



Kepler-167e as a Probe of the Formation Histories of Cold Giants with Inner Super-Earths

Yayaati Chachan¹ , Paul A. Dalba^{2,3,10} , Heather A. Knutson¹ , Benjamin J. Fulton⁴ , Daniel Thorngren⁵ , Charles Beichman^{6,7} , David R. Ciardi⁷ , Andrew W. Howard⁸ , and Judah Van Zandt⁹

¹ Division of Geological and Planetary Sciences, California Institute of Technology, 1200 E. California Blvd., Pasadena, CA 91125, USA; yachachan@caltech.edu

² Department of Earth and Planetary Sciences, University of California Riverside, 900 University Ave., Riverside, CA 92521, USA

³ Department of Astronomy and Astrophysics, University of California Santa Cruz, 1156 High St., Santa Cruz, CA, USA

⁴ NASA Exoplanet Science Institute/Caltech-IPAC, 1200 E. California Blvd., Pasadena, CA 91125, USA

⁵ Institute for Research on Exoplanets (iREx), Université de Montréal, Montréal, QC, Canada

⁶ Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Dr., Pasadena, CA 91109, USA

⁷ Caltech/IPAC-NASA Exoplanet Science Institute, 770 S. Wilson Ave., Pasadena, CA 91106, USA

⁸ Cahill Center for Astrophysics, California Institute of Technology, 1216 E. California Blvd., Pasadena, CA 91125, USA

⁹ Department of Physics & Astronomy, University of California Los Angeles, Los Angeles, CA 90095, USA

Received 2021 August 25; revised 2021 November 22; accepted 2021 November 29; published 2022 February 11

Abstract

The observed correlation between outer giant planets and inner super-Earths is emerging as an important constraint on planet formation theories. In this study, we focus on Kepler-167, which is currently the only system known to contain both inner transiting super-Earths and a confirmed outer transiting gas giant companion beyond 1 au. Using long-term radial velocity monitoring, we measure the mass of the gas giant Kepler-167e ($P = 1071$ days) to be $1.01^{+0.16}_{-0.15} M_J$, thus confirming it as a Jupiter analog. We refit the Kepler photometry to obtain updated radii for all four planets. Using a planetary structure model, we estimate that Kepler-167e contains $66 \pm 19 M_{\oplus}$ of solids and is significantly enriched in metals relative to its solar-metallicity host star. We use these new constraints to explore the broader question of how systems like Kepler-167 form in the pebble accretion framework for giant planet core formation. We utilize simple disk evolution models to demonstrate that more massive and metal-rich disks, which are the most favorable sites for giant planet formation, can also deliver enough solids to the inner disk to form systems of super-Earths. We use these same models to constrain the nature of Kepler-167's protoplanetary disk and find that it likely contained $\gtrsim 300 M_{\oplus}$ of dust and was $\gtrsim 40$ au in size. These values overlap with the upper end of the observed dust mass and size distributions of Class 0 and I disks and are also consistent with the observed occurrence rate of Jupiter analogs around Sun-like stars.

Unified Astronomy Thesaurus concepts: Planet formation (1241); Super Earths (1655); Extrasolar gaseous giant planets (509); Radial velocity (1332); Protoplanetary disks (1300)

Supporting material: machine-readable table

1. Introduction

The relative rarity of Jupiter analogs around Sun-like stars suggests that only $\sim 10\%$ of protoplanetary disks provide the conditions needed for their formation (Cumming et al. 2008; Wittenmyer et al. 2016, 2020; Fulton et al. 2021). In contrast, close-in super-Earths and sub-Neptunes, which are $< 4 R_{\oplus}$ in size, appear to be much more common (30%–50% occurrence rate for Sun-like stars; Batalha et al. 2013; Fressin et al. 2013; Petigura et al. 2018). It was initially thought that distant gas giants and close-in super-Earths were unlikely to occur in the same system, as the growing gas giant planet was expected to prevent the formation of massive inner planets (Izidoro et al. 2015; Ormel et al. 2017). However, there is now growing observational evidence suggesting that cold gas giants are frequently accompanied by inner super-Earths (Zhu & Wu 2018; Bryan et al. 2019; Herman et al. 2019). This suggests that the same protoplanetary disk properties that allow for the formation of distant giant planets are also compatible with the formation of inner super-Earths, and

that the presence of an outer gas giant does not disrupt super-Earth formation. This observed correlation between inner super-Earths and outer giants therefore provides an important constraint on planet formation theories, as they must explain the formation of both types of planets in the same system.

Systems with multiple transiting super-Earths are particularly valuable for testing formation and migration models, as the transit photometry can be used to characterize their mutual inclinations and corresponding dynamical histories (e.g., Masuda et al. 2020). When combined with radial velocity (RV) or transit timing follow-up to determine masses, we can additionally measure their average densities and calculate the corresponding masses in solids for these planets (e.g., Dubber et al. 2019; Santerne et al. 2019; Dalba et al. 2021). Transit surveys like Kepler (Borucki et al. 2010) and the Transiting Exoplanet Survey Satellite (TESS; Ricker et al. 2015) are more sensitive than RV surveys (e.g., Howard et al. 2010b; Mayor et al. 2011; Rosenthal et al. 2021) to the presence of close-in planets in edge-on orbits with masses less than $10 M_{\oplus}$ (Winn 2018),¹¹ making it easier to obtain a relatively complete census of the inner regions of these planetary systems. However, the probability of seeing a transit

¹⁰ NSF Astronomy and Astrophysics Postdoctoral Fellow.

¹¹ The lower average sensitivity of RV surveys to planets below $10 M_{\oplus}$, which dominate the population of close-in planets, may explain why the correlation between super-Earth and cold Jupiter occurrence rates appears weaker in RV-only samples (Barbato et al. 2018; Rosenthal et al. 2021).

decreases with increasing semimajor axis, and the need to observe multiple transits imposes a hard limit on detectability that is a function of the duration of the survey. To date, only the Kepler survey has had the sensitivity to detect transiting planets beyond ~ 1 au, and they constitute a tiny fraction of the sample of known transiting planets (Wang et al. 2015; Foreman-Mackey et al. 2016; Uehara et al. 2016; Kawahara & Masuda 2019). It is therefore quite rare to find transiting outer companions to close-in super-Earths, but that may not necessarily indicate that this configuration is rare.

Kepler-167 is unique among the sample of transiting planetary systems detected by Kepler, as it contains three close-in super-Earths accompanied by a confirmed transiting $0.9 R_J$ gas giant planet at 1.9 au (Kipping et al. 2016; Dalba & Tamburo 2019). However, the measured radius of the outer gas giant is consistent with more than an order of magnitude range in its predicted mass (Stevenson 1982), making it difficult to predict its dynamical effect on the inner super-Earths. In Section 2, we present RV observations of the Kepler-167 system collected over 4 yr with the High Resolution Echelle Spectrometer (HIRES) instrument on the Keck telescope. In Section 3, we carry out a joint analysis of the RV and transit data, which we use to place tight constraints on the mass and eccentricity of Kepler-167e. We also derive revised stellar properties using new Gaia data and refit the Kepler data for the inner super-Earths in order to provide updated radii for these planets. In Section 4, we use Kepler-167e’s measured mass and radius to constrain its bulk metallicity using the methods described in Thorngren & Fortney (2019). Since the super-Earths are not detected in our RV data, we estimate their masses using a nonparametric mass–radius (M – R) relationship (Ning et al. 2018). This allows us to obtain an estimate of the total metal mass contained in the Kepler-167 planets and the corresponding minimum solid mass required to assemble this system.

In Section 5, we examine the implications of these results for the formation history of the Kepler-167 system. We know that the gas giant core must have formed early in order to undergo runaway accretion. In the pebble accretion paradigm, the core grows until it reaches the “isolation mass.” This mass marks the point where the core is massive enough to perturb the gas disk, forming a pressure trap beyond its orbit that effectively halts the accretion of pebbles. This pressure trap also blocks the transport of pebbles to the inner disk (Morbidelli & Nesvorný 2012; Lambrechts & Johansen 2014), reducing the reservoir of solids available to form super-Earths (Ormel et al. 2017; Lambrechts et al. 2019). However, pebble accretion is known to be a fairly lossy process (Ormel 2017; Lin et al. 2018). That is, prior to reaching the isolation mass, a substantial amount of solids can flow past the growing giant planet core.

We use simple dust evolution models (Birnstiel et al. 2010, 2012) to model the growth of the giant planet core in the outer disk and track the evolution of the solid-mass reservoir in the inner disk (Ormel & Klahr 2010; Lambrechts & Johansen 2012). We vary the effective pebble flux by changing key disk properties such as initial solid mass and size. This allows us to determine which disks are able to form giant planets and to quantify the effect that the formation of the outer gas giant has on the amount of dust that reaches the inner disk. We use these models to relate the estimated solid masses of Kepler-167e and the inner super-Earths to the likely properties of the primordial disk. More broadly, we place constraints on the types of disks that can

produce systems of inner super-Earths and outer gas giant companions under the pebble accretion paradigm. We summarize our conclusions in Section 6.

2. Observations

2.1. Archival Photometry

Archival photometry of Kepler-167 exists from the Kepler mission (e.g., Borucki et al. 2010), the Spitzer spacecraft (Dalba & Tamburo 2019), and TESS (Ricker et al. 2015). The transiting planets in this system were initially discovered in the Kepler data (Kipping et al. 2016). The Spitzer observations specifically targeted a transit of Kepler-167e, but they only spanned part of the transit light curve. While these observations did not improve the precision of Kepler-167e’s transit ephemeris, they did establish that the transit occurred at the expected time. This discovery significantly mitigated some of the uncertainty inherent to long-period exoplanets with only two observed transits, where the existence of transit timing variations (TTVs) can significantly bias initial estimates of the orbital period (e.g., Dalba & Muirhead 2016; Santerne et al. 2019). Kepler-167 was also observed by TESS in Cycle 2 of its primary mission and Cycle 4 of its extended mission.

Our analysis of Kepler-167 archival photometry only uses the Kepler data. The Spitzer observations mitigate uncertainties in the orbital period due to possible TTVs but do not improve the precision of the outer planet’s ephemeris when we assume a constant ephemeris. The epochs of the TESS photometry span transits of the inner planets but not the outer one. However, TESS was designed to survey stars much brighter than Kepler-167 ($V \approx 14$), and the TESS observations are too imprecise to improve our constraints on the physical properties of the small inner super-Earths in this system.

The Kepler spacecraft observed Kepler-167 during 17 quarters of its primary mission (2009 May through 2013 May). Observations in quarters 1–8 were long cadence (30 minutes), while those in quarters 9–17 were short cadence (1 minute). This observation window spanned dozens to hundreds of transits of the three inner planets and two transits of the outer giant planet (Kipping et al. 2016). We accessed the Pre-search Data Conditioning Simple Aperture Photometry (PDCSAP; Jenkins et al. 2010; Smith et al. 2012; Stumpe et al. 2012) through the Mikulski Archive for Space Telescopes¹² using the `lightkurve`¹³ package (Lightkurve Collaboration et al. 2018). Although the PDCSAP data are corrected for many sources of systematic noise, we noticed a quasiperiodic variability signal in the corrected photometry for this target that is likely due to stellar rotation. We modeled this signal using Gaussian process (GP) regression as implemented in the `celerite2`¹⁴ package built into the `exoplanet`¹⁵ tool kit (Foreman-Mackey et al. 2017, 2021; Foreman-Mackey 2018; Luger et al. 2019; Agol et al. 2020). We fit the long- and short-cadence data with quasiperiodic kernels of different widths but the same period. We determined the maximum a posteriori parameters for the GP (see Section 5.1) with a numerical optimization method (Salvatier et al. 2016). Then, we subtracted the GP signal from the long- and short-cadence data before fitting for the transits.

¹² <https://archive.stsci.edu/>

¹³ <https://docs.lightkurve.org/>

¹⁴ <https://celerite2.readthedocs.io/en/latest/>

¹⁵ <https://docs.exoplanet.codes>

2.2. Keck-HIRES Spectroscopy

We obtained spectroscopic observations of Kepler-167 using HIRES (Vogt et al. 1994) on the Keck I telescope at the W. M. Keck Observatory. We first processed a moderate signal-to-noise ratio ($S/N \sim 40$) reconnaissance spectrum with `SpecMatch`¹⁶ (Petigura 2015; Petigura et al. 2017). This spectrum was originally acquired for the Kepler-167 discovery effort (Kipping et al. 2016) and processed with the Stellar Parameter Classification pipeline (Buchhave et al. 2012). The stellar properties that we obtain from `SpecMatch` are in close agreement with those published by Kipping et al. (2016). Specifically, the stellar metallicity ($[Fe/H]$), effective temperature (T_{eff}), and projected rotational velocity ($v \sin i$) are -0.02 ± 0.09 dex, 4830 ± 100 K, and 2.0 ± 1.0 km s⁻¹, respectively.

We acquired 13 additional spectra between 2017 August 23 and 2020 December 4 with S/Ns ranging between 40 and 52. In these observations, the starlight was passed through a heated iodine cell in order to allow us to obtain a more precise wavelength calibration. Owing to the faintness of Kepler-167 ($V \approx 14$), we did not utilize a high-S/N template spectrum for our RV analysis and instead substituted a best-match template from another star (HD 16160; $T_{\text{eff}} = 4720 \pm 110$ K, $\log g = 4.57 \pm 0.10$, $[Fe/H] = -0.02 \pm 0.09$) in the HIRES spectral library (Yee et al. 2017; Dalba et al. 2020). Aside from the template substitution, the data reduction and Doppler analysis followed the standard procedures of the California Planet Search (Howard et al. 2010a; Howard & Fulton 2016). This analysis produced RV measurements of Kepler-167 with a median internal precision of 3–5 m s⁻¹. We then added an additional 6.2 m s⁻¹ error in quadrature, which is a conservative estimate of the average uncertainty incurred by the match-template technique (Dalba et al. 2020). The resulting individual RV measurements are listed in Table 1. We also include the corresponding S_{HK} activity indicators derived from the Ca II H and K spectral lines (Wright et al. 2004; Isaacson & Fischer 2010). We see no evidence for any covariance between the measured RVs and this activity indicator.

2.3. Constraints on Companion Properties

High-resolution images from the UK Infrared Telescope Survey and Keck NIRC2 reveal the presence of a companion 2'' to the NE of Kepler-167 (Kipping et al. 2016). This companion is ~ 5 mag fainter than Kepler-167 in the Kepler bandpass. Kipping et al. (2016) could not establish whether this companion is bound to Kepler-167. However, the parallax ($\Delta\varpi = 0.15 \pm 0.22$ mas) and proper motion ($\Delta\mu_{\text{R.A.}} = 0.43 \pm 0.27$, $\Delta\mu_{\text{decl.}} = 1.2 \pm 0.3$ mas yr⁻¹) measurements from Gaia DR3 for these two sources agree to within the uncertainties (Gaia Collaboration et al. 2021), thus indicating that they are comoving. The companion has a spectral type of M4V and is $\sim 0.2 M_{\odot}$, which is consistent with the $J-K$ color measured in Kipping et al. (2016). Given the large separation of 2'' (~ 700 au, $P > 15,000$ yr) and the small stellar mass, the companion's effect on the RVs is negligible. Additionally, although the companion's light contaminates the Kepler light curves, its effect on the measured planetary radii is insignificant. The companion is 100 times fainter than Kepler-167 in the Kepler bandpass; thus, the true radii of Kepler-167's planets are only $\sqrt{1.01/1} = 1.005$ larger than our measurements

Table 1
RV Measurements of Kepler-167

BJD _{TDB}	RV (m s ⁻¹)	S_{HK}
2,457,988.955812	6.1 ± 7.2	0.195 ± 0.001
2,458,300.972784	16.1 ± 7.0	0.248 ± 0.001
2,458,328.854650	22.6 ± 7.0	0.235 ± 0.001
2,458,363.864254	6.3 ± 7.2	0.267 ± 0.001
2,458,385.796973	18.1 ± 7.2	0.149 ± 0.001
2,458,645.987639	-11.8 ± 7.3	0.239 ± 0.001
2,458,662.992470	-26.3 ± 7.0	0.236 ± 0.001
2,458,710.986137	-24.2 ± 8.0	0.167 ± 0.001
2,458,797.818538	-23.2 ± 8.1	0.148 ± 0.001
2,459,072.003261	5.6 ± 7.3	0.168 ± 0.001
2,459,101.942839	2.1 ± 7.2	0.166 ± 0.001
2,459,118.833084	14.4 ± 7.1	0.186 ± 0.001
2,459,187.714084	29.9 ± 7.3	0.247 ± 0.001

(This table is available in its entirety in machine-readable form.)

(Ciardi et al. 2015). This effect is an order of magnitude smaller than our uncertainties on the radii.

3. Model Fitting and Parameter Estimation

The combined Kepler data set for the Kepler-167 system contains over 100,000 individual data points measuring signals from four separate transiting planets. When combined with the complexity of a model combining stellar, RV, and transit data, this data volume makes it computationally intractable to fit a single joint model. Instead, we separated the modeling of this system into three parts: the stellar parameters, the long-period gas giant Kepler-167e, and the three inner super-Earths.

We first determined the stellar parameters by fitting archival photometry of Kepler-167 from Gaia (Gaia Collaboration et al. 2018), the Two Micron All Sky Survey (Cutri et al. 2003), and the Wide-field Infrared Survey Explorer (Cutri et al. 2014) with a model spectral energy distribution (SED) to constrain the stellar properties. We employed the EXOFASTv2 modeling suite (Eastman et al. 2013, 2019) to conduct this fit. The fit included the MESA Isochrones and Stellar Tracks (MIST) stellar evolution models (Paxton et al. 2011, 2013, 2015; Choi et al. 2016; Dotter 2016), which provided constraints on the stellar mass and age. We placed normal priors on $[Fe/H]$ (-0.02 ± 0.09) and T_{eff} (4830 ± 100 K) from the `SpecMatch` analysis of the HIRES iodine-free spectrum (Section 2.2) and an upper limit on the line-of-sight extinction ($A_V < 0.4204$) from galactic reddening maps (Schlafly & Finkbeiner 2011). The parallax of Kepler-167 as measured by the Gaia Early Data Release 3 (Gaia Collaboration et al. 2021) and corrected according to Lindgren et al. (2021) is 2.944 ± 0.018 mas, which we applied as a normal prior in the fit. We also enforced a noise floor of 2% on the bolometric flux, as suggested by Tayar et al. (2020). We checked that this fit met the default criteria for convergence in EXOFASTv2, which includes at least 1000 independent draws from the posterior and a Gelman–Rubin statistic below 1.01 for each parameter. The resulting stellar parameters and their corresponding priors are summarized in Table 2. All of our stellar parameters are consistent with those derived by Kipping et al. (2016) to the 2σ level. We inflated the widths of the uncertainties on the stellar mass and radius priors to 5% and 4%, respectively, prior to using these values to calculate absolute planetary parameters.

¹⁶ <https://github.com/petigura/specmatch-syn>

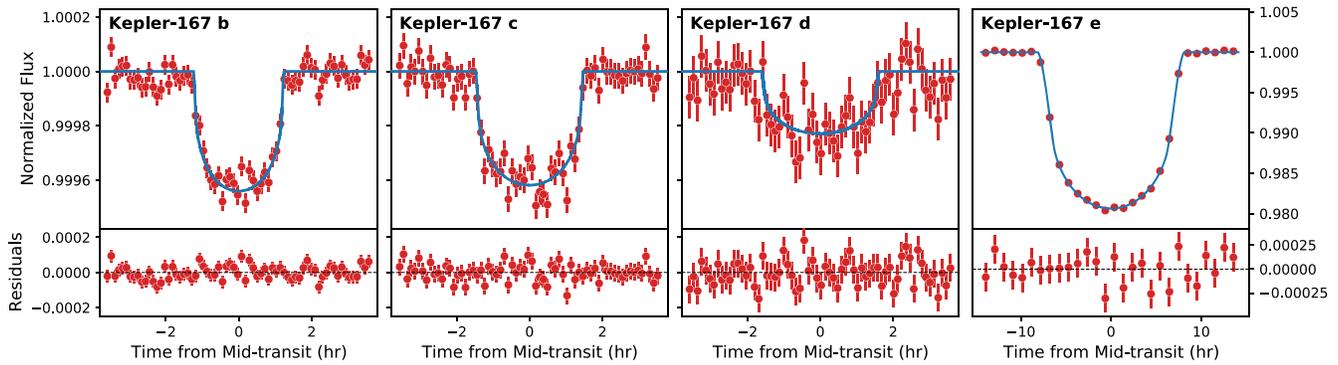


Figure 1. Optical Kepler transit photometry of the four Kepler-167 planets folded on their best-fit ephemeris along with their best-fit models (blue). Note the difference in scale in the panel for Kepler-167e. The data (red) are shown in bins of 10 minutes for the inner planets and 1 hr for the outer planet, although we note that the models were fit to the unbinned data.

Table 2

Median Values and 68% Confidence Intervals for the Stellar Parameters for Kepler-167

Parameter	Description	Values
Informative Priors		
T_{eff}	Effective temperature (K)	(4830, 100)
[Fe/H]	Metallicity (dex)	(- 0.02, 0.09)
ϖ	Parallax (mas)	(2.944, 0.018)
A_V	V-band extinction (mag)	(0, 0.4204)
Stellar Parameters from SED-only Fit		
M_*	Mass (M_{\odot})	$0.777^{+0.034}_{-0.031}$
R_*	Radius (R_{\odot})	0.749 ± 0.020
L_*	Luminosity (L_{\odot})	$0.289^{+0.017}_{-0.020}$
F_{bol}	Bolometric flux (cgs)	$8.02 \times 10^{-11+4.7} \cdot 10^{-12}$ $- 5.7 \cdot 10^{-12}$
ρ_*	Density (cgs)	$2.60^{+0.23}_{-0.20}$
$\log g$	Surface gravity (cgs)	$4.579^{+0.027}_{-0.025}$
T_{eff}	Effective temperature (K)	4884^{+69}_{-75}
[Fe/H]	Metallicity (dex)	0.020 ± 0.067
[Fe/H] ₀	Initial metallicity ^a	$0.024^{+0.069}_{-0.067}$
Age	Age (Gyr)	$7.1^{+4.4}_{-4.6}$
EEP	Equal evolutionary phase ^b	339^{+12}_{-28}
A_V	V-band extinction (mag)	$0.277^{+0.098}_{-0.13}$
σ_{SED}	SED photometry error scaling	$1.23^{+0.48}_{-0.32}$
ϖ	Parallax (mas)	2.945 ± 0.018
d	Distance (pc)	339.6 ± 2.1

Notes. See Table 3 in Eastman et al. (2019) for a detailed description of all parameters and default (noninformative) priors beyond those specified here. The (a, b) denotes a normal distribution with mean a and variance b^2 , and (a, b) denotes a uniform distribution over the interval $[a, b]$.

^a Initial metallicity is that of the star when it formed.

^b Corresponds to static points in a star's evolutionary history. See Section 2 of Dotter (2016).

This inflation accounts for systematic uncertainty floors set by imperfect models of stellar evolution (Tayar et al. 2020).

Second, we conducted a joint fit of the Keck-HIRES RVs and the two Kepler transits of Kepler-167e using *exoplanet* (Foreman-Mackey et al. 2021). We allowed for orbital eccentricity and a long-term acceleration (slope) in the RV. For the long-cadence photometry, we numerically integrated the transit model over the appropriate time bin in order to account for the effect of these longer integrations on the shape of the transit light curve. We checked for convergence using both the effective sample size and the Gelman–Rubin statistic,

which we required to be greater than 1000 and less than 1.01 for all parameters, respectively. The resulting parameters and models are provided in Table 3 and Figures 1 and 2.

Finally, we conducted a separate fit to the transit photometry of the inner three planets in the Kepler-167 system. In order to simplify the fit and reduce the convergence time, we fixed the orbital eccentricity of these planets to zero. This is likely a valid assumption for Kepler-167b and c, which both orbit within 0.1 au and either have been tidally circularized or have a sufficiently low eccentricity that the effect on the transit shape is negligible (e.g., Mills et al. 2019). However, Kepler-167d is far enough from these two planets to maintain some orbital eccentricity. Using the same Kepler data set, Kipping et al. (2016) demonstrated an upper limit on eccentricity of 0.12. This indicates that there is no detectable deviation from the expected transit shape for a circular orbit in the Kepler photometry, and we therefore should not introduce any additional error by fixing this planet's eccentricity to zero in our fits. As in the previous fit, we numerically integrated the model to account for the integration time when fitting the long-cadence data. We applied the priors on limb-darkening parameters from our fit to the Kepler-167e photometry, which has a much higher S/N than that of the inner super-Earths. We again gauged convergence with the effective sample size and the Gelman–Rubin statistic, for which we achieved >1000 and <1.01 for all parameters. The best-fit models are shown along with the transit and RV data in Figures 1 and 2, respectively. The resulting planetary parameters are listed in Table 3. All of the shared planetary parameters between our work and that of Kipping et al. (2016) are consistent at the 2σ level. We note that the median values of all of the planetary radii are slightly larger than those from Kipping et al. (2016) owing to the increase in stellar radii derived from the updated Gaia parallax.

4. A Closer Look at the Kepler-167 Planets

4.1. Kepler-167e in the Context of Other Cold Giants

Our new RV observations allow us to detect the RV signal from Kepler-167e with high statistical significance (6σ). We find a measured mass of $1.01 \pm 0.16 M_J$ and a 3σ upper limit of 0.29 for the orbital eccentricity. This new mass measurement for Kepler-167e allows us to place it in the context of the broader population of long-period gas giant planets ($>0.5 M_J$) from RV surveys. We focus our comparison on the sample of planets detected in the California Legacy Survey (CLS; Rosenthal et al. 2021), shown in Figure 3. Although there

Table 3
Median Values and 68% Confidence Interval of the Parameters for Each of the Kepler-167 Planets

Parameter	Description	Values			
Planetary Parameters		b	c	d	e
P	Period (days)	$4.3931539^{+0.0000048}_{-0.0000046}$	7.406106 ± 0.000010	$21.80379^{+0.00013}_{-0.00018}$	$1071.23205^{+0.00059}_{-0.00058}$
T_C	Time of conjunction (BJD_{TDB})	$2,455,831.78065 \pm 0.00039$	$2,455,552.15797^{+0.00070}_{-0.00071}$	$2,455,669.7888^{+0.0028}_{-0.0023}$	$2,455,253.28756 \pm 0.00039$
R_p	Radius (R_E)	1.718 ± 0.070	1.674 ± 0.069	1.238 ± 0.064	10.16 ± 0.42
a	Semimajor axis (au)	0.04825 ± 0.00070	0.0684 ± 0.0010	0.1404 ± 0.0020	1.883 ± 0.027
i	Inclination (deg)	$88.3^{+1.6}_{-1.2}$	$88.48^{+0.88}_{-1.0}$	89.26 ± 0.50	$89.9720^{+0.0069}_{-0.0079}$
b	Transit impact parameter	$0.41^{+0.38}_{-0.28}$	$0.52^{+0.30}_{-0.35}$	0.52 ± 0.35	$0.271^{+0.051}_{-0.073}$
Depth	Transit depth	0.0004407 ± 0.0000078	$0.0004187^{+0.0000090}_{-0.0000089}$	0.000229 ± 0.000015	$0.01540^{+0.00027}_{-0.00032}$
T_{eq}	Equilibrium temperature ^a (K)	918 ± 27	771 ± 23	538 ± 16	134.4 ± 4.0
e	Eccentricity ^b	0	0	0	<0.29
ω_*	Argument of periastron (rad)	$-0.4^{+2.0}_{-1.0}$
K	RV semiamplitude (m s^{-1})	$23.7^{+3.7}_{-3.5}$
M_p	Mass (M_J)	$1.01^{+0.16}_{-0.15}$
ρ_p	Density (g cm^{-3})	$1.68^{+0.34}_{-0.33}$
g	RV slope ($\text{m s}^{-1} \text{ day}^{-1}$)	$-0.0049^{+0.0068}_{-0.0066}$
Kepler Parameters					
u_1	Linear limb-darkening coefficient	$0.616^{+0.036}_{-0.034}$			
u_2	Quadratic limb-darkening coefficient	0.126 ± 0.096			
Keck-HIRES Parameters					
γ_{rel}	Relative RV offset (m s^{-1})	$-2.49^{+0.95}_{-0.92}$			
σ_J	RV jitter (m s^{-1})	$3.1^{+2.5}_{-1.6}$			

Notes.

^a Assumes a Bond albedo of 0.3 for Kepler-167e and 0 for all other planets. Assumes perfect heat redistribution.

^b Fixed to 0 for all planets except Kepler-167e, for which we report the 3σ upper limit.

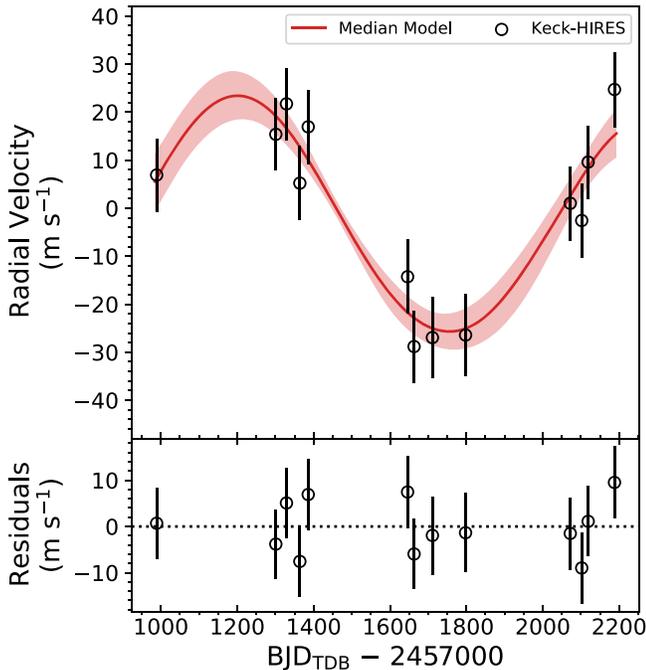


Figure 2. Keck-HIRES RV time-series data with the median model for Kepler-167e and the corresponding 68% credible interval overplotted.

have been several other long-term RV surveys capable of detecting Jupiter analogs (e.g., Rowan et al. 2016; Wittenmyer

et al. 2016, 2020), the CLS includes data sets for more than 700 stars with baselines spanning close to three decades, making it one of the largest and most complete surveys of Jupiter analogs published to date. This allows us to obtain a (relatively) unbiased sample of long-period gas giant planets both with and without inner super-Earth companions. We mark giant planets from the Rosenthal et al. (2021) sample that harbor an inner companion ($<0.1 M_J$) in black. We additionally supplement this sample with the set of (RV) giant planet companions to transiting super-Earths with resolved orbits in Bryan et al. (2019). Although Bryan et al. (2019) had much better sensitivity to small inner companions compared to CLS, the systems in their sample have shorter RV baselines, and the companion orbits are therefore biased toward shorter orbital periods than in the CLS sample.

We find that the mass and eccentricity of Kepler-167e are fairly typical of other long-period giant planets, both with and without detected inner super-Earths. Despite the faintness of the host star and the relative sparseness of our RV data, our joint fit with the transit data results in constraints on mass and eccentricity that are comparable in precision to those of the nontransiting planets detected in these RV surveys. We find that the sample of RV planets spans a wide range of masses at Kepler-167e's location, although these data are relatively insensitive to planets smaller than $\sim 0.5 M_J$ (Rosenthal et al. 2021) at these large separations. While Kepler-167e's mass may be typical of planets at these separations, it appears to have a relatively low orbital eccentricity. Although it is possible that a higher orbital eccentricity might have destabilized the system

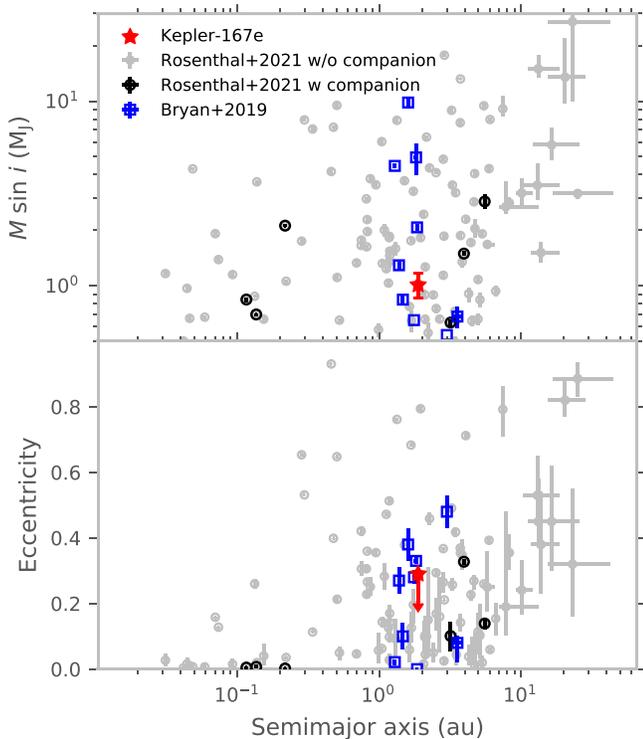


Figure 3. Planet mass and eccentricity vs. semimajor axis for planets more massive than $0.5 M_J$ in the CLS sample (Rosenthal et al. 2021). Giant planets in this sample are shown in two different colors, depending on whether or not they have a detectable inner companion less than $0.1 M_J$. We also plot resolved giant planet companions of super-Earth systems studied in Bryan et al. (2019). Kepler-167e is marked in red (3σ upper limit for eccentricity) and has properties that are typical of long-period giant planets.

of inner super-Earths, we note that there are multiple examples of eccentric gas giants in the Bryan et al. (2019) sample with inner transiting super-Earths.

The fact that Kepler-167e also transits its host star provides us with a unique opportunity to use its measured mass, radius, and age to constrain its bulk metallicity and absolute metal content. Since Kepler-167e has a mass that is indistinguishable from that of Jupiter but 10% smaller than Jupiter in radius, we can immediately surmise that the planet has a higher bulk metal content. We use the giant planet interior structure and evolution model described in Thorngren et al. (2016) and Thorngren & Fortney (2019) to quantify the corresponding bulk metallicity for Kepler-167e. In this approach, a Bayesian statistical model is used to infer a planet’s bulk metallicity from its mass, age, and radius. The planetary mass and age observations are used as priors, and the planetary radius is the independent measurement that yields an estimate of the bulk metallicity. For more details on the giant planet interior structure model and the Bayesian method used to make these inferences, we refer the reader to Thorngren & Fortney (2019). The only difference between the models described there and the ones used here is that we updated the equation of state for hydrogen and helium to the one given in Chabrier et al. (2019). Figure 4 shows the resulting posterior probability distribution for bulk metallicity and its covariance with the observational priors on planetary mass and age.

With a bulk metallicity of 0.21 ± 0.05 , Kepler-167e is significantly enriched in metals relative to its host star, which has a metallicity $[\text{Fe}/\text{H}]$ of 0.02 ± 0.07 (bulk metallicity $Z_* = 0.015 \pm 0.003$ assuming solar $Z = 0.014$;

Asplund et al. 2009). A bulk metallicity of 0.21 ± 0.05 and planet mass of $1.01^{+0.16}_{-0.15} M_J$ translates to an absolute metal content of $66^{+20}_{-18} M_\oplus$. Kepler-167e has a bulk metallicity that is fairly typical for gas giant planets in this mass range; transiting planets with measured bulk metallicities in the mass range $0.85\text{--}1.17 M_J$ have a median bulk metallicity of 0.22 (Thorngren & Fortney 2019). Although we cannot tell how the metals are partitioned between the core and the envelope, the total metal content of the planet nonetheless provides a useful constraint on the solid inventory of the disk that Kepler-167e formed in. We discuss the implications of this measurement in more detail in Section 5.1.

To date, only four other transiting giant planets with orbital periods >200 days have precisely measured masses: the circumbinary planets Kepler-16b (Doyle et al. 2011) and Kepler-34b (Welsh et al. 2012), Kepler-1514b (Dalba et al. 2020), and Kepler-1704b (Dalba et al. 2021). Among the planets around single stars, Kepler-1704b is a massive giant ($4.15 M_J$) on a highly eccentric ($e = 0.92$) orbit and therefore unlikely to harbor any undetected inner companions. Kepler-1514b is also quite massive ($5.3 M_J$) and has a high eccentricity ($e = 0.4$), but it is accompanied by a single transiting inner super-Earth. Kepler-167e is the only planet in this sample that has multiple transiting inner super-Earths. Both Kepler-1514b ($1.11 \pm 0.02 R_J$) and Kepler-1704b ($1.07 \pm 0.04 R_J$) are larger than Jupiter in size, and therefore their bulk metallicities ($0.06^{+0.03}_{-0.02}$ and 0.12 ± 0.04 , respectively) are lower than Kepler-167e’s. However, given their large masses, the bulk metallicities of Kepler-1514b and Kepler-1704b translate to ~ 100 and $\sim 160 M_\oplus$, respectively, in absolute metal content.

4.2. How Massive Are the Inner Super-Earths?

Our RV data set has relatively sparse sampling and is therefore not very sensitive to the RV signals of the three inner super-Earths. We quantify the expected RV semiamplitudes for each planet by using the nonparametric M – R relationship for Kepler planets from Ning et al. (2018) to calculate predicted masses for these planets using their measured radii. We use the `mr-exo` package to obtain posteriors for the masses of the super-Earths using normal distributions for the radii with mean values and distribution widths from Table 3 (Ning et al. 2018; Kanodia et al. 2019). The predicted planet masses for Kepler-167b, c, and d are $4.5^{+6.5}_{-2.6}$, $4.4^{+6.3}_{-2.6}$, and $3.6^{+5.2}_{-2.1} M_\oplus$, respectively. These uncertainties are dominated by the relatively large measurement errors and correspondingly large intrinsic scatter for planets with measured masses in this size range. For median mass estimates of these three planets, we would expect RV semiamplitudes of 2.1, 1.7, and 1.0 m s^{-1} , respectively, which are well below the noise floor of our data.

In order to comment on the solid budget required to form the inner super-Earths, we also need to estimate the total amount of solids contained in these planets. In Figure 5, we show the posterior for the total mass contained in the three planets ($15.7^{+11.6}_{-6.5} M_\oplus$), as well as the total mass estimates for median planetary radii, assuming they are made of pure rock or iron (Zeng et al. 2019). The measured radii and orbital periods of these planets place them at or below the location of the radius valley (Fulton et al. 2017). It is therefore unlikely that they host significant hydrogen-rich envelopes (Rogers 2015; Owen & Wu 2017; Ma & Ghosh 2021). The $15 M_\oplus$ peak of the posterior probability distribution is equivalent to the predicted value for

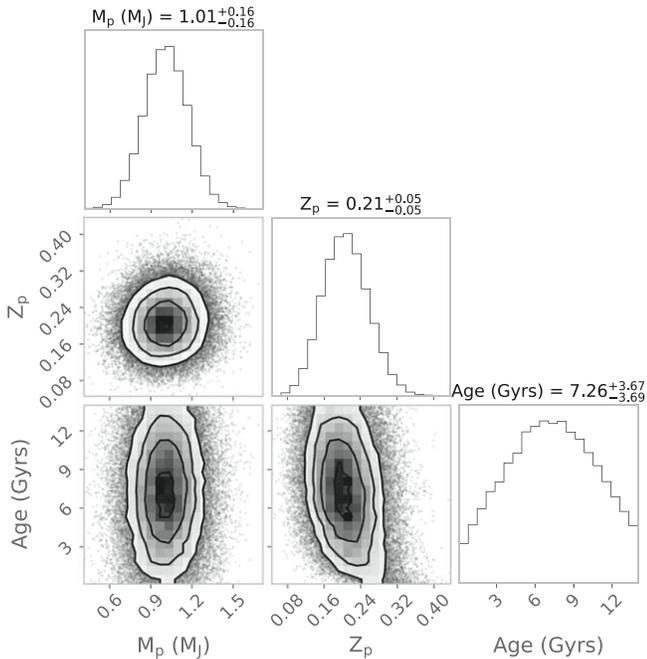


Figure 4. Posterior probability distribution for bulk metallicity and its covariance with the observational priors on planetary mass and age. Kepler-167e’s radius, mass, and age are used to infer its bulk metal content from planetary evolution modeling as described in Thorngren et al. (2016) and Thorngren & Fortney (2019).

Earth-like rock–iron compositions, and we therefore adopt it as our baseline value for all subsequent calculations. How does this mass compare with the solid-mass budget in the inner disk? Since the disk density profiles are poorly constrained by the observations, we use the minimum-mass solar nebula (MMSN) and minimum-mass extrasolar nebula (MMEN) as baselines to estimate the dust content of the inner disk (Chiang & Youdin 2010; Chiang & Laughlin 2013; Dai et al. 2020). The MMSN and MMEN predict 7 and 36 M_{\oplus} of solids within the orbit of the giant planet, respectively. Even in the more optimistic MMEN, the formation of Kepler-167’s super-Earths would require dust to be converted to planets with a fairly high efficiency of 40%. The predicted efficiency of converting dust to super-Earths by either pebble accretion or planetesimal accretion is instead expected to be 10%–20% (Drażkowska et al. 2016; Liu & Ormel 2018; Ormel & Liu 2018; Lenz et al. 2020). Moreover, dust in the inner disk is likely to drift into the star on very short timescales. This suggests that the initial solid budget of the inner disk was very likely supplemented by the addition of small solids from regions exterior to Kepler-167e’s orbit, which could have migrated inward via radial drift. We explore this scenario in more detail below.

5. Formation of Inner Super-Earths with Outer Gas Giant Companions

We use our observational constraints on the properties of the Kepler-167 planets to explore potential formation scenarios for this system. In particular, we are interested in how the presence of a growing giant planet core affects the dust distribution in the disk, since the dust content of the inner disk determines the potential for close-in super-Earth formation. For the giant planet core, we assume that it grows by accreting the marginally coupled

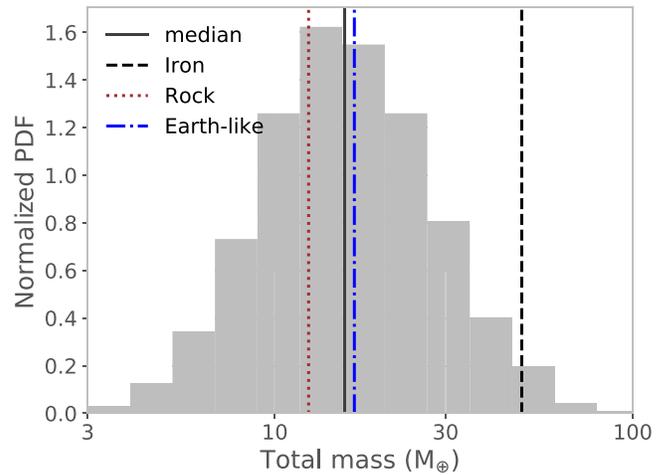


Figure 5. Posterior for the total mass contained in the three inner super-Earths obtained using `mr-exo` (Kanodia et al. 2019), which utilizes the M – R relationship from Ning et al. (2018). The median of the distribution is shown with a gray line. We also mark the total mass contained in three planets assuming they are pure rock (MgSiO_3 ; brown), iron (black), or Earth-like (blue) using the median radii of these planets from Table 3 and M – R relations from Zeng et al. (2019).

“pebbles.” In the pebble accretion paradigm, the growth of the giant planet core depends on the pebble flux through the disk, and a threshold pebble flux is typically required to form a sufficiently large core prior to the dispersal of the gas disk (e.g., Bitsch et al. 2019). We do not consider planetesimal accretion for the formation of the cold giant planet’s core (e.g., Schlecker et al. 2021) because it is highly inefficient at the relevant orbital distances unless the planetesimals are assumed to be small and the turbulent stirring is assumed to be very weak (Johansen & Bitsch 2019). For the inner super-Earths, both pebble and planetesimal accretion appear to be feasible. However, super-Earth progenitors are likely to dynamically evolve and merge after reaching pebble/planetesimal isolation masses (Dawson et al. 2016; Lambrechts et al. 2019), which significantly complicates their formation modeling. We therefore do not explicitly model their formation and instead impose the condition that the amount of solids that reaches the inner disk must be sufficient to form a system of close-in super-Earths (see Section 5.3.2).

In order to understand the formation of the Kepler-167 system, we must therefore first understand the dynamical evolution of solids throughout the disk, which determines the local pebble flux. These pebbles could be directly accreted by the growing protoplanet or could form planetesimals. The pebble flux is very sensitive to the assumed protoplanetary disk properties, such as disk mass, size, metallicity, and turbulence, as well as material properties, such as the fragmentation velocity of grains (e.g., Drażkowska et al. 2021). Since our knowledge of these properties is incomplete, we explore a broad parameter space of potential disk models. Although these models are motivated by a desire to explain the origin of the Kepler-167 system, we do not make any star-specific assumptions other than the stellar mass. This means that the models presented here are broadly applicable to all Sun-like stars.

5.1. Protoplanetary Disk Model

We utilize a simple two-population dust evolution model (Birnstiel et al. 2012) as implemented in the publicly available

two-poppy code¹⁷ to determine which disks are most conducive for giant planet core formation and calculate the amount of solids that reaches the inner disk. This model is described in Birnstiel et al. (2012), who demonstrated that the dust evolution in state-of-the-art numerical simulations is well approximated by splitting the dust population into two groups: one with a spatially and temporally constant size a_0 (assumed monomer size = $0.1 \mu\text{m}$, corresponding to the mass-weighted average of the grains in the interstellar medium; e.g., Laor & Draine 1993) and surface density Σ_0 and the other with size a_1 and surface density Σ_1 . The size of the larger grains (a_1) is set by growth, drift, and fragmentation and varies as a function of time and location in the disk. This approach allows us to model the dust evolution with a simple advection–diffusion equation,

$$\frac{\partial \Sigma_d}{\partial t} + \frac{1}{r} \frac{\partial}{\partial r} \left[r \left(S_d \bar{u} - D_{\text{gas}} \Sigma_g \frac{\partial}{\partial r} \left(\frac{S_d}{S_g} \right) \right) \right] = S_d, \quad (1)$$

where r is the cylindrical distance from the star, $\Sigma_d = \Sigma_0 + \Sigma_1$, \bar{u} is the mass-weighted velocity of the dust grains in the radial direction, Σ_g is the gas density, D_{gas} is the diffusivity of the gas, and S_d is the sink term due to accretion of dust by a growing core. The S_d is related to the pebble accretion rate, which is discussed in more detail in Section 5.2.

The gas disk is assumed to evolve viscously according to the equations in Lynden-Bell & Pringle (1974), and its initial surface density profile is derived from the self-similar solution to these equations at time $t = 0$. The radial dependence of the density profile is set by the power-law exponent p of viscosity ν , which we fix at unity. The viscosity ν is calculated as $\alpha_t c_s H_g$, where α_t is the Shakura–Sunyaev turbulence parameter (Shakura & Sunyaev 1973), c_s is the sound speed, and H_g is the gas scale height. The turbulence parameter α_t and the fragmentation velocity v_{frag} of the grains determine the Stokes number of the largest grains, which is given by $\text{St}_{\text{frag}} = v_{\text{frag}}^2 / 3a_1 c_s^2$, in fragmentation-limited regions of the disk. Since we are primarily interested in understanding the effect of giant planet core formation on the dust flux in the inner disk, we fix $\alpha_t = 10^{-3}$ and $v_{\text{frag}} = 10 \text{ m s}^{-1}$ to reduce the number of free parameters. Although the values of v_{frag} and α_t are both uncertain, these values are reasonably well supported by the literature (e.g., Pinte et al. 2016; Flaherty et al. 2017; Gundlach & Blum 2015). This combination of values also ensures that giant planet cores can reach the pebble isolation mass prior to the dispersal of the gas disk. If we wished to form cores in disks with lower values of v_{frag} , it would require a correspondingly lower α_t (e.g., Venturini et al. 2020).

We explore models for gas disks with characteristic radii of $R_{\text{disk}} \in \{20, 60, 200\} \text{ au}$, masses $M_{\text{disk}} \in \{0.05, 0.1, 0.2\} M_*$, and metallicities $Z \in \{0.005, 0.01, 0.02\}$. This allows us to quantify how the pebble flux and growth rate of the giant planet core vary as a function of each disk parameter. Although we vary M_{disk} and Z as separate model parameters, our results are primarily presented in terms of $M_{\text{dust}} = Z M_{\text{disk}}$, as this quantity plays an important role in controlling the outcome of our models. We fix the mass of the central star to $0.75 M_\odot$, which is representative of Kepler-167. We calculate the temperature profile of the disk by assuming that it is passively irradiated

(e.g., Chiang & Goldreich 1997; D’Alessio et al. 1998); this is a reasonable approximation at the location of our giant planet progenitors (see Section 5.2),

$$T(r) = \left[f T_*^4 \left(\frac{R_*}{r} \right)^2 + T_0^4 \right]^{1/4}, \quad (2)$$

with a flaring angle $\phi = 0.05$, $T_* = 4180 \text{ K}$, $R_* = 2.126 R_\odot$, and $T_0 = 7 \text{ K}$. We obtain the stellar temperature and radius from a MIST model for a $0.75 M_\odot$ star at an age of 1 Myr. Although the inner regions of the disk where super-Earths might form are likely to be viscously heated, this has no effect on the integrated pebble flux that reaches the inner disk from the outer disk.

5.2. Core Formation with Pebble Accretion

To model the formation of a gas giant core, we introduce a seed of mass $0.01 M_\oplus$ at time t_{seed} , which then grows by accreting solids that drift past its location. Although this initial seed mass is somewhat larger than the predicted masses of planetesimals formed by the streaming instability, this choice allows us to circumvent potential complications related to the unknown initial planetesimal distribution and early growth rates of small planetesimals (Johansen et al. 2015; Simon et al. 2016). We vary $t_{\text{seed}} \in \{10^4, 10^5, 10^6\} \text{ yr}$ to study the effect of this assumption on the timescale over which the seed reaches the isolation mass. Our chosen t_{seed} values are motivated by a desire to span a wide range for the time at which a lunar-mass seed might form in the disk. We model the growth of cores with final masses of $\in \{10, 15, 20\} M_\oplus$. We do not consider more massive cores, even though our model fits indicate that Kepler-167e contains $66 \pm 19 M_\oplus$ of metals (see Section 4.1), because the formation timescales for such cores become prohibitively long unless the disks are extremely massive. We adopt the expression for M_{iso} from Lambrechts et al. (2014):

$$M_{\text{iso}} = 20 \left(\frac{H_g/r}{0.05} \right)^3 \left(\frac{M_*}{M} \right) M_A. \quad (3)$$

Although there are updated expressions for M_{iso} that account for its dependence on other properties of the disk (Ataiee et al. 2018; Bitsch et al. 2018), they are in reasonably good agreement with the simpler expression we adopt and depend on parameters we keep fixed in our work (e.g., turbulence, pressure gradient). We implicitly determine the location of our seeds via this expression for the pebble isolation mass. Our initial seeds are therefore placed at 3.2, 5.6, and 8.2 au, respectively, in order to produce cores of 10, 15, and $20 M_\oplus$. Although these seeds might alternatively have started farther out and then migrated inward as they grew, the predicted migration rates are uncertain and depend sensitively on local disk properties (e.g., Rafikov 2002; Li et al. 2009; Paardekooper et al. 2010; Benítez-Llambay et al. 2015). We therefore elect to keep the location of each seed fixed in our models. This simplifies our calculations and gives us a conservative lower limit on the dust mass that reaches the inner disk, as an inward-migrating core that starts further out will reach the same isolation mass later, thus increasing the amount of solids that reaches the inner disk. We calculate the growth rate of the core

¹⁷ The original code is available at <https://github.com/birnstiel/two-pop-py>, and the modified version used in this paper is available at <https://github.com/y-chachan/two-pop-py/tree/kepler-167>.

as the accretion rate of dust of size a_1 (larger grain population),

$$M = f_{3D} M_{2D}, \quad (4)$$

where M_{2D} is the standard 2D pebble accretion rate in the Hill (shear) regime (Lambrechts & Johansen 2014),

$$M_{2D} = 2 \left(\frac{\min(\text{St}_1, 0.1)}{0.1} \right)^{2/3} R_{\text{Hill}}^2 \Omega_K \Sigma, \quad (5)$$

where R_{Hill} is the core's Hill radius, Ω_K is the orbital frequency at the core's location, and St_1 is the Stokes number of grains of size a_1 . Our assumption of accretion in the Hill regime is justified because the transition mass between the Bondi and Hill regimes (Johansen & Lambrechts 2017) is smaller than our adopted seed mass ($0.01 M_{\oplus}$) through most of our simulated domain (< 8 au; our outermost seed is located only slightly further out at 8.2 au). The factor f_{3D} in Equation (4) accounts for the effect of the relative magnitudes of the dust scale height and the core's R_{Hill} on the accretion rate (Morbidelli et al. 2015),

$$f_{3D} = \min \left(1, \frac{1}{2} \sqrt{\frac{\bar{\rho}}{2}} \left(\frac{\text{St}_1}{0.1} \right)^{1/3} \frac{R_{\text{Hill}}}{H_{d,1}} \right), \quad (6)$$

where $H_{d,1} = H_g \sqrt{a_1 / (\bar{q} + \text{St}_1)}$ is the scale height of the large dust grains (Dubrulle et al. 1995). We allow the core to grow until it reaches M_{iso} and record the corresponding time t_{iso} . We assume that this event effectively truncates the flow of solids to the inner disk.

5.3. Model Results

5.3.1. Which Disks Form Giant Planets?

We determine which of our models are able to successfully form gas giant planets by checking to see which cores reach the isolation mass prior to the dispersal of the gas disk, which we assume occurs at 10 Myr. Figure 6 shows results from the full grid of disk models for $M_{\text{iso}} = 15 M_{\oplus}$. Disks where the core reaches the isolation mass are marked with filled circles, while those where it does not are shown as open circles. The importance of the disk's initial solid reservoir is readily apparent (e.g., Bitsch et al. 2019; Schlecker et al. 2021). If the seed of the giant planet core is introduced early ($t_{\text{seed}} = 10^4$ yr; top panel), its ability to reach the isolation mass is determined by the initial solid mass for all but the most compact disk models.

If the seed is introduced later ($t_{\text{seed}} = 10^5$ yr; middle panel), it can only reach the pebble isolation mass if it is located in a relatively large disk. This is because a larger fraction of solids are distributed further out in larger disks, and it takes correspondingly longer for the solids to drain onto the star. This means that seeds that are introduced later can still accrete enough solids to reach the isolation mass (see also Johansen et al. 2019). For a fixed solid-mass reservoir and disk size, we find that the influence of the total disk mass and dust-to-gas ratio, which we only vary by a factor of a few in these models, is relatively weak. It is the product of disk gas mass and dust-to-gas ratio that really matters.

When the core seed is introduced very late ($t_{\text{seed}} = 10^6$ yr; bottom panel), it does not reach the isolation mass in any of the models in our grid. We conclude that $t_{\text{seed}} \lesssim 10^5$ yr is a requirement for lunar-mass seeds to turn into giant planet cores

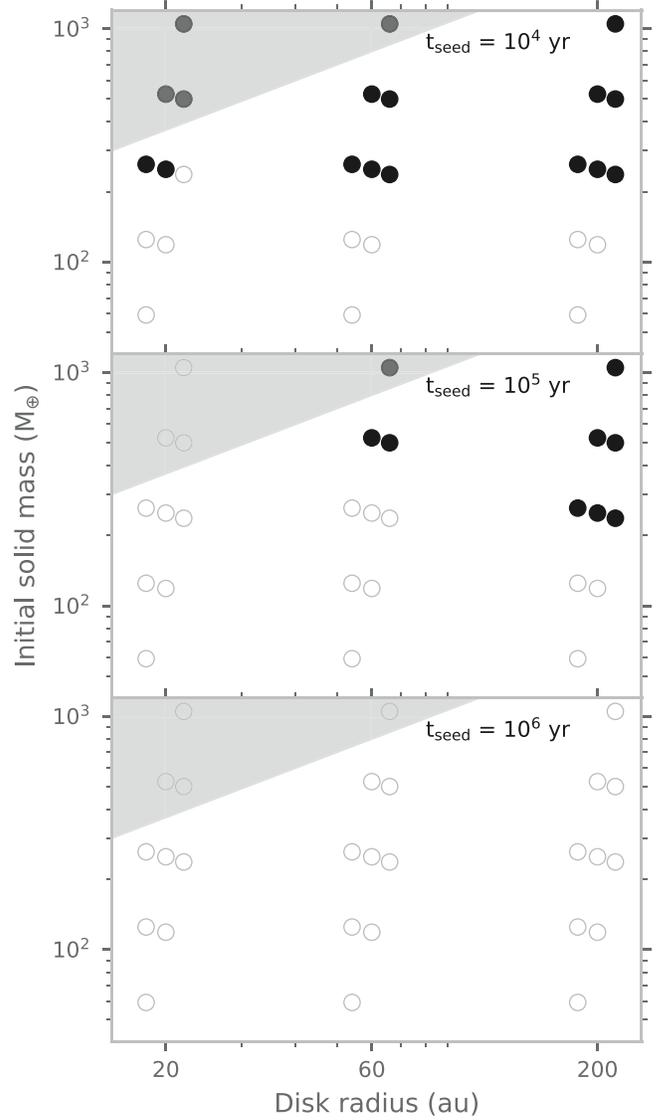


Figure 6. Filled (open) circles mark disk models for which a $0.01 M_{\oplus}$ seed at 5.6 au does (does not) reach the isolation mass of a $15 M_{\oplus}$ core. We show results from models with different initial solid masses M_{dust} (y-axis; product of M_{disk} and Z) and disk sizes (x-axis) and three different t_{seed} . The gray shading indicates combinations of solid mass and disk size that are unlikely to exist in nature. For each combination of disk radius and initial solid mass, we use small offsets to show results for different $M_{\text{disk}} \in \{0.05, 0.1, 0.2\} M_{\star}$ and $Z \in \{0.005, 0.01, 0.02\}$, with M_{disk} increasing in the horizontal direction and Z increasing in the vertical direction.

in the framework considered here. We find that seeds introduced at $\lesssim 10^5$ yr typically reach isolation mass by $\lesssim 1$ Myr (Figure 8). This is consistent with the detection of gaps in protoplanetary disks as young as a few megayears (e.g., Andrews et al. 2018; Long et al. 2018), which are likely opened by planets that are already larger than the pebble isolation mass by this time.

5.3.2. Which Systems with Outer Gas Giant Planets Also Form Inner Super-Earths?

When the giant planet core reaches the pebble isolation mass, the solid reservoir available for planet formation interior to the giant planet's orbit becomes effectively isolated from the outer disk. Although there may still be a modest inward flux of

dust across the gap opened by the planet, it is expected to be a few orders of magnitude smaller than the flux in a smooth disk (e.g., Morbidelli & Nesvorny 2012; Lambrechts et al. 2014; Drazkowska et al. 2019). This means that the mass budget for planet formation in the inner disk is simply the sum of the initial solid reservoir and the cumulative amount of solids delivered from the outer disk before the gas giant core reaches the isolation mass. The initial solid reservoir in the inner disk is typically negligible compared to the flux from the outer disk for all but the smallest disks.

The next question that arises is: how much solid mass needs to be delivered to the inner disk for close-in super-Earths to form? The required mass depends on how super-Earths are formed and is likely to be model-dependent. We adopt two illustrative limits from the pebble and planetesimal accretion paradigms that provide us with useful estimates of the dust mass needed to form super-Earths. Assuming super-Earths form by accretion of “dry” pebbles onto lunar-mass seeds, Lambrechts et al. (2019) showed that an integrated pebble flux $\gtrsim 190 M_{\oplus}$ is necessary to form systems of super-Earths with masses and orbital architectures comparable to those observed by Kepler. In their models, a factor of 2 increase in the pebble flux (from 100 to 200 $M_{\oplus} \text{ Myr}^{-1}$) changes the final outcome from widely spaced terrestrial planets to compact systems of close-in super-Earths. Accounting for 50% mass loss for pebbles across the water snow line, a higher pebble accretion efficiency of our prescription, and the inverse dependence of the accretion efficiency on stellar mass (see Appendix A), we modify this threshold to $190/0.5 \times 3/10 \times 0.75 \sim 86 M_{\oplus}$.¹⁸ This modification is likely to be imperfect because the pebble-mass threshold is sensitive to various time- and space-dependent quantities. Nonetheless, the key point is to compare the pebble-mass threshold for super-Earth formation and the accretion efficiency of a cold giant planet core in the same framework, which we endeavor to do in our study. A less (more) efficient pebble accretion prescription would increase (decrease) the threshold mass for super-Earth formation, but it would also increase (decrease) the pebble mass that filters past the cold giant planet core and reaches the inner disk.

Alternatively, super-Earths might form by planetesimal accretion. To quantify the dust mass needed to form super-Earths in this paradigm, we need to know the efficiencies with which (i) dust is converted into planetesimals and (ii) planetesimals are converted into super-Earths. Unfortunately, quantifying the efficiency of converting dust into planetesimals is quite challenging, and there are few estimates in the literature. Here we use the results of Drazkowska et al. (2016), who used global dust evolution models coupled with planetesimal formation by the streaming instability to show that $\sim 23\%$ of their dust mass is converted into planetesimals. The planetesimals in Drazkowska et al. (2016) form interior to their location of the water snow line, so we additionally account for 50% mass loss of the pebbles that form these planetesimals. Assuming that planetesimals are converted into planets by mutual collision and growth with 100% efficiency (commonly assumed in this paradigm), the combined mass of $\sim 15 M_{\oplus}$ of the Kepler-167 super-Earths translates to

¹⁸ We note that Lambrechts et al. (2019) quantified the pebble mass required to form super-Earths after lunar-mass seeds had already formed. This pebble-mass threshold does not include the pebble mass required to form the seeds in the first place.

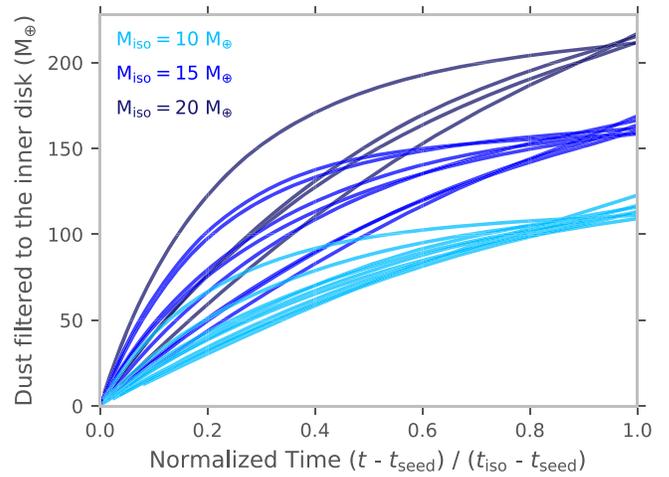


Figure 7. Dust mass that filters through to the inner disk between t_{seed} (the time at which a lunar-mass seed is introduced) and t_{iso} (when the core reaches M_{iso}) for different M_{iso} and a fixed $t_{\text{seed}} = 10^5$ yr. All disk models in which a lunar-mass seed reaches M_{iso} are shown. The filtered dust mass is primarily a function of M_{iso} and does not strongly depend on the assumed disk properties.

$\sim 15/0.23/0.5 \sim 130 M_{\oplus}$ of solids required for formation by planetesimal accretion.

Figure 7 shows the dust mass delivered to the inner disk between t_{seed} and t_{iso} , i.e., while the core is growing, for different disks in which the seed reaches M_{iso} . We find that this integrated dust mass is primarily a function of M_{iso} and depends only weakly on disk properties (disk size and dust mass). It also has a weak dependence on t_{seed} itself, i.e., when the seed is introduced (not shown in the plot). This is because the amount of dust mass filtered through to the inner disk while the core is forming depends on the pebble accretion efficiency ϵ (see Appendix A), and most of the parameters that affect ϵ are constant for our disk models (e.g., α_t , temperature structure). A larger M_{iso} results in the delivery of a larger amount of solids to the inner disk because seeds take longer to reach a larger M_{iso} . This is because the larger M_{iso} places the seed at a larger orbital separation, where the disk aspect ratio is higher and pebble accretion efficiency is lower. The amount of dust mass delivered to the inner disk between t_{seed} and t_{iso} is $\sim 10 \times M_{\text{iso}}$, which implies a cumulative $\epsilon \sim 10\%$ for our disk models. With such efficiencies, the dust mass delivered between t_{seed} and t_{iso} alone is enough to form inner super-Earths via planetesimal or pebble accretion for $M_{\text{iso}} \gtrsim 10 M_{\oplus}$. This inflowing material is augmented by the initial dust located interior to the giant planet’s orbit, as well as the dust mass delivered before t_{seed} .

Figure 8 shows the pebble isolation time for the outer giant companion and the corresponding total solid mass available in the inner disk for a range of t_{seed} and $M_{\text{iso}} = 15 M_{\oplus}$ (left panel) and a fixed $t_{\text{seed}} = 10^5$ yr with varying M_{iso} (right panel). For a fixed M_{iso} (left panel), the amount of solids that reaches the inner disk generally increases with t_{seed} . This is primarily because of the increase in dust mass supplied to the inner disk by radial drift before t_{seed} , and not because of differences in the dust mass delivered between t_{seed} and t_{iso} . We note that for a given t_{seed} , models with different disk gas masses and dust-to-gas ratios but the same total dust mass have fairly different t_{iso} , even though they allow roughly the same mass of solids to reach the inner disk. Although we consider t_{seed} values as low as 10^4 yr, we find that there are many potential disk models with enough solids to form super-Earths. This implies that no

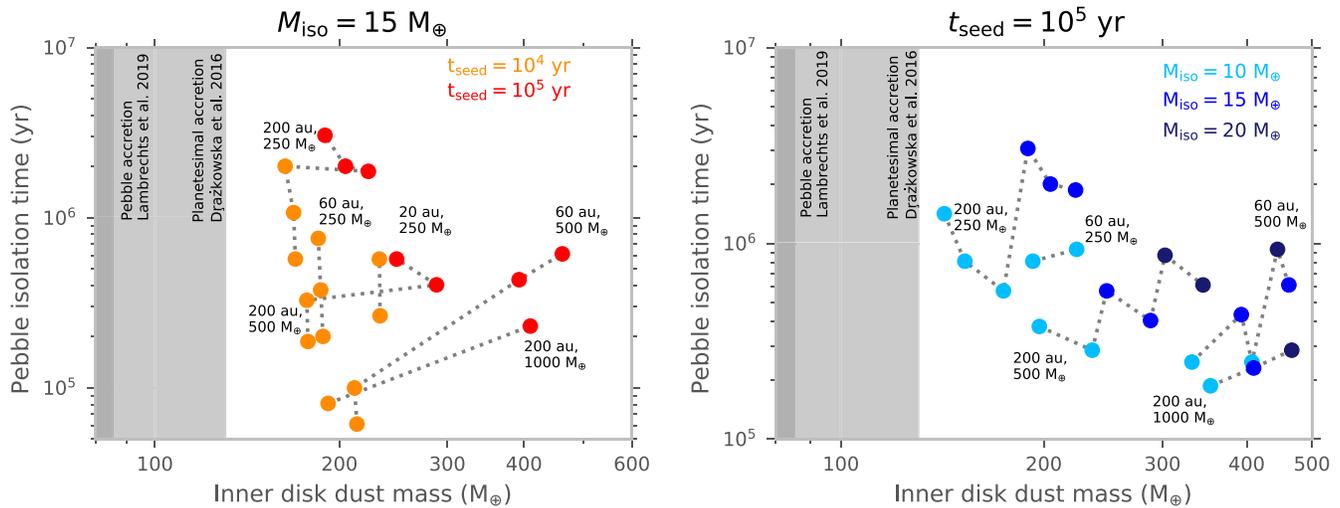


Figure 8. Pebble isolation time vs. total solid mass available in the inner disk for all disk models in which a $0.01 M_{\oplus}$ seed reaches M_{iso} (i.e., those containing an outer gas giant). In the left panel, we vary t_{seed} and fix M_{iso} to $15 M_{\oplus}$. In the right panel, t_{seed} is fixed to 10^5 yr, and M_{iso} is varied. The initial seeds are placed at 3.2, 5.6, and 8.2 au to produce cores of 10, 15, and $20 M_{\oplus}$, respectively. We mark the estimated dust masses that are required for super-Earth formation in the pebble (Lambrechts et al. 2019) and planetesimal (Drazkowska et al. 2016) accretion paradigms using shaded regions. Models with the same disk size and initial solid mass are connected via dotted lines.

temporal fine-tuning in the giant planet core’s formation is necessary in order to enable the formation of inner super-Earths. For a fixed t_{seed} (10^5 yr in the right panel of Figure 8), a larger M_{iso} results in the availability of a larger amount of solids for super-Earth formation in the inner disk. In this panel, models with the same M_{iso} but different disk properties have a different total solid mass available in the inner disk due to the disk-dependent contribution of radial drift before t_{seed} .

Variations in disk properties, t_{seed} , and M_{iso} lead to a large range in the dust mass available for planet formation in the inner disk. For the most massive disks, the dust mass supplied to the inner disk can significantly exceed the threshold dust mass required to form super-Earths. This might lead to the formation of inner planets with higher masses. For example, Lambrechts et al. (2019) showed that increasing the total available pebble mass from 190 to $340 M_{\oplus}$ moved the range of planet masses produced from 2–20 to 5–30 M_{\oplus} and increased the mean mass of the planets that form by a factor of 2. The most massive disks may therefore also allow for the formation of massive sub-Neptunes and Neptunes interior to a cold giant planet’s orbit. This could possibly explain how planetary systems with such architectures emerge (e.g., HAT-P-11, Yee et al. 2018; HD 47186, Bouchy et al. 2009).

5.3.3. Constraints on Kepler-167’s Protoplanetary Disk Properties

We can use our grid of disk models to constrain the properties of Kepler-167’s protoplanetary disk. We know that (1) Kepler-167e’s core must reach the isolation mass well before the dissipation of the disk, (2) enough solids must be delivered to the inner disk prior to this point to allow for super-Earth formation, and (3) after Kepler-167e’s core reaches the isolation mass, there must be enough solids still present beyond its orbit to account for its remaining bulk metal content ($\sim 66 M_{\oplus} - M_{\text{iso}}$). By taking these three conditions into account, we can place a lower limit on the initial dust mass of the disk as a function of disk size. For condition 1, we adopt a stricter limit of 1 Myr rather than our prior 10 Myr for t_{iso} , as we know that Kepler-167e had enough time to accrete a relatively massive (i.e., Jupiter-like) gaseous envelope. This limit is also in better

agreement with the observational constraints on the average disk lifetimes for isolated Sun-like stars, which are around 3 Myr (e.g., Mamajek 2009; Williams & Cieza 2011). Our new upper limit on t_{iso} excludes scenarios with a t_{seed} of 10^6 , leaving us with a choice between a t_{seed} of 10^4 and 10^5 yr. However, for this exercise, we only use $t_{\text{seed}} = 10^5$ yr, as 10^4 yr is likely too early for a lunar-mass seed to form (Lambrechts & Johansen 2012; Visser & Ormel 2016; Johansen & Lambrechts 2017). For condition 2, we adopt the higher limit of $130 M_{\oplus}$ for the dust mass required for super-Earth formation that is imposed by planetesimal accretion (Drazkowska et al. 2016).

Given the sensitivity of gas accretion rates to core mass (e.g., Lee 2019), we also limit our models to M_{iso} of 15 and $20 M_{\oplus}$, which are more likely to produce a Jovian-mass planet. We note that our results are not qualitatively different for an M_{iso} of $10 M_{\oplus}$. Although this requires the giant planet core to accrete additional solids after reaching M_{iso} in order to match the bulk metal content of Kepler-167e ($\sim 66 M_{\oplus}$), this is a more plausible scenario than models in which the pebble isolation mass is set to $66 M_{\oplus}$. Cores of this size can only form in the most massive and largest disk in our grid ($\sim 1000 M_{\oplus}$ solids, disk size of 200 au). Since Kepler-167e’s bulk metal content is typical for planets in its mass range (see Section 4.1 and Thorngren & Fortney 2019), it seems unlikely that all of these giant planets formed with such a large M_{iso} . Planets that have reached the pebble isolation mass may continue to accrete solids in the form of planetesimals (as suggested for Jupiter; Alibert et al. 2018), or they might accrete the pebbles that grow from the dust present the circumplanetary disks (e.g., Drazkowska & Szulágyi 2018). Indeed, Thorngren et al. (2016) argued that the late-stage accretion of planetesimals is needed in order to explain the mass–metallicity relation observed for extrasolar giant planets. We do not model this process explicitly here but simply require that the remaining solid content at orbital separations beyond the giant planet’s core is equal to or greater than $66 M_{\oplus} - M_{\text{iso}}$ at the time when the core reaches the isolation mass.

The initial dust mass of the disk is a product of the disk dust-to-gas ratio and gas mass. In Section 5.3.2, we showed that

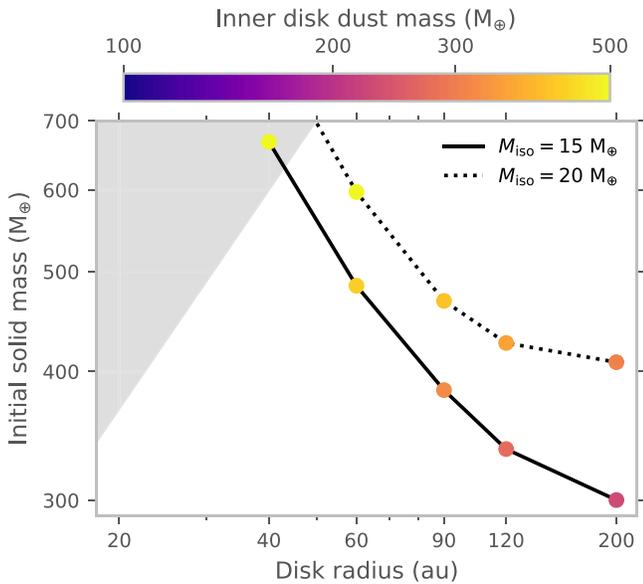


Figure 9. Initial solid mass and size of protoplanetary disks that can produce the Kepler-167 planetary system assuming $t_{\text{seed}} = 10^5$ yr. The color of the points indicates the total amount of solids that is available in the inner disk. We find that $\geq 165 M_{\oplus}$ of solids reach the inner disk for all of our models; thus, they all exceed the super-Earth formation threshold. We gray out the region corresponding to small disks with very large solid masses, as these disks are unlikely to exist in practice.

varying the disk dust-to-gas ratio and gas mass while keeping the total dust mass constant does not affect the solid mass that reaches the inner disk. We therefore reduce the dimensionality of our original grid by fixing the dust-to-gas ratio to 0.015, taking the median stellar $[\text{Fe}/\text{H}] = 0.02$ and assuming solar $[\text{Fe}/\text{H}] = 0.014$ (Asplund et al. 2009). We are left with a grid in which we vary M_{iso} , t_{seed} , disk size, and disk gas mass. We then identify the subset of models in this grid that fulfill the three conditions listed above. In practice, we find that the second condition ($\gtrsim 130 M_{\oplus}$ supplied to the inner disk) is automatically met when the first and third conditions are satisfied.

Figure 9 shows the resulting constraints on the size and initial solid mass of Kepler-167’s protoplanetary disk. We find that we require an initial solid mass larger than $\sim 300 M_{\oplus}$ and a radius larger than 40 au in order to explain this system’s present-day properties. Within this range, disks with a larger M_{iso} require higher initial solid masses in order to form Kepler-167e. For our chosen $t_{\text{seed}} = 10^5$ yr, the requisite dust mass rises sharply with decreasing disk size. This is primarily driven by the need to have sufficient solid mass beyond the giant planet to explain its bulk metal content (condition 3). Since smaller disks have shorter radial drift timescales and dust rapidly drains out of their outer regions, they need to have larger dust masses to meet this requirement.

We next consider whether or not these constraints are consistent with results from protoplanetary disk surveys. In Figure 10, we plot the Atacama Large Millimeter/submillimeter Array (ALMA) and Very Large Array (VLA) disk radii and dust masses estimated for Class 0 and I sources in the Orion cluster (Tobin et al. 2020) and compare them to the theoretical constraints from our models. Since we are interested in the initial dust mass and size, we exclude Class II disks, which show significant signs of processing, especially for dust mass (e.g., Tychoniec et al. 2020). If we consider the VLA and ALMA measurements in isolation, we find that very few disks lie above

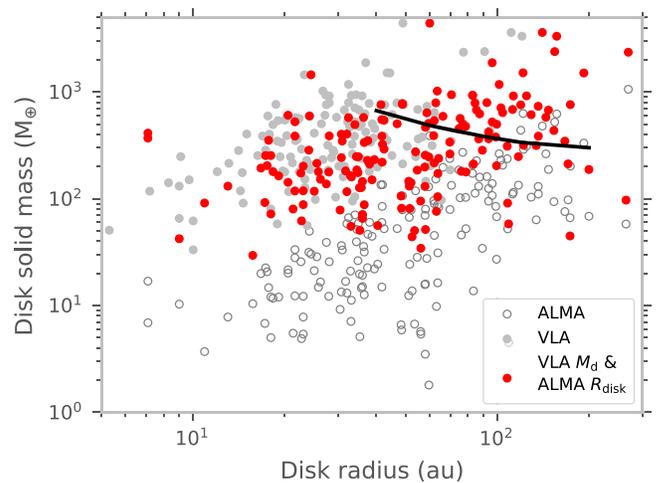


Figure 10. Disk dust mass and radius estimates for Class 0 and I sources in the Orion cluster that are detected with both ALMA (0.87 mm) and the VLA (9 mm; Tobin et al. 2020). We plot the threshold contour above which disks can form systems like Kepler-167 (corresponding to $M_{\text{iso}} = 15 M_{\oplus}$ and the $t_{\text{seed}} = 10^5$ yr curve in Figure 9). Since disks tend to be optically thin in the VLA bandpass, dust mass estimates obtained from these observations are closer to the true estimates. However, disk sizes obtained from ALMA are likely closer to the characteristic disk size that is used in our modeling. We therefore plot dust mass estimates from the VLA against disk radii from ALMA in red.

the planet formation threshold contours we have for Kepler-167. However, disk radii estimated from the VLA and disk dust masses estimated from ALMA tend to be underestimates. This is evident when we instead plot dust masses derived from the VLA against disk radii obtained from ALMA for the same disks (shown in red). Doing so moves the ALMA points up in dust mass and the VLA points to larger radii. Dust masses derived from the VLA are likely closer to the true values, as the disks are much more likely to be optically thin at 9 mm than at 0.87 mm. Similarly, since disks tend to appear smaller in continuum emission at larger wavelengths, the radii estimated from ALMA are likely to be closer to the characteristic disk radii that we have in our models. When we combine dust masses from the VLA with disk radii from ALMA, we find that a substantial fraction of the disks meet the threshold dust mass and disk size necessary for the formation of the Kepler-167 system.

The formation threshold for the Kepler-167 system is primarily driven by the properties of Kepler-167e, in particular its bulk metal content and the need to form it early. Since Kepler-167e is fairly representative of giant planets beyond several au around FGK stars, we can roughly quantify the fraction of disks around single FGK stars that lie above our formation threshold (f_{disk}) and compare it with the corresponding occurrence rate of giant planets (e.g., Wittenmyer et al. 2020; Fulton et al. 2021). Of the 425 disks targeted by ALMA in Tobin et al. (2020), 45 disks lie above the $M_{\text{iso}} = 15 M_{\oplus}$ threshold. However, this sample is likely to contain both massive and low-mass stars that will bias our estimate of f_{disk} . Correcting for this contamination, as well as the presence of close companions to FGK stars that likely go undetected in Tobin et al. (2020; see Appendix B for details of this correction), we find that $f_{\text{disk}} \approx 14\%$.

We conclude that it is reasonably probable that a star with Kepler-167’s mass might host a disk with an initial solid mass and radius that lie above the thresholds indicated by our disk models. If we take the giant planet occurrence rate beyond

several au around FGK stars (e.g., Wittenmyer et al. 2020; Fulton et al. 2021) as a proxy for the occurrence rate of Kepler-167-like systems, we find that this value is broadly consistent with our estimated massive disk fraction of 10%–20%. Our models also suggest that there is likely to be a strong correlation between outer gas giants and inner super-Earths, as most disks that met our conditions for giant planet formation also delivered enough material to the inner disk to form short-period super-Earths. We note that massive metal-rich disks are also more likely to form multiple gas giants, which in turn can pump up the eccentricities of the gas giants and destabilize the system of inner super-Earths. That is, postformation dynamical evolution might reduce the strength of the correlation between inner super-Earths and outer gas giants for metal-rich stars with massive disks. Nonetheless, our results are consistent with observational studies, which find a strong empirical correlation between these two populations (Zhu & Wu 2018; Bryan et al. 2019; Herman et al. 2019).

6. Conclusions

The fact that close-in super-Earths often accompany cold giant planets provides us with valuable insights into the planet formation process. Systems such as Kepler-167, which can be characterized in detail, serve as an important bridge between observed planetary properties and planet formation models. In this work, we refit the Kepler photometry in order to derive updated parameters for both the host star and the four transiting planets in the system. We also obtain RV measurements spanning more than 3 yr in order to measure the mass of the outer transiting gas giant, Kepler-167e. We determine that Kepler-167e is a true Jupiter analog with a mass of $1.01_{-0.15}^{+0.16} M_J$. Its mass and semimajor axis are typical of gas giant planets detected by RV surveys, but it appears to have a relatively low orbital eccentricity (3σ upper limit of 0.29).

We fit Kepler-167e’s measured mass and radius using a giant planet evolution model and find that this planet is more metal-rich than Jupiter, with a bulk metallicity of 0.21 ± 0.05 . This translates to an impressive $66_{-18}^{+20} M_{\oplus}$ of metals in its interior. Although our RV data are not precise enough to place any constraints on the masses of the inner super-Earths, we use the nonparametric M – R relationship from Ning et al. (2018) to estimate that the three planets are predicted to contain $15.7_{-6.5}^{+11.6} M_{\oplus}$ in total. Dust in the inner disk drifts into the star on very short timescales, and converting the local dust content into planets requires a rather high formation efficiency. It therefore seems unlikely that these planets could have formed from the material initially located inside Kepler-167e’s orbit; instead, the dust budget must have been supplemented by the migration of solids from the outer disk. Taken together, these two quantities constrain the initial solid budget of Kepler-167’s protoplanetary disk.

We quantify the conditions required to form the Kepler-167 system by exploring a simple grid of protoplanetary disk models spanning a range of initial solid masses and disk radii. We find that giant planets like Kepler-167e preferentially form in fairly massive (in dust content) and large disks, in good agreement with the results from previous studies (e.g., Bitsch et al. 2019; Schlecker et al. 2021). Our models assume that once the giant planet reaches the isolation mass, the flow of solids to the inner disk is effectively truncated. Despite this constraint, we find that most disks that form outer gas giants are nonetheless able to supply enough solids to the inner disk to

also form super-Earths in both the pebble (Lambrechts et al. 2019) and planetesimal (Drażkowska et al. 2016) accretion paradigms. This remains true regardless of the time at which the giant planet seed is introduced, and we find consistent results across a range of different isolation masses for the giant planet core.

When we incorporate the additional constraint provided by Kepler-167e’s bulk metallicity, we find that we require disks that contain $\gtrsim 300 M_{\oplus}$ of solids and are $\gtrsim 40$ au in size to form this planet. We compare these constraints with the observed properties of Class 0 and I disks in the Orion cluster as measured by ALMA and the VLA (Tobin et al. 2020). We find that 10%–20% of FGK stars should have disks with masses and radii large enough to form the Kepler-167 system, even after accounting for contamination from a range of stellar populations. This strengthens the plausibility of our constraints on Kepler-167’s protoplanetary disk properties.

Further efforts to characterize the Kepler-167 system will enhance our understanding of the origin of its planetary configuration. In particular, there is a pressing need for mass measurements of the inner planets. Our ignorance of the super-Earth planet masses hinders our ability to estimate the accretion efficiency of pebbles and/or planetesimals. Future observations with next-generation instruments such as the Keck Planet Finder (KPF; Gibson et al. 2016) will allow us to measure the masses of the super-Earths and put our formation scenario on a firmer footing. The RV semi-amplitudes corresponding to an Earth-like planetary composition for Kepler-167b and c are expected to be accessible with the KPF. Additionally, characterization of Kepler-167e’s atmospheric composition would be a useful probe of its formation history and location. However, Kepler-167 is a faint star, and Kepler-167e has a high surface gravity and cold atmosphere and is a good candidate for photochemical hazes. We also cannot stack multiple transits or eclipses because it transits so infrequently. Taken together, these factors mean that it does not appear to be a good target for atmospheric characterization with the James Webb Space Telescope.

Although it is outside the scope of this study, we note that further characterization of Class 0 and I disks would be particularly valuable for bridging the gap between disk properties and planet formation. It is fortuitous that we can obtain observational constraints on disk dust masses and radii, as these two quantities have a significant impact on planet formation. There is a growing consensus in the field that planet formation starts earlier than the Class II stage (e.g., Tychoniec et al. 2020; Segura-Cox et al. 2020), and it would therefore be particularly useful to carry out additional comprehensive surveys targeting other young star-forming regions with a significant population of Class 0 and I disks. This would allow us to more accurately assess the distribution of disk properties at early times, which we can use to make predictions for giant planet occurrence rates. On the modeling end, our Class II disk model is unlikely to be appropriate for the early stages of disk evolution; therefore, more accurate models are needed for these initial epochs. A deeper understanding of the connection between the collapse of protostellar cores and the initial properties of protoplanetary disks, such as disk sizes and the timescale over which dust and gas are delivered, would also help to better elucidate the environment in which planets first begin to form (see Lebreuilly et al. 2020, 2021; Lee et al. 2021; Schib et al. 2021, for recent attempts in this direction).

We are grateful to the referee for a timely, thoughtful, and instructive report that helped improve this paper. The authors would like to express their gratitude to Dan Foreman-Mackey for assistance with the `exoplanet` package. Y.C. is grateful to Michael Greklek-McKeon, Michiel Lambrechts, Eve Lee, Lee Rosenthal, and Shreyas Vissapragada for helpful discussions. The authors thank all of the observers in the California Planet Search team for their many hours of hard work. H.K. acknowledges support from NSF CAREER grant 1555095. P.D. is supported by a National Science Foundation (NSF) Astronomy and Astrophysics Postdoctoral Fellowship under award AST-1903811.

This research has made use of the NASA Exoplanet Archive, which is operated by the California Institute of Technology, under contract with the National Aeronautics and Space Administration under the Exoplanet Exploration Program. Some of the data presented in this paper were obtained from the Mikulski Archive for Space Telescopes (MAST) at the Space Telescope Science Institute. The specific observations analyzed can be accessed via [10.17909/T9059R](https://doi.org/10.17909/T9059R). Funding for the Kepler mission is provided by the NASA Science Mission Directorate. The STScI is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 526555.

Some of the data presented herein were obtained at the W. M. Keck Observatory, which is operated as a scientific partnership among the California Institute of Technology, the University of California, and NASA. The Observatory was made possible by the generous financial support of the W. M. Keck Foundation. Finally, the authors recognize and acknowledge the cultural role and reverence that the summit of Maunakea has within the indigenous Hawaiian community. We are deeply grateful to have the opportunity to conduct observations from this mountain.

Facilities: Keck:I (HIRES), Kepler.

Software: EXOFASTv2 (Eastman et al. 2013, 2019), `lightkurve` (Lightkurve Collaboration et al. 2018), `SpecMatch` (Petigura 2015; Petigura et al. 2017), `SpecMatch-Emp` (Yee et al. 2017), `exoplanet` (Foreman-Mackey et al. 2021), `pymc3` (Salvatier et al. 2016), `theano` (Theano Development Team et al. 2016), `starry` (Luger et al. 2019; Agol et al. 2020), `celerite` (Foreman-Mackey et al. 2017; Foreman-Mackey 2018), `astropy` (Astropy Collaboration et al. 2018), `twopoppy` (Birnstiel et al. 2012).

Appendix A Pebble Accretion Efficiency

The pebble flux threshold determined by Lambrechts et al. (2019) for super-Earth formation is applicable for a solar-mass star. To obtain the integrated pebble flux threshold for super-Earth formation around a less massive star, we need to determine the stellar mass dependence of the pebble accretion efficiency. The pebble accretion efficiency is equal to the pebble accretion rate divided by the radial pebble flux at the growing planet’s location. Depending on whether the pebble scale height is smaller or larger than the Hill radius of the planet (2D or 3D regime), this efficiency ϵ is given by

$$\epsilon_{2D/3D} = \frac{M_{2D/3D}}{2\pi r v_r S_d}, \quad (\text{A1})$$

where $M_{3D} = f_{3D} M_{2D}$ when $f_{3D} < 1$ (see Equations (4)–(6)). In the expression for the radial pebble flux, r is the radial distance

from the star, Σ_d is the surface density of the pebbles (equal to Σ_1 in our two-population model), and $v_r = 2\eta r \Omega_K \text{St}$ is the radial drift velocity of the pebbles. Here $h = -0.5 d \ln P / d \ln r (H_g/r)^2$ is a measure of the deviation of the gas orbital velocity from the Keplerian velocity. Following through, we obtain the following expressions for pebble accretion efficiency in the 2D and 3D regimes:

$$\epsilon_{2D} = \left(\frac{10}{3}\right)^{2/3} \frac{1}{2\pi} \frac{q^{2/3}}{h \text{St}^{1/3}} \gg 0.36 \frac{q^{2/3}}{\text{St}^{1/3}}, \quad (\text{A2})$$

$$\epsilon_{3D} = \left(\frac{5}{6}\right) \frac{1}{\sqrt{2\pi}} \frac{q r}{h H_d} \gg 0.33 \frac{q}{h h_d}. \quad (\text{A3})$$

Here $q = M_p/M_*$ is the mass ratio of the growing planet, and $h_d = H_d/r$. The numerical coefficients and physical dependencies we obtain match those given in Table 2 of Ormel & Liu (2018; listed under Morbidelli et al. 2015 and Lambrechts & Johansen 2014, which matches our prescription). The numerical coefficients are in reasonable agreement with the values determined from 3D simulations in Ormel & Liu (2018; our ϵ_{2D} is 50% higher, and ϵ_{3D} is 15% lower; most of the seeds in our models accrete in 3D, so ϵ_{3D} is the relevant value). Our expression for ϵ_{3D} is larger by a factor of 10/3 compared to the value obtained by Lambrechts et al. (2019).

Super-Earths form in the inner disk and accrete in the 3D regime in Lambrechts et al. (2019), where the pebble accretion efficiency $\epsilon_{3D} \propto 1/M_*$. Therefore, accounting for the higher pebble accretion efficiency (factor of 10/3), the lower stellar mass ($0.75 M_\odot$) in our models, and 50% mass loss due to sublimation of water ice from pebbles, the threshold for super-Earth formation by pebble accretion is roughly $190 \times 3/10 \times 0.75/0.5 \sim 86 M_\oplus$. We note that this is the integrated pebble flux that must reach the inner disk after the super-Earth seeds have formed. This threshold therefore does not include the pebble mass required to form the seeds in the first place. We have also ignored the stellar mass dependence of other quantities in the expression, most notably the disk aspect ratio and the Stokes number of pebbles in the inner disk.

Appendix B Disk Fraction and Formation Threshold

Of the 425 disks targeted by ALMA in Tobin et al. (2020), 45 lie above the $M_{\text{iso}} = 15 M_\oplus$ threshold. However, this sample is likely to contain both massive and low-mass stars that will bias our estimate of f_{disk} . Restricting the observational sample of protostars to a range of masses is notoriously difficult, as protostellar masses are highly uncertain for Class 0 and I sources (e.g., Dunham et al. 2014; Fischer et al. 2017). Instead, we correct for contamination from massive and low-mass stars using the measured initial mass function (IMF) for the Orion nebula. In order to do so, we need to know how well the ALMA sample captures the IMF, including the threshold stellar mass below which we miss most protostellar objects. Given the wide range (10^{-2} – $10^3 L_\odot$) of protostellar luminosities exhibited by the sources in Tobin et al. (2020), it is likely that we are only missing protostars that reach $\lesssim 0.1$ – $0.2 M_\odot$ at the end of accretion (e.g., Offner & McKee 2011; Hartmann et al. 2016). Adopting the IMF from Da Rio et al. (2012) along with the modification suggested by Krumholz et al. (2012) to account

for stars more massive than $2 M_{\odot}$, we find that these stars constitute $\sim 20\%$ of the stellar population.

Therefore, assuming that the ALMA survey likely samples only the top 80% of the IMF, stars more massive than $1.4 M_{\odot}$ should comprise approximately 12.5% of our sample (53 sources). We expect a substantial fraction of disks above our formation threshold to be hosted by these massive stars. However, nearly $\sim 50\%$ of massive stars have a companion within $\log P = 5$ (Moe & Di Stefano 2017), which is associated with lower disk masses and/or lifetimes (Kraus et al. 2012). This means that, at most, half of the massive stellar population should host massive disks in our survey sample. We must also correct for contamination from M stars that lie above the luminosity threshold. We assume that disks around M stars are likely to have lower masses,¹⁹ and we therefore do not expect a substantial number of these disks to lie above our formation threshold. Using the assumption adopted earlier that we only miss the bottom 20% of the IMF, we would expect half the stars in our sample to be $< 0.5 M_{\odot}$ (212 sources). Finally, we expect $\sim 20\%$ of the remaining FGK stars (32 out of the remaining 160) to have close-in companions ($\lesssim 50$ au; Moe & Di Stefano 2017) that we cannot detect,²⁰ and we must therefore remove these from the denominator. Performing all of these steps, we find $f_{\text{disk}} = (45 - 53/2)/(425 - 53 - 212 - 32) \approx 14\%$.

ORCID iDs

Yayaati Chachan  <https://orcid.org/0000-0003-1728-8269>
 Paul A. Dalba  <https://orcid.org/0000-0002-4297-5506>
 Heather A. Knutson  <https://orcid.org/0000-0002-5375-4725>
 Benjamin J. Fulton  <https://orcid.org/0000-0003-3504-5316>
 Daniel Thorngren  <https://orcid.org/0000-0002-5113-8558>
 David R. Ciardi  <https://orcid.org/0000-0002-5741-3047>
 Andrew W. Howard  <https://orcid.org/0000-0001-8638-0320>
 Judah Van Zandt  <https://orcid.org/0000-0002-4290-6826>

References

- Agol, E., Luger, R., & Foreman-Mackey, D. 2020, *AJ*, 159, 123
 Alibert, Y., Venturini, J., Helled, R., et al. 2018, *NatAs*, 2, 873
 Andrews, S. M., Rosenfeld, K. A., Kraus, A. L., & Wilner, D. J. 2013, *ApJ*, 771, 129
 Andrews, S. M., Huang, J., Pérez, L. M., et al. 2018, *ApJL*, 869, L41
 Ansdell, M., Williams, J. P., van der Marel, N., et al. 2016, *ApJ*, 828, 46
 Asplund, M., Grevesse, N., Sauval, A. J., & Scott, P. 2009, *ARA&A*, 47, 481
 Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, *AJ*, 156, 123
 Ataiee, S., Baruteau, C., Alibert, Y., & Benz, W. 2018, *A&A*, 615, A110
 Barbato, D., Sozzetti, A., Desidera, S., et al. 2018, *A&A*, 615, A175
 Batalha, N. M., Rowe, J. F., Bryson, S. T., et al. 2013, *ApJS*, 204, 24
 Benítez-Llambay, P., Masset, F., Koenigsberger, G., & Szulágyi, J. 2015, *Natur*, 520, 63
 Birnstiel, T., Dullemond, C. P., & Brauer, F. 2010, *A&A*, 513, A79
 Birnstiel, T., Klahr, H., & Ercolano, B. 2012, *A&A*, 539, A148
 Bitsch, B., Izidoro, A., Johansen, A., et al. 2019, *A&A*, 623, A88
 Bitsch, B., Morbidelli, A., Johansen, A., et al. 2018, *A&A*, 612, A30
 Borucki, W. J., Koch, D., Basri, G., et al. 2010, *Sci*, 327, 977
 Bouchy, F., Mayor, M., Lovis, C., et al. 2009, *A&A*, 496, 527
 Bryan, M. L., Knutson, H. A., Lee, E. J., et al. 2019, *AJ*, 157, 52
 Buchhave, L. A., Latham, D. W., Johansen, A., et al. 2012, *Natur*, 486, 375
 Chabrier, G., Mazevet, S., & Soubiran, F. 2019, *ApJ*, 872, 51
 Chiang, E., & Laughlin, G. 2013, *MNRAS*, 431, 3444
 Chiang, E., & Youdin, A. 2010, *AREPS*, 38, 493
 Chiang, E. I., & Goldreich, P. 1997, *ApJ*, 490, 368
 Choi, J., Dotter, A., Conroy, C., et al. 2016, *ApJ*, 823, 102
 Ciardi, D. R., Beichman, C. A., Horch, E. P., & Howell, S. B. 2015, *ApJ*, 805, 16
 Cumming, A., Butler, R. P., Marcy, G. W., et al. 2008, *PASP*, 120, 531
 Cutri, R. M., Skrutskie, M. F., van Dyk, S., et al. 2003, *yCat*, 2246, 0
 Cutri, R. M., Wright, E. L., Conrow, T., et al. 2014, *yCat*, 2328, 0
 Da Rio, N., Robberto, M., Hillenbrand, L. A., Henning, T., & Stassun, K. G. 2012, *ApJ*, 748, 14
 Dai, F., Winn, J. N., Schlaufman, K., et al. 2020, *AJ*, 159, 247
 Dalba, P. A., Fulton, B., Isaacson, H., Kane, S. R., & Howard, A. W. 2020, *AJ*, 160, 149
 Dalba, P. A., & Muirhead, P. S. 2016, *ApJL*, 826, L7
 Dalba, P. A., & Tamburo, P. 2019, *ApJL*, 873, L17
 Dalba, P. A., Kane, S. R., Li, Z., et al. 2021, *AJ*, 162, 154
 Dalba, P. A., Kane, S. R., Isaacson, H., et al. 2021, *AJ*, 161, 103
 D'Alessio, P., Cantó, J., Calvet, N., & Lizano, S. 1998, *ApJ*, 500, 411
 Dawson, R. I., Lee, E. J., & Chiang, E. 2016, *ApJ*, 822, 54
 Dotter, A. 2016, *ApJS*, 222, 8
 Doyle, L. R., Carter, J. A., Fabrycky, D. C., et al. 2011, *Sci*, 333, 1602
 Drażkowska, J., Alibert, Y., & Moore, B. 2016, *A&A*, 594, 105
 Drażkowska, J., Li, S., Birnstiel, T., Stammer, S. M., & Li, H. 2019, *ApJ*, 885, 91
 Drażkowska, J., Stammer, S. M., & Birnstiel, T. 2021, *A&A*, 647, 1
 Drażkowska, J., & Szulágyi, J. 2018, *ApJ*, 866, 142
 Dubber, S. C., Mortier, A., Rice, K., et al. 2019, *MNRAS*, 490, 5103
 Dubrulle, B., Morfill, G., & Sterzik, M. 1995, *Icar*, 114, 237
 Dunham, M. M., et al. 2014, in *Protostars and Planets VI*, ed. H. Beuther et al. (Tucson, AZ: Univ. Arizona Press), 195
 Eastman, J., Gaudi, B. S., & Agol, E. 2013, *PASP*, 125, 83
 Eastman, J. D., Rodriguez, J. E., Agol, E., et al. 2019, arXiv:1907.09480
 Fischer, W. J., Megeath, S. T., Furlan, E., et al. 2017, *ApJ*, 840, 69
 Flaherty, K. M., Hughes, A. M., Rose, S. C., et al. 2017, *ApJ*, 843, 150
 Foreman-Mackey, D. 2018, *RNAAS*, 2, 31
 Foreman-Mackey, D., Agol, E., Ambikasaran, S., & Angus, R. 2017, *AJ*, 154, 220
 Foreman-Mackey, D., Morton, T. D., Hogg, D. W., Agol, E., & Schölkopf, B. 2016, *AJ*, 152, 206
 Foreman-Mackey, D., Luger, R., Agol, E., et al. 2021, *JOSS*, 6, 3285
 Fressin, F., Torres, G., Charbonneau, D., et al. 2013, *ApJ*, 766, 81
 Fulton, B. J., Petigura, E. A., Howard, A. W., et al. 2017, *AJ*, 154, 109
 Fulton, B. J., Rosenthal, L. J., Hirsch, L. A., et al. 2021, *ApJS*, 255, 14
 Gaia Collaboration, Brown, A. G. A., Ballenari, A., et al. 2018, *A&A*, 616, A1
 Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2021, *A&A*, 649, A1
 Gibson, S. R., Howard, A. W., Marcy, G. W., et al. 2016, *Proc. SPIE*, 9908, 990870
 Gundlach, B., & Blum, J. 2015, *ApJ*, 798, 34
 Hartmann, L., Hercezeg, G., & Calvet, N. 2016, *ARA&A*, 54, 135
 Herman, M. K., Zhu, W., & Wu, Y. 2019, *AJ*, 157, 248
 Howard, A. W., & Fulton, B. J. 2016, *PASP*, 128, 114401
 Howard, A. W., Johnson, J. A., Marcy, G. W., et al. 2010a, *ApJ*, 721, 1467
 Howard, A. W., Marcy, G. W., Johnson, J. A., et al. 2010b, *Sci*, 330, 653
 Isaacson, H., & Fischer, D. 2010, *ApJ*, 725, 875
 Izidoro, A., Morbidelli, A., Raymond, S. N., Hersant, F., & Pierens, A. 2015, *A&A*, 582, A99
 Jenkins, J. M., Caldwell, D. A., Chandrasekaran, H., et al. 2010, *ApJL*, 713, L87
 Johansen, A., & Bitsch, B. 2019, *A&A*, 631, A70
 Johansen, A., Ida, S., & Brasser, R. 2019, *A&A*, 622, A202
 Johansen, A., & Lambrechts, M. 2017, *AREPS*, 45, 359
 Johansen, A., Mac Low, M.-M., Lacerda, P., & Bizzarro, M. 2015, *SciA*, 1, 1500109
 Kanodia, S., Wolfgang, A., Stefansson, G. K., Ning, B., & Mahadevan, S. 2019, *ApJ*, 882, 38
 Kawahara, H., & Masuda, K. 2019, *AJ*, 157, 218
 Kipping, D. M., Torres, G., Henze, C., et al. 2016, *ApJ*, 820, 112
 Kraus, A. L., Ireland, M. J., Hillenbrand, L. A., & Martinache, F. 2012, *ApJ*, 745, 19
 Krumholz, M. R., Klein, R. I., & McKee, C. F. 2012, *ApJ*, 754, 71

¹⁹ This is at least well supported by observations of Class II disks (e.g., Andrews et al. 2013; Ansdell et al. 2016). For Class 0 and I disks, Tobin et al. (2020) observed a weak correlation between disk mass and protostellar luminosity ($M_{\text{dust}} \propto L_{\text{bol}}^{0.31 \pm 0.05}$ for nonmultiple sources; see their Figure 8), where the latter serves as a relatively poor tracer of stellar mass.

²⁰ Although Tobin et al. (2020) indicated whether or not a set of sources are part of a multiple system, they could only resolve companions separated by $\gtrsim 40$ au. Using the position coordinates, Gaia distances to the sources, and disk radii provided in Tobin et al. (2020), we find that there is likely only a single pair of bona fide binary stars in their sample with separation < 50 au.

- Lambrechts, M., & Johansen, A. 2012, *A&A*, **544**, A32
- Lambrechts, M., & Johansen, A. 2014, *A&A*, **572**, A107
- Lambrechts, M., Johansen, A., & Morbidelli, A. 2014, *A&A*, **572**, A35
- Lambrechts, M., Morbidelli, A., Jacobson, S. A., et al. 2019, *A&A*, **627**, A83
- Laor, A., & Draine, B. T. 1993, *ApJ*, **402**, 441
- Lebreuilly, U., Commerçon, B., & Laibe, G. 2020, *A&A*, **641**, A112
- Lebreuilly, U., Hennebelle, P., Colman, T., et al. 2021, *ApJL*, **917**, L10
- Lee, E. J. 2019, *ApJ*, **878**, 36
- Lee, Y.-N., Charnoz, S., & Hennebelle, P. 2021, *A&A*, **648**, A101
- Lenz, C. T., Klahr, H., Birnstiel, T., Kretke, K., & Stammerl, S. 2020, *A&A*, **640**, A61
- Li, H., Lubow, S. H., Li, S., & Lin, D. N. C. 2009, *ApJL*, **690**, L52
- Lightkurve Collaboration, Cardoso, J. V. M., Hedges, C., et al. 2018, *Lightkurve: Kepler and TESS time series analysis in Python*, Astrophysics Source Code Library, ascl:1812.013
- Lin, J. W., Lee, E. J., & Chiang, E. 2018, *MNRAS*, **480**, 4338
- Lindgren, L., Klioner, S. A., Hernández, J., et al. 2021, *A&A*, **649**, A2
- Liu, B., & Ormel, C. W. 2018, *A&A*, **615**, A138
- Long, F., Pinilla, P., Herczeg, G. J., et al. 2018, *ApJ*, **869**, 17
- Luger, R., Agol, E., Foreman-Mackey, D., et al. 2019, *AJ*, **157**, 64
- Lynden-Bell, D., & Pringle, J. E. 1974, *MNRAS*, **168**, 603
- Ma, Q., & Ghosh, S. K. 2021, *MNRAS*, **505**, 3853
- Mamajek, E. E. 2009, in *AIP Conf. Ser. 1158, Exoplanets and Disks: Their Formation and Diversity*, ed. T. Usuda, M. Tamura, & M. Ishii (Melville, NY: AIP), 3
- Masuda, K., Winn, J. N., & Kawahara, H. 2020, *AJ*, **159**, 38
- Mayor, M., Marmier, M., Lovis, C., et al. 2011, arXiv:1109.2497
- Mills, S. M., Howard, A. W., Weiss, L. M., et al. 2019, *AJ*, **157**, 145
- Moe, M., & Di Stefano, R. 2017, *ApJS*, **230**, 15
- Morbidelli, A., Lamrechts, M., Jacobson, S., & Bitsch, B. 2015, *Icar*, **258**, 418
- Morbidelli, A., & Nesvorný, D. 2012, *A&A*, **546**, A18
- Ning, B., Wolfgang, A., & Ghosh, S. 2018, *ApJ*, **869**, 5
- Offner, S. S. R., & McKee, C. F. 2011, *ApJ*, **736**, 53
- Ormel, C. W. 2017, in *Formation, Evolution, and Dynamics of Young Solar Systems*, Astrophysics and Space Science Library, ed. M. Pessah & O. Gressel, Vol. 445 (Cham: Springer), 197
- Ormel, C. W., & Klahr, H. H. 2010, *A&A*, **520**, A43
- Ormel, C. W., & Liu, B. 2018, *A&A*, **615**, A178
- Ormel, C. W., Liu, B., & Schoonenberg, D. 2017, *A&A*, **604**, A1
- Owen, J. E., & Wu, Y. 2017, *ApJ*, **847**, 29
- Paardekooper, S. J., et al. 2010, *MNRAS*, **401**, 1950
- Paxton, B., Bildsten, L., Dotter, A., et al. 2011, *ApJS*, **192**, 3
- Paxton, B., Cantiello, M., Arras, P., et al. 2013, *ApJS*, **208**, 4
- Paxton, B., Marchant, P., Schwab, J., et al. 2015, *ApJS*, **220**, 15
- Petigura, E. A. 2015, PhD thesis, Univ. California, Berkeley
- 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
- Pinte, C., Dent, W. R. F., Ménard, F., et al. 2016, *ApJ*, **816**, 25
- Rafikov, R. R. 2002, *ApJ*, **572**, 566
- Ricker, G. R., Winn, J. N., Vanderspek, R., et al. 2015, *JATIS*, **1**, 014003
- Rogers, L. A. 2015, *ApJ*, **801**, 41
- Rosenthal, L. J., Fulton, B. J., Hirsch, L. A., et al. 2021, *ApJS*, **255**, 8
- Rowan, D., Meschiari, S., Laughlin, G., et al. 2016, *ApJ*, **817**, 104
- Salvatier, J., Wiecki, T. V., & Fonnesbeck, C. 2016, *PeerJ Comput. Sci.*, **2**, e55
- Santerne, A., Malavolta, L., Kosiarek, M. R., et al. 2019, arXiv:1911.07355
- Schib, O., Mordasini, C., Wenger, N., Marleau, G.-D., & Helled, R. 2021, *A&A*, **645**, A43
- Schlafly, E. F., & Finkbeiner, D. P. 2011, *ApJ*, **737**, 103
- Schlecker, M., Mordasini, C., Emsenhuber, A., et al. 2021, *A&A*, **656**, A71
- Segura-Cox, D. M., Schmiedeke, A., Pineda, J. E., et al. 2020, *Natur*, **586**, 228
- Shakura, N. I., & Sunyaev, R. A. 1973, *A&A*, **24**, 337
- Simon, J. B., Armitage, P. J., Li, R., Youdin, A. N., et al. 2016, *ApJ*, **822**, 55
- Smith, J. C., Stumpe, M. C., Van Cleve, J. E., et al. 2012, *PASP*, **124**, 1000
- Stevenson, D. J. 1982, *AREPS*, **10**, 257
- Stumpe, M. C., Smith, J. C., Van Cleve, J. E., et al. 2012, *PASP*, **124**, 985
- Tayar, J., Clayton, Z. R., Huber, D., & van Saders, J. 2020, arXiv:2012.07957
- Theano Development Team, Al-Rfou, R., Alain, G., et al. 2016, arXiv:1605.02688
- Thorngren, D., & Fortney, J. J. 2019, *ApJL*, **874**, L31
- Thorngren, D. P., Fortney, J. J., Murray-Clay, R. A., & Lopez, E. D. 2016, *ApJ*, **831**, 64
- Tobin, J. J., Sheehan, P. D., Megeath, S. T., et al. 2020, *ApJ*, **890**, 130
- Tychoniec, Ł., Manara, C. F., Rosotti, G. P., et al. 2020, *A&A*, **640**, A19
- Uehara, S., Kawahara, H., Masuda, K., Yamada, S., & Aizawa, M. 2016, *ApJ*, **822**, 2
- Venturini, J., Guilera, O. M., Ronco, M. P., & Mordasini, C. 2020, *A&A*, **644**, A174
- Visser, R. G., & Ormel, C. W. 2016, *A&A*, **586**, A66
- Vogt, S. S., Allen, S. L., Bigelow, B. C., et al. 1994, *Proc. SPIE*, **2198**, 362
- Wang, J., Fischer, D. A., Barclay, T., et al. 2015, *ApJ*, **815**, 127
- Welsh, W. F., Orosz, J. A., Carter, J. A., et al. 2012, *Natur*, **481**, 475
- Williams, J. P., & Cieza, L. A. 2011, *ARA&A*, **49**, 67
- Winn, J. N. 2018, in *Handbook of Exoplanets*, ed. H. J. Deeg & J. A. Belmonte (Cham: Springer), 195
- Wittenmyer, R. A., Butler, R. P., Tinney, C. G., et al. 2016, *ApJ*, **819**, 28
- Wittenmyer, R. A., Wang, S., Horner, J., et al. 2020, *MNRAS*, **492**, 377
- Wright, J. T., Marcy, G. W., Butler, R. P., & Vogt, S. S. 2004, *ApJS*, **152**, 261
- Yee, S. W., Petigura, E. A., & von Braun, K. 2017, *ApJ*, **836**, 77
- Yee, S. W., Petigura, E. A., Fulton, B. J., et al. 2018, *AJ*, **155**, 255
- Zeng, L., Jacobsen, S. B., Sasselov, D. D., et al. 2019, *PNAS*, **116**, 9723
- Zhu, W., & Wu, Y. 2018, *AJ*, **156**, 92