJ, 157, 32
Montet, B. T., Morton, T. D., Foreman-Mackey, D., et al. 2015, ApJ, 809, 25
Morales, J. C., Mustill, A. J., Ribas, I., et al. 2019, Science, 365, 1441
Morton, T. D. 2012, ApJ, 761, 6
—. 2015, VESPA: False positive probabilities calculator, Astrophysics Source Code Library, ascl:1503.011
Morton, T. D., & Swift, J. 2014, ApJ, 791, 10
Mulders, G. D., Pascucci, I., & Apai, D. 2015, ApJ, 798, 112
Owen, J. E., & Adams, F. C. 2019, MNRAS, 490, 15
Owen, J. E., & Wu, Y. 2013, ApJ, 775, 105
—. 2017, ApJ, 847, 29
Pecaut, M. J., & Mamajek, E. E. 2013, ApJs, 208, 9
Persson, C. M., Fridlund, M., Barragán, O., et al. 2018, A&A, 618, A33
Petigura, E. A., Howard, A. W., & Marcy, G. W. 2013, Proceedings of the National Academy of Science, 110, 19273
Petigura, E. A., Howard, A. W., Marcy, G. W., et al. 2017, AJ, 154, 107
Petigura, E. A., Marcy, G. W., Winn, J. N., et al. 2018, AJ, 155, 89
Quirrenbach, A., Amado, P. J., Caballero, J. A., et al. 2014, in SPIE, Vol. 9147, Ground-based and Airborne Instrumentation for Astronomy V, 91471F
Rauer, H., Catala, C., Aerts, C., et al. 2014, Experimental Astronomy, 38, 249
Rice, K., Malavolta, L., Mayo, A., et al. 2019, MNRAS, 484, 3731
Ricker, G. R., Winn, J. N., Vanderspek, R., et al. 2015, Journal of Astronomical Telescopes, Instruments, and Systems, 1, 014003
Rogers, L. A. 2015, ApJ, 801, 41
Stassun, K. G., Oelkers, R. J., Paegert, M., et al. 2019, AJ, 158, 138
Twicken, J. D., Catanzarite, J. H., Clarke, B. D., et al. 2018, Publications of the Astronomical Society of the Pacific, 130, 064502
Van Eylen, V., Agentoft, C., Lundkvist, M. S., et al. 2018, MNRAS, 479, 4786
Vanderburg, A., & Johnson, J. A. 2014, PASP, 126, 948
Weiss, L. M., & Marcy, G. W. 2014, ApJL, 783, L6
Winters, J. G., Henry, T. J., Jao, W.-C., et al. 2019, AJ, 157, 216
Wu, Y. 2019, ApJ, 874, 91
Youdin, A. N. 2011, ApJ, 742, 38
Zeng, L., Sasselov, D. D., & Jacobsen, S. B. 2016, ApJ, 819, 127
Zhu, W., Petrovich, C., Wu, Y., Dong, S., & Xie, J. 2018, ApJ, 860, 101