

## A 38 Million Year Old Neptune-Sized Planet in the Kepler Field

<sup>2</sup> L. G. BOUMA,<sup>1,2,\*</sup> J. L. CURTIS,<sup>3,4</sup> K. MASUDA,<sup>5</sup> L. A. HILLENBRAND,<sup>1</sup> G. STEFANSSON,<sup>2,†</sup> H. ISAACSON,<sup>6</sup>  
<sup>3</sup> N. NARITA,<sup>7,8,9,10</sup> A. FUKUI,<sup>7,10</sup> M. IKOMA,<sup>11</sup> M. TAMURA,<sup>12,9,11</sup> A. L. KRAUS,<sup>13</sup> E. FURLAN,<sup>14</sup>  
<sup>4</sup> C. L. GNILKA,<sup>15,14</sup> K. V. LESTER,<sup>15</sup> AND S. B. HOWELL<sup>15</sup>

<sup>5</sup> Cahill Center for Astrophysics, California Institute of Technology, Pasadena, CA 91125, USA

<sup>6</sup> Department of Astrophysical Sciences, Princeton University, 4 Ivy Lane, Princeton, NJ 08540, USA

<sup>7</sup> Department of Astronomy, Columbia University, 550 West 120th Street, New York, NY 10027, USA

<sup>8</sup> Department of Astrophysics, American Museum of Natural History, New York, NY 10024, USA

<sup>9</sup> Department of Earth and Space Science, Osaka University, Osaka 560-0043, Japan

<sup>10</sup> Astronomy Department, University of California, Berkeley, CA 94720, USA

<sup>11</sup> Komaba Institute for Science, The University of Tokyo, Tokyo 153-8902, Japan

<sup>12</sup> Japan Science and Technology Agency, PRESTO, Tokyo 153-8902, Japan

<sup>13</sup> Astrobiology Center, Tokyo 181-8588, Japan

<sup>14</sup> Instituto de Astrofísica de Canarias (IAC), 38205 La Laguna, Tenerife, Spain

<sup>15</sup> Division of Science, National Astronomical Observatory of Japan, Tokyo 181-8588, Japan

<sup>16</sup> Department of Astronomy, The University of Tokyo, Tokyo 113-0033, Japan

<sup>17</sup> Department of Astronomy, The University of Texas at Austin, Austin, TX 78712, USA

<sup>18</sup> NASA Exoplanet Science Institute, Caltech/IPAC, Pasadena, CA 91125, USA

<sup>19</sup> NASA Ames Research Center, Moffett Field, CA 94035, USA

(Received 2021 Oct 12; Revised —; Accepted —)

Submitted to AAS Journals

### ABSTRACT

Kepler 1627A is a G8V star previously known to host a  $3.8 R_{\oplus}$  (**Replaced:** mini-Neptune replaced with: planet) on a 7.2 day orbit. The star was observed by the Kepler space telescope because it is nearby ( $d = 329$  pc) and it resembles the Sun. Here we show using Gaia kinematics, TESS stellar rotation periods, and spectroscopic lithium abundances that Kepler 1627 is a member of the  $38^{+6}_{-5}$  Myr old  $\delta$  Lyr cluster. To our knowledge, this makes Kepler 1627Ab the youngest planet with a precise age yet found by the (**Replaced:** main replaced with: prime) Kepler mission. The Kepler photometry shows two peculiarities: the average transit profile is asymmetric, and the individual transit times might be correlated with the local light curve slope. We discuss possible explanations for each anomaly. More importantly, the  $\delta$  Lyr cluster is one of  $\sim 10^3$  coeval groups whose properties have been clarified by Gaia. Many other exoplanet hosts are candidate members of these clusters; their ages can be verified (**Replaced:** through replaced with: with) the trifecta of Gaia, TESS, and ground-based spectroscopy.

**Keywords:** exoplanet evolution (491), open star clusters (1160), stellar ages (1581)

### 1. INTRODUCTION

While thousands of exoplanets have been discovered orbiting nearby stars, the vast majority of

them are several billion years old. This makes it difficult to test origin theories for the different families of planets, since many evolutionary processes are expected to operate on timescales of less than 100 million years.

For instance, the “mini-Neptunes”, thought to be made of metal cores, silicate mantles (Kite et al. 2020), and extended hydrogen-dominated atmospheres, are expected to shrink in size by fac-

Corresponding author: L. G. Bouma  
luke@astro.caltech.edu

\* 51 Pegasi b Fellow

† Henry Norris Russell Fellow

tors of several over their first  $10^8$  years. Specifically, in the models of Owen & Wu (2016) and Owen (2020), the  $\approx 5 M_{\oplus}$  planets start with sizes of  $4 - 12 R_{\oplus}$  shortly after the time of disk dispersal ( $\lesssim 10^7$  years), and shrink to sizes of  $2 - 4 R_{\oplus}$  by  $10^8$  years. While the majority of this change is expected to occur within the first few million years after the disk disperses (Ikoma & Hori 2012), stellar irradiation and internal heat can also power gradual outflows which, if strong enough, can deplete or entirely strip the atmosphere (Lopez et al. 2012; Owen & Wu 2013; Ginzburg et al. 2018).  
**(Added:** The photoevaporative and core-powered outflows are thought to persist for  $\approx 10^8$  to  $\approx 10^9$  years, though the details depend on the planetary masses, the irradiation environments, and the initial atmospheric mass fractions (Owen & Wu 2017; Gupta & Schlichting 2020; Rogers & Owen 2021; King & Wheatley 2021).**)** Discovering young planets, measuring their masses, and detecting their atmospheric outflows are key steps toward testing this paradigm, which is often invoked to explain the observed radius distribution of mature exoplanets (Fulton et al. 2017; Van Eylen et al. 2018).

The K2 and TESS missions have now enabled the detection of about ten close-in planets younger than 100 million years, all smaller than Jupiter (Mann et al. 2016; David et al. 2016, 2019; Newton et al. 2019; Bouma et al. 2020; Plavchan et al. 2020; Rizzuto et al. 2020; Martioli et al. 2021). The Kepler mission however has not yielded any planets with precise ages below one gigayear (Meibom et al. 2013). The reason is that during the **(Replaced: main replaced with: prime)** Kepler mission (2009–2013), only four open clusters were known in the Kepler field, with ages spanning 0.7 Gyr to 9 Gyr (Meibom et al. 2011). Though isochronal, gyrochronal, and lithium-based analyses suggest that younger Kepler planets do exist (Walkowicz & Basri 2013; Berger et al. 2018; David et al. 2021), accurate and precise age measurements typically require an ensemble of stars. Fortunately, recent analyses of the Gaia data have greatly expanded our knowledge of cluster memberships (e.g., Cantat-Gaudin et al. 2018; Zari et al. 2018; Kounkel & Covey 2019; Meingast et al. 2021; Kerr et al. 2021). As part of our Cluster Difference Imaging Photometric Survey (CDIPS, Bouma et al. 2019), we concatenated the available analyses from the literature, which yielded a list of candidate young and age-dated stars (see Appendix A).

Matching our young star list against stars observed by Kepler revealed that Kepler observed a portion of the  $\delta$  Lyr cluster (Stephenson-1; Theia 73). More specifically, a clustering analysis

of the Gaia data by Kounkel & Covey (2019) reported that Kepler 1627 (KIC 6184894; KOI 5245) is a  $\delta$  Lyr cluster member. Given the previous statistical validation of the close-in (**Replaced: mini-Neptune replaced with: Neptune-sized planet**) Kepler 1627b (Tenenbaum et al. 2012; Morton et al. 2016; Thompson et al. 2018), we begin by scrutinizing the properties of the cluster (Section 2). We find that the  $\delta$  Lyr cluster is  $38^{+6}_{-5}$  Myr old, and in Section 3 show that Kepler 1627 is both a binary and also a member of the cluster. Focusing on the planet (Section 4), we confirm that despite the existence of the previously unreported (**Replaced: M2.5 replaced with: M3**)V companion, hereafter Kepler 1627B, the planet orbits the G-dwarf primary, Kepler 1627A. We also analyze an asymmetry in the average transit profile, and a possible correlation between the individual transit times and the local light curve slope. We conclude by discussing broader implications for our ability to age-date a larger sample of planets (Section 5).

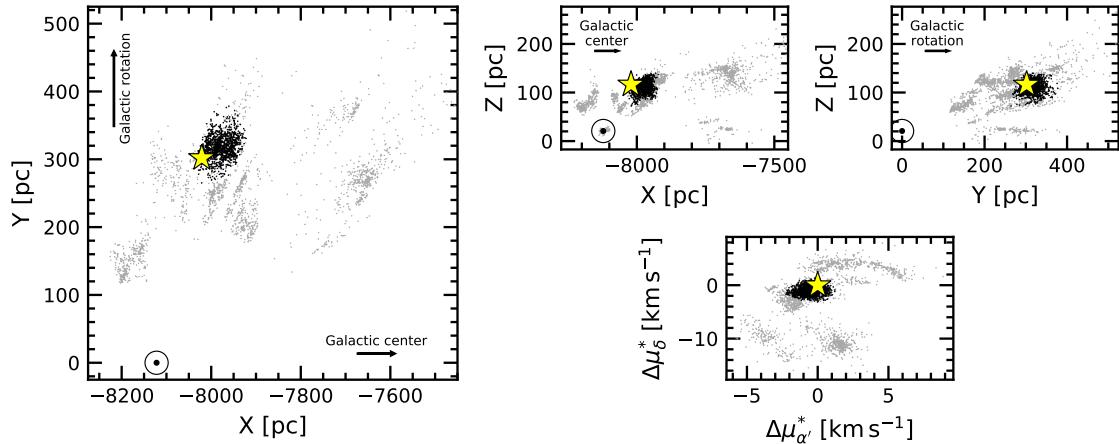
## 2. THE CLUSTER

To measure the age of the  $\delta$  Lyr cluster, we first selected a set of candidate cluster members (Section 2.1), and then analyzed these stars using a combination of the isochronal and gyrochronal techniques (Section 2.2).

### 2.1. Selecting Cluster Members

**Kounkel & Covey (2019)** applied an unsupervised clustering algorithm to Gaia DR2 on-sky positions, proper motions, and parallaxes for stars within the nearest kiloparsec. For the  $\delta$  Lyr cluster (Theia 73), they reported 3,071 candidate members. We matched these stars against the latest Gaia EDR3 observations using the dr2\_neighbourhood table from the ESA archive, taking the stars closest in proper motion and epoch-corrected angular distance as the presumed match (Gaia Collaboration et al. 2021a).  
**(Replaced: In Figure 1, (Added: we )have shown galactic positions only for the stars with parallax signal-to-noise exceeding 20.** replaced with: For plotting purposes, we focused only on the stars with parallax signal-to-noise exceeding 20.)  
**(Added: We calculated the tangential velocities for each of these stars relative to Kepler 1627 ( $\Delta\mu^*$ ) by subtracting the observed proper motion from what the proper motion at each star’s position would be if it were co-moving with Kepler 1627. )**

**(Replaced: The replaced with: Figure 1 shows that the)** reported cluster members (gray and black points) extend over a much larger volume (**Added: in both physical and kinematic space**) than the cluster previously identified by Stephenson (1959)



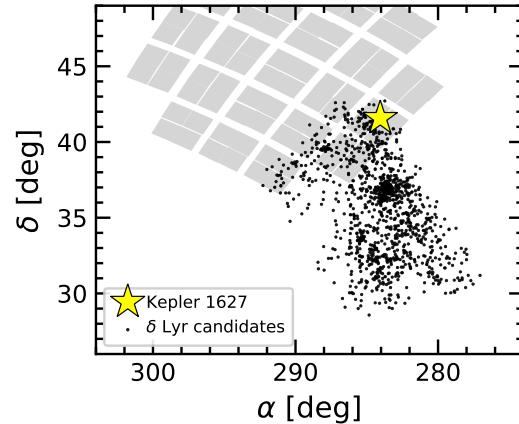
**Figure 1. Galactic positions and tangential velocities of stars in the  $\delta$  Lyr cluster.** Points are reported cluster members from Kounkel & Covey (2019). The tangential velocities relative to Kepler 1627 (bottom right) are computed assuming that every star has the same three-dimensional spatial velocity as Kepler 1627. Our analysis considers stars (black points) in the spatial and kinematic vicinity of Kepler 1627 (yellow star). The question of whether the other candidate cluster members (gray points) are part of the cluster is outside our scope. The location of the Sun is ( $\odot$ ) is shown (Added: to clarify the direction along which parallax uncertainties are expected to produce erroneous clusters).

and later corroborated by Eggen (1968). While the non-uniform “clumps” of stars might comprise a *bona fide* cluster of identically-aged stars, they could also be heavily contaminated by field stars. (Added: One reason to suspect this is that the spread in tangential velocities exceeds typical limits for kinematic coherence (Meingast et al. 2021)). We therefore considered stars only in the immediate kinematic and spatial vicinity of Kepler 1627 as candidate cluster members. We performed (Replaced: the replaced with: this) selection cut(Deleted: s ) manually, by drawing lassos with the interactive glue visualization tool (Beaumont et al. 2014) in the four projections shown in Figure 1. The overlap between the Kepler field and the resulting candidate cluster members is shown in Figure 2. While this method will include some field interlopers in the “cluster star” sample, and vice-versa, it should suffice for our aim of verifying the existence of the cluster in the vicinity of Kepler 1627.

## 2.2. The Cluster’s Age

### 2.2.1. Color-Absolute Magnitude Diagram

We measured the isochrone age using an empirical approach. The (Added: upper )left panel of Figure 3 shows the color-absolute magnitude diagram (CAMD) of candidate  $\delta$  Lyr cluster members, IC 2602(Added: ( $\approx$  38 Myr)), the Pleiades(Added: ( $\approx$  115 Myr)), and the field. The stars from the Pleiades and IC 2602 were adopted from Cantat-Gaudin et al. (2018), and the field stars are from the Gaia EDR3 Catalog of Nearby Stars (Gaia Collaboration et al. 2021b).(Added: We also compared against  $\mu$ -



**Figure 2. Kepler’s view of the  $\delta$  Lyr cluster, shown in equatorial coordinates.** Each black circle is a candidate cluster member selected based on its position and kinematics (Figure 1). Of the 1,201 candidate cluster members, 58 have at least one quarter of Kepler data. TESS has also observed most of the cluster, for one to two lunar months to date.

Tau ( $62 \pm 7$  Myr; Gagné et al. 2020) and the Upper-Centaurus-Lupus (UCL) component of the Sco OB2 association ( $\approx$ 16 Myr; Pecaut & Mamajek 2016). We adopted the UCL members from Damiani et al. (2019). For visual clarity, the latter two clusters are not shown in Figure 3.) We cleaned (Replaced: these replaced with: the membership lists) following the data filtering criteria from Gaia Collaboration et al. (2018a, Appendix B), except that we weakened the parallax precision requirement to  $\varpi/\sigma_\varpi > 5$ . (Added: This also involved cuts on the photometric signal to noise ratio, the number of vis-

ability periods used, the astrometric  $\chi^2$  of the single-source solution, and the  $G_{\text{BP}} - G_{\text{RP}}$  color excess factor.) These filters were designed to include genuine binaries while omitting instrumental artifacts.

(Replaced: We then corrected for extinction by querying replaced with: To correct for extinction, we queried) the 3-dimensional maps of Capitanio et al. (2017) and Lallement et al. (2018)<sup>1</sup>, and applied the extinction coefficients  $k_X \equiv A_X/A_0$  computed by Gaia Collaboration et al. (2018a) assuming that  $A_0 = 3.1E(B-V)$ . For (Added: UCL, )IC 2602, the Pleiades, and the  $\delta$  Lyr cluster, this procedure yielded a respective mean and standard deviation for the reddening of (Replaced:  $E(B-V) = \{0.020 \pm 0.003, 0.045 \pm 0.008, 0.032 \pm 0.006\}$  replaced with:  $E(B-V) = \{0.084 \pm 0.041, 0.020 \pm 0.003, 0.045 \pm 0.008, 0.032 \pm 0.006\}$ ). These values (Replaced: agree reasonably well with replaced with: are within a factor of two of) previously reported values (Replaced: from replaced with: in) the literature (Pecaut & Mamajek 2016; Gaia Collaboration et al. 2018a; Kounkel & Covey 2019; Bossini et al. 2019)(Added: , and are all small enough that the choice of whether to use them vs. other extinction estimates does not affect our primary conclusions).

Figure 3 shows that the  $\delta$  Lyr cluster and IC 2602 overlap, and therefore are approximately the same age. (Replaced: In our exploration, we also compared against  $\mu$ -Tau ( $62 \pm 7$  Myr; Gagné et al. 2020) and the Upper-Centaurus-Lupus (UCL) component of the Sco OB2 association ( $\approx 16$  Myr; Pecaut & Mamajek 2016). The pre-main-sequence M dwarfs of the  $\delta$  Lyr cluster were intermediate between the latter two clusters. replaced with: The pre-main-sequence M dwarfs of the  $\delta$  Lyr cluster were also seen to be “below” UCL and “above” the Pleiades.) To turn this heuristic interpolation into (Replaced: an age replaced with: a quantitative age) measurement, we used the empirical method developed by Gagné et al. (2020). In brief, we fitted the pre-main-sequence loci of a set of reference clusters, and the locus of the target  $\delta$  Lyr cluster was then modeled as a piecewise linear combination of these reference clusters. For our reference clusters, (Replaced: we adopted members of UCL, IC 2602, and the Pleiades from Damiani et al. (2019) and Cantat-Gaudin et al. (2018) respectively replaced with: we used UCL, IC 2602, and the Pleiades). We removed binaries by requiring

$\text{RUWE} < 1.3$ ,  $\text{radial\_velocity\_error}$  below the 80<sup>th</sup> percentile of each cluster’s distribution, and excluded stars that were obvious photometric binaries in the CAMD.<sup>2</sup> We then passed a moving box average and standard deviation across the CAMD in 0.10 mag bins, fitted a univariate spline to the binned values, and assembled a piecewise grid of (Deleted: hybrid) isochrones spanning the ages between UCL (Replaced: to replaced with: and) the Pleiades using Equation(Replaced: s 6 and 7 replaced with: 6) from Gagné et al. (2020). (Added: To derive a probability distribution function for the age of  $\delta$  Lyr cluster, we then assumed a Gaussian likelihood that treated the interpolated isochrones as the “model” and the  $\delta$  Lyr cluster’s isochrone as the “data” (Equation 7 from Gagné et al. 2020). The cluster’s age and its statistical uncertainty are then quoted as the the mean and standard deviation of this age posterior.)

The ages returned by this procedure depend on the ages assumed for each reference cluster. We adopted a 115 Myr age for the Pleiades (Dahm 2015), and a 16 Myr age for UCL (Pecaut & Mamajek 2016). The age of IC 2602 however is the most important ingredient, since it receives the most weight in the interpolation. Plausible ages for IC 2602 span 30 Myr to 46 Myr, with older ages being preferred by the lithium-depletion-boundary (LDB) measurements (Dobbie et al. 2010; Randich et al. 2018) and younger ages by the main-sequence turn-off (Stauffer et al. 1997; David & Hillenbrand 2015; Bossini et al. 2019). If we were to adopt the 30 Myr age for IC 2602, then the  $\delta$  Lyr cluster would be  $31^{+5}_{-4}$  Myr old. For the converse extreme of 46 Myr, the  $\delta$  Lyr cluster would be  $44^{+8}_{-7}$  Myr old. We adopt an intermediate 38 Myr age for IC 2602, which yields an age for the  $\delta$  Lyr cluster of  $38^{+6}_{-5}$  Myr.<sup>3</sup> Follow-up studies of the LDB or main-sequence turn-off in the  $\delta$  Lyr cluster could help determine a more precise and accurate age for the cluster, and are left for future work.

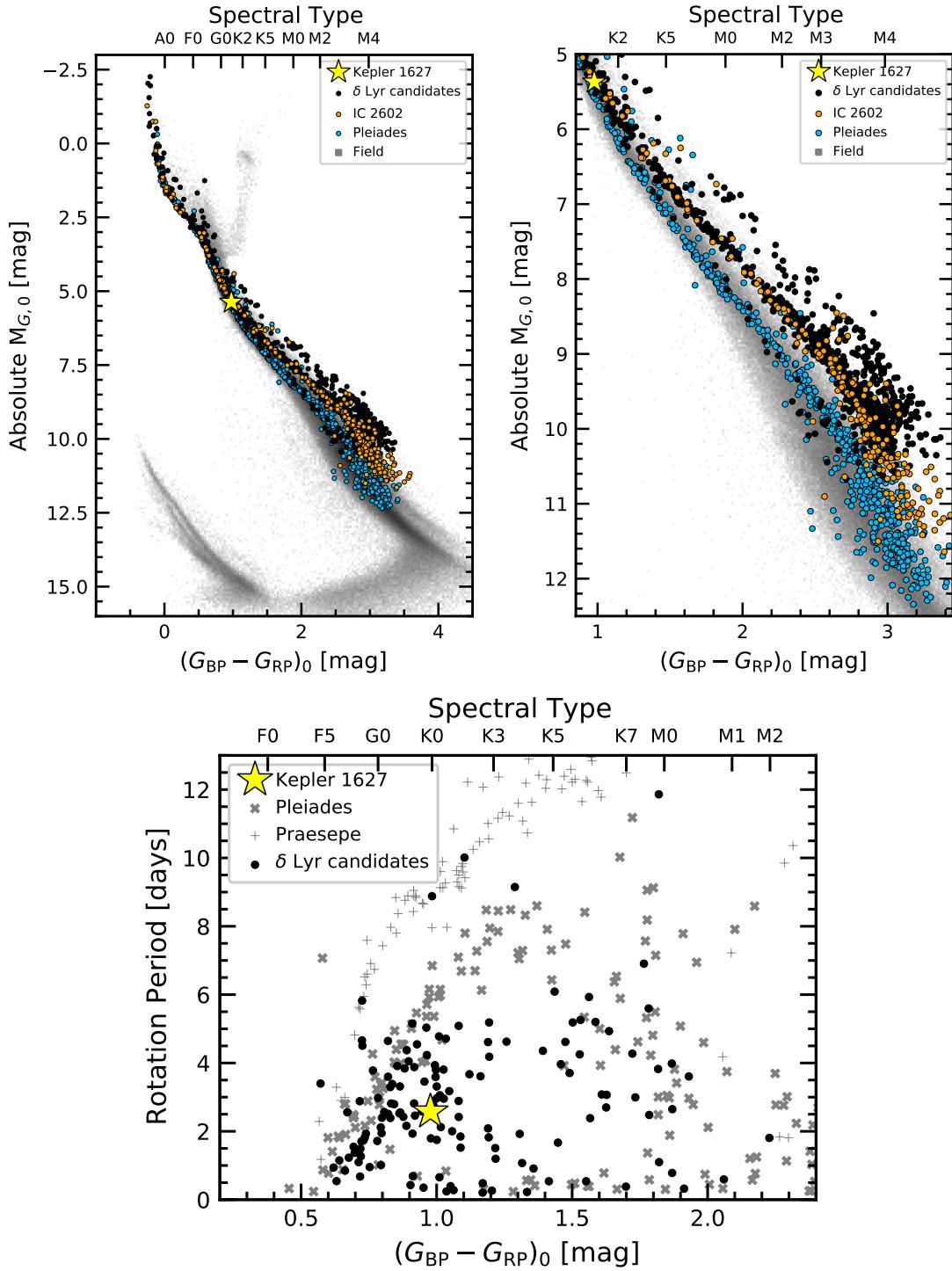
### 2.2.2. Stellar Rotation Periods

Of the 3,071 candidate  $\delta$  Lyr cluster members reported by Kounkel & Covey (2019), 924 stars were amenable to rotation period measurements ( $G < 17$  and  $(G_{\text{BP}} - G_{\text{RP}})_0 > 0.5$ ) using the TESS full frame

<sup>2</sup> For a description of the renormalized unit weight error (RUWE), see the GAIA DPAC technical note [http://www.rssd.esa.int/doc\\_fetch.php?id=3757412](http://www.rssd.esa.int/doc_fetch.php?id=3757412).

<sup>3</sup> Our exploration of the PARSEC and MIST isochrone models over a grid of ages, metallicities, and reddenings, yielded the best agreement for this  $\approx 38$  Myr age as well, given  $[\text{Fe}/\text{H}] = +0.1$  and  $A_V = 0.2$  (Bressan et al. 2012; Choi et al. 2016); this preferred CAMD reddening is higher than the Lallement et al. (2018) value by a factor of two.

<sup>1</sup> <https://stilism.obspm.fr/>, 2021/09/25



**Figure 3.** The  $\delta$  Lyr cluster is  $38^{+6}_{-5}$  Myr old. *Top:* Color-absolute magnitude diagram of candidate  $\delta$  Lyr cluster members, in addition to stars in IC 2602 ( $\approx 38$  Myr), the Pleiades ( $\approx 115$  Myr), and the Gaia EDR3 Catalog of Nearby Stars (gray background). The zoomed right panel highlights the pre-main-sequence. The  $\delta$  Lyr cluster and IC 2602 are approximately the same isochronal age. *Bottom:* TESS and Kepler stellar rotation period versus dereddened Gaia color, with the Pleiades and Praesepe (650 Myr) shown for reference (Rebull et al. 2016; Douglas et al. 2017). Most candidate  $\delta$  Lyr cluster members are gyrochronologically younger than the Pleiades; outliers are probably field interlopers.

image data.(Added: As a matter of scope, we restricted our attention to the 391 stars discussed in Section 2.1 in the spatial and kinematic proximity of Kepler 1627.) We extracted light curves from the TESS images using the nearest pixel to each star, and regressed them against systematics with the causal pixel model implemented in the unpopular package (Hattori et al. 2021). We then measured candidate rotation periods using a Lomb-Scargle periodogram (Lomb 1976; Scargle 1982; Astropy Collaboration et al. 2018). To enable cuts on crowding, we queried the Gaia source catalog for stars within a  $21.^{\circ}0$  radius of the target star (a radius of 1 TESS pixel). Within this radius, we recorded the number of stars with greater brightness than the target star, and with brightness within 1.25 TESS magnitudes of the target star.

We then cleaned the candidate TESS rotation period measurements through a combination of automated and manual steps.(Deleted: —As—a matter of scope, we restricted our attention to the 391 stars discussed in Section 2.1 in the spatial and kinematic proximity of Kepler 1627.)(Deleted: Kepler rotation periods were derived by MeQuillan et al. (2014) for 28 of our 391 stars; for these cases, we simply adopted the Kepler rotation period.)(Added: First, to validate the TESS rotation periods, we compared against 28 stars from McQuillan et al. (2014) that were also observed by Kepler. Of the 23 stars with Kepler periods below 10 days, 21 of the TESS periods agreed with the Kepler rotation periods; the other 2 were measured at the double-period harmonic. Of the remaining 5 stars with Kepler rotation periods above 10 days, none were correctly recovered by TESS, and 3 were near the half-period harmonic. We therefore adopted a general stance of skepticism for TESS rotation periods exceeding 10 days, and used the Kepler rotation periods whenever possible.) For the remaining stars with only TESS data, we focused only on the stars for which no companions were known with a brightness exceeding one-tenth of the target star in a  $21.^{\circ}0$  radius. There were 192 stars that met these crowding requirements, and that had TESS data available. For plotting purposes we then imposed a selection based on the strength of the signal itself: we required the Lomb Scargle power to exceed 0.2, and the period to be below 15 days.

The lower panel of Figure 3 shows the resulting 145 stars. The majority of these stars fall below the “slow sequence” of the Pleiades, consistent with a gyrochronal age for the  $\delta$  Lyr cluster below 100 Myr. In fact, the rotation-color distributions of other 30 Myr to 50 Myr clusters (e.g., IC 2602 and IC 2391) are indistinguishable (Douglas et al.

2021). Approximately 10 of the  $\delta$  Lyr cluster stars appear as outliers above the “slow sequence”. Assuming that they are all false positives (*i.e.*, field interlopers), our rotation period detection fraction would be  $135/192 \approx 70\%$ . (Replaced: The other stars are likely to be field contaminants. replaced with: Although some of these outlier stars might be unresolved F+K binaries that are in the cluster (Stauffer et al. 2016), assuming that they are field contaminants provides a more secure lower bound of the rotation period detection fraction.) A final possible confounding factor – binarity – is known to affect the “fast sequence” of stars beneath the slow sequence (Meibom et al. 2007; Gillen et al. 2020; Bouma et al. 2021). We do not expect it to change the central conclusion regarding the cluster’s age.

### 3. THE STARS

#### 3.1. Kepler 1627A

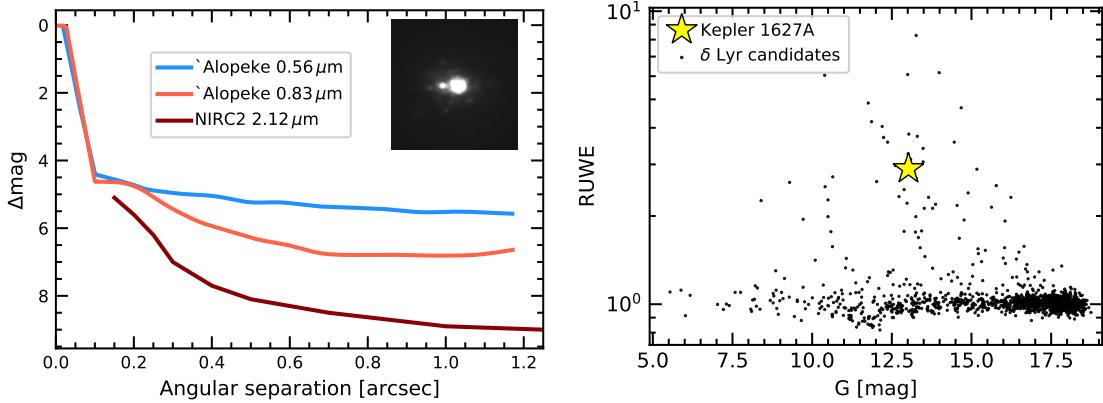
##### 3.1.1. Age

Based on the spatial and kinematic association of Kepler 1627 with the  $\delta$  Lyr cluster, and the assumption that the planet formed shortly after the star, it seems likely that Kepler 1627 is the same age as the cluster. There are two consistency checks on whether this is true: rotation and lithium. Based on the Kepler light curve, the rotation period is  $2.642 \pm 0.042$  days, where the quoted uncertainty is based on the scatter in rotation periods measured from each individual Kepler quarter. This is consistent with comparable cluster members (Figure 3).

To infer the amount of Li I from the  $6708 \text{ \AA}$  doublet (e.g., Soderblom et al. 2014), we acquired an iodine-free spectrum from Keck/HIRES on the night of 2021 March 26 using the standard setup and reduction techniques of the California Planet Survey (Howard et al. 2010). Following the equivalent width measurement procedure described by Bouma et al. (2021), we find  $\text{EW}_{\text{Li}} = 233^{+5}_{-7} \text{ m\AA}$ . This value does not correct for the Fe I blend at  $6707.44 \text{ \AA}$ . Nonetheless, given the stellar effective temperature (Table 1), this measurement is in agreement with expectations for a  $\approx 40$  Myr star (e.g., as measured in IC 2602 by Randich et al. 2018). It is also larger than any lithium equivalent widths measured by Berger et al. (2018) in their analysis of 1,301 Kepler-star spectra.

##### 3.1.2. Stellar Properties

The adopted stellar parameters are listed in Table 1. The stellar mass, radius, and effective temperature are found by interpolating against a 38 Myr MIST isochrone (Choi et al. 2016). The statistical uncertainties are propagated from the absolute magnitude (mostly originating from the



**Figure 4. Kepler 1627 is a binary.** *Left:* High-resolution imaging from Gemini-North/‘Alopeke and Keck/NIRC2 shows an  $\approx$ (Replaced: M2.5 replaced with: M3)V companion at  $\rho \approx 0.^{\prime\prime}16$ , which corresponds to a projected separation of  $53 \pm 4$  AU. The inset shows a cutout of the stacked NIRC2 image (North is up, East is left, scale is set by the separation of the binary). The lines show  $5-\sigma$  contrast limits for the ‘Alopeke filters, and  $6-\sigma$  contrast limits for NIRC2 outside of  $0.^{\prime\prime}15$ . *Right:* Gaia EDR3 renormalized unit weight error (RUWE) point estimates for candidate  $\delta$  Lyr cluster members. Since other members of the cluster with similar brightnesses have comparable degrees of photometric variability, the high RUWE independently suggests that Kepler 1627 is a binary.

417 parallax uncertainty) and the color; the systematic  
418 uncertainties are taken to be the difference  
419 between the PARSEC (Bressan et al. 2012) and  
420 MIST isochrones. Reported uncertainties are a  
421 quadrature sum of the statistical and systematic  
422 components. As a consistency check, we ana-  
423 lyzed the aforementioned Keck/HIRES spectrum  
424 from the night of 2021 March 26 using a com-  
425 bination of SpecMatch-Emp for stellar prop-  
426 erties, and SpecMatch-Synth for  $v\sin i$  (Yee  
427 et al. 2017). This procedure yielded  $T_{\text{eff}} = 5498 \pm$   
428 100 K,  $\log g = 4.6 \pm 0.1$ ,  $[\text{Fe}/\text{H}] = 0.15 \pm 0.10$  from  
429 SpecMatch-Emp, and  $v\sin i = 18.9 \pm 1.0$  from  
430 SpecMatch-Synth. These values are within the  
431  $1-\sigma$  uncertainties of our adopted values from the  
432 isochrone interpolation.

### 3.2. Kepler 1627B

434 We first noted the presence of a close neighbor in  
435 the Kepler 1627 system on 2015 July 22 when we  
436 acquired adaptive optics imaging using the NIRC2  
437 imager on Keck-II. We used the narrow camera  
438 (FOV =  $10.2''$ ) to obtain 8 images in the  $K'$  fil-  
439 ter ( $\lambda = 2.12 \mu\text{m}$ ) with a total exposure time of  
440 160 s. We analyzed these data following Kraus  
441 et al. (2016), which entailed using PSF-fitting to  
442 measure the separation, position angle, and con-  
443 trast of the candidate companion. The best-fitting  
444 empirical PSF template was identified from among  
445 the near-contemporaneous observations of single  
446 stars in the same filter. The mean values inferred  
447 from the 8 images are reported in Table 1. To es-  
448 timate the detection limits, we analyzed the resi-  
449 duals after subtracting the empirical PSF template.  
450 Within each residual image, the flux was measured

451 through 40 mas apertures centered on every pixel,  
452 and then the noise as a function of radius was esti-  
453 mated from the RMS within concentric rings. Fi-  
454 nally, the detection limits were estimated from the  
455 strehl-weighted sum of the detection significances  
456 in the image stack, and we adopted the  $6-\sigma$  thresh-  
457 old as the detection limit for ruling out additional  
458 companions.

459 We also observed Kepler 1627 on Gemini-North  
460 using the ‘Alopeke speckle imager on 2021 June  
461 24. ‘Alopeke is a dual-channel speckle interfer-  
462 ometer that uses narrow-band filters centered at  
463  $0.83 \mu\text{m}$  and  $0.56 \mu\text{m}$ . We acquired three sets  
464 of  $1000 \times 60$  msec exposures during good seeing  
465 ( $0.45''$ ), and used the autocorrelation function of  
466 these images to reconstruct a single image and  $5-$   
467  $\sigma$  detection limits (see Howell et al. 2011). This  
468 procedure yielded a detection of the companion in  
469 the  $0.83 \mu\text{m}$  notch filter, but not the  $0.56 \mu\text{m}$  filter.  
470 The measured projected separation and magnitude  
471 difference are given in Table 1.

472 Figure 4 summarizes the results of the high-  
473 resolution imaging. The Gaia EDR3 parallax for  
474 the primary implies a projected separation of  $53 \pm$   
475 4 AU, assuming the companion is bound. Al-  
476 though the companion is unresolved in the Gaia  
477 source catalog (there are no comoving, codistant  
478 candidate companions brighter than  $G < 20.5$  mag  
479 within  $\rho < 120''$ ), its existence was also suggested  
480 by the primary star’s large (Replaced: renorm-  
481 alized unit weight error (RUWE), replaced with:  
482 RUWE) relative to other members of the  $\delta$  Lyr  
483 cluster (Added: (RUWE $\approx 2.9$ ; roughly the 98<sup>th</sup>  
484 percentile of the cluster’s distribution)). Based  
485 on the apparent separation, the binary orbital pe-

486 riod is of order hundreds of years. The large  
 487 RUWE is therefore more likely to be caused by a  
 488 PSF-mismatch skewing the Gaia centroiding dur-  
 489 ing successive scans, rather than true astrometric  
 490 motion. Regardless, given the low geometric prob-  
 491 ability that a companion imaged at  $\rho \approx 0.^{\circ}16$  is a  
 492 chance line-of-sight companion, we proceed under  
 493 the assumption that the companion is bound,  
 494 and that Kepler 1627 is a binary. Given the dis-  
 495 tance and age, the models of Baraffe et al. (2015)  
 496 imply a companion mass of (**Replaced:**  $M_B \approx$   
 497  $0.33M_{\odot}$  replaced with:  $M_B \approx 0.30M_{\odot}$ ) and com-  
 498 panion temperature of (**Replaced:**  $T_{\text{eff},B} \approx 3450$   
 499 replaced with:  $T_{\text{eff},B} \approx 3408$ ) K. The correspond-  
 500 ing spectral type is roughly (**Replaced:** M2.5 re-  
 501 placed with: M3)V (Pecaut & Mamajek 2013).  
 502 These models combined with the NIRC2 contrast  
 503 limits imply physical limits on tertiary companions  
 504 of  $M_{\text{ter}} < 50M_{\text{Jup}}$  at  $\rho = 50$  AU,  $M_{\text{ter}} < 20M_{\text{Jup}}$  at  
 505  $\rho = 100$  AU, and  $M_{\text{ter}} < 10M_{\text{Jup}}$  at  $\rho = 330$  AU.

## 506 4. THE PLANET

### 507 4.1. Kepler Light Curve

508 The Kepler space telescope observed Kepler 1627  
 509 at a 30-minute cadence from 2009 May 2 until  
 510 2013 April 8. Data gaps during quarters 4, 9, and  
 511 13 led to an average duty cycle over the 3.9 year  
 512 interval of 67%. Kepler 1627 was also observed at  
 513 1-minute cadence from 2012 Oct 5 until 2013 Jan  
 514 11. The top panel of Figure 5 shows a portion of the  
 515 30-minute cadence PDCSAP light curve. Nonas-  
 516 trophysical variability has been removed using the  
 517 methods discussed by Smith et al. (2017); the de-  
 518 fault optimal aperture was assumed (Smith et al.  
 519 2016). Cadences with non-zero quality flags (9%  
 520 of the data) have been omitted. The resulting pho-  
 521 tometry is dominated by a quasi-periodic starspot  
 522 signal with a peak-to-peak amplitude that varies  
 523 between 2% and 8%. (**Added:** Given that the  
 524 secondary companion's brightness in the Kepler  
 525 band is  $\approx 1.5\%$  that of the primary, source con-  
 526 fusion for the rotation signal is not expected to  
 527 be an issue.) Previous analyses have identified  
 528 and characterized the smaller transit signal (Tenen-  
 529 baum et al. 2012; Thompson et al. 2018), validated  
 530 its planetary nature (Morton et al. 2016), and even  
 531 searched the system for transit timing variations  
 532 (Holczer et al. 2016). Nonetheless, since the clus-  
 533 ter membership provides us with more precise stel-  
 534 lar parameters than those previously available, we  
 535 opted to reanalyze the light curve.

#### 536 4.1.1. Transit and Stellar Variability Model

537 We fitted the Kepler long cadence time se-  
 538 ries with a model that simultaneously included  
 539 the planetary transit and the stellar variabil-  
 540 ity. The stellar variability was modeled with

541 the RotationTerm Gaussian Process kernel in  
 542 exoplanet (Foreman-Mackey et al. 2020). This  
 543 kernel assumes that the variability is generated by  
 544 a mixture of two damped simple harmonic oscil-  
 545 lators with characteristic frequencies set by  $1/P_{\text{rot}}$   
 546 and its first harmonic. We additionally included a  
 547 jitter term to inflate the flux uncertainties in a man-  
 548 ner that accounted for otherwise unmodeled excess  
 549 white noise, and let the eccentricity float. For the  
 550 limb-darkening, we assumed a quadratic law, and  
 551 sampled using the uninformative prior suggested  
 552 by Kipping (2013).

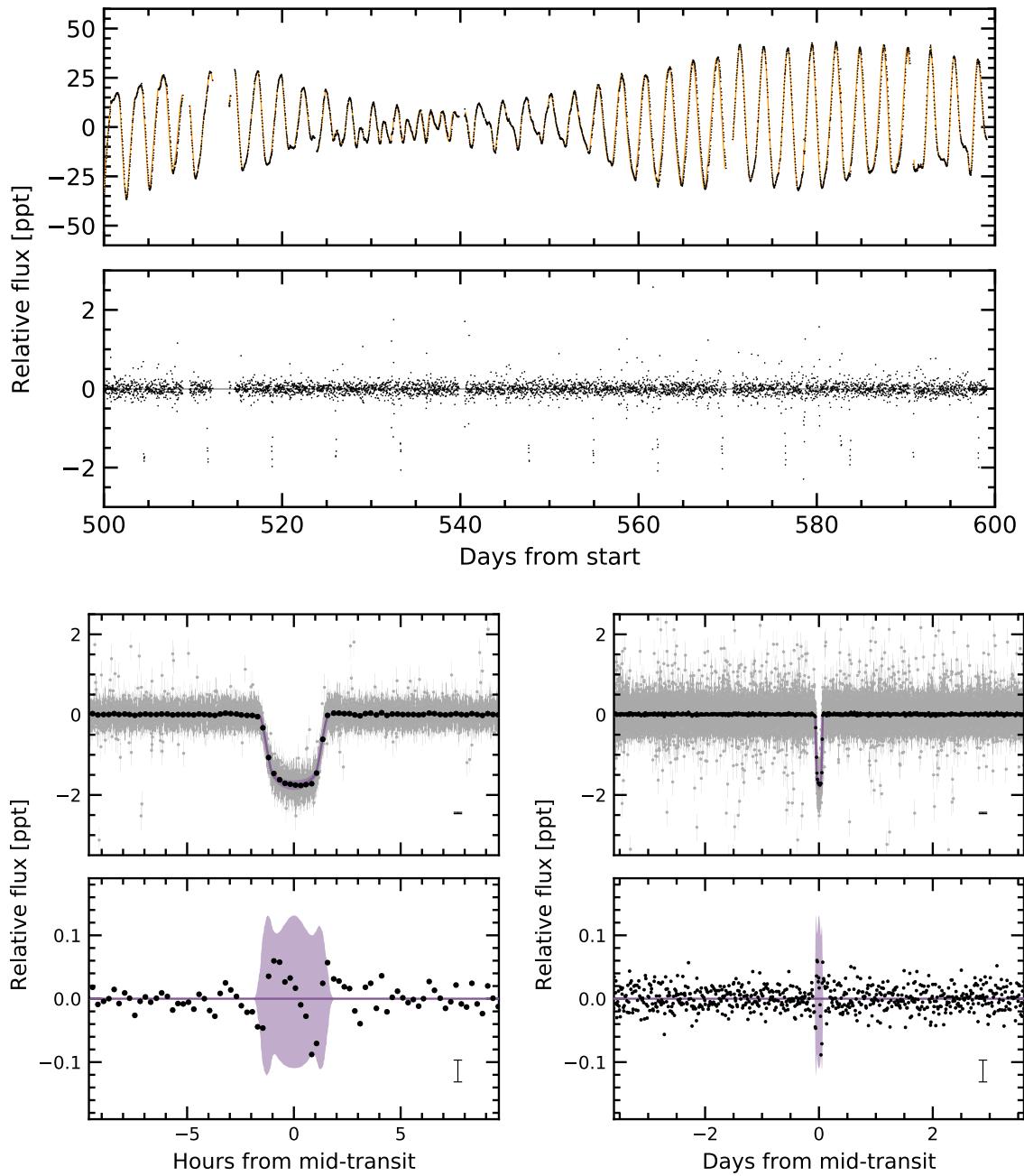
553 Our model therefore included 10 free parameters  
 554 for the transit ( $\{P, t_0, \delta, b, u_1, u_2, R_{\star}, \log g, e, \omega\}$ ), 2  
 555 free parameters for the light curve normalization  
 556 and a white noise jitter ( $\{\langle f \rangle, \sigma_f\}$ ), and 5 hyper-  
 557 parameters for the GP ( $\{\sigma_{\text{rot}}, P_{\text{rot}}, Q_0, dQ, f\}$ ). We  
 558 also considered including an additive SHOTerm  
 559 kernel to account for stochastic noise, but found  
 560 that this did not affect the results, and so opted for  
 561 the simpler GP kernel. We fitted the models us-  
 562 ing PyMC3 (Salvatier et al. 2016; Theano Devel-  
 563 opment Team 2016), and accounted for the finite  
 564 integration time of each exposure in the numerical  
 565 integration when evaluating the model light curve  
 566 (see Kipping 2010). We assumed a Gaussian like-  
 567 lihood, and after initializing each model with the  
 568 parameters of the maximum *a posteriori* model,  
 569 we sampled using PyMC3’s gradient-based No-U-  
 570 Turn Sampler (Hoffman & Gelman 2014) in the  
 571 bases indicated in Table 2. We used  $\hat{R}$  as our con-  
 572 vergence diagnostic (Gelman & Rubin 1992).

573 Figure 5 shows the resulting best-fit model in  
 574 orange (top) and purple (bottom). The model  
 575 parameters and their uncertainties, given in Ta-  
 576 ble 2, are broadly consistent with a (**Replaced:**  
 577 **mini-Neptune** replaced with: Neptune-)sized  
 578 planet ( $3.82 \pm 0.16R_{\oplus}$ ) on a close-in circular<sup>4</sup>  
 579 orbit around a G8V host star ( $0.88 \pm 0.02R_{\odot}$ ). This  
 580 best-fit planet size is consistent with those previ-  
 581 ously reported by Morton et al. (2016) and Berger  
 582 et al. (2018), and corrects for the small amount of  
 583 flux dilution from Kepler 1627B.

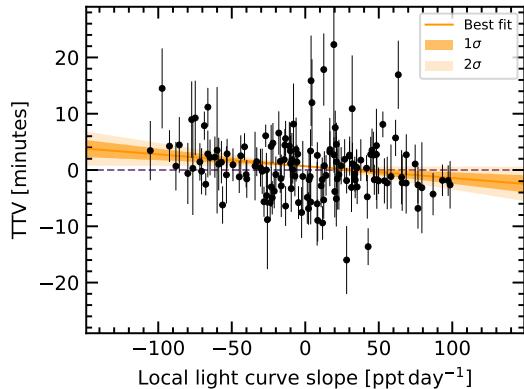
#### 584 4.1.2. Transit Asymmetry

585 The transit fit however is not perfect: the lower  
 586 panels of Figure 5 show an asymmetric residual in  
 587 the data relative to the model: the measured flux  
 588 is high during the first half of transit, and low in  
 589 the second half. The semi-amplitude of this devi-  
 590 ation is  $\approx 50$  ppm, which represents a  $\approx 3\%$  dis-  
 591 tortion of the transit depth ( $\delta = 1759 \pm 62$  ppm).  
 592 Note that although this asymmetry is within the 2-  
 593  $\sigma$  model uncertainties, the model has a jitter term

<sup>4</sup> Our transit fitting yields  $e < 0.48$  at  $2\sigma$ ; the constraints on the eccentricity are not particularly strong.



**Figure 5. The light curve of Kepler 1627.** *Top:* The Kepler data span 1,437 days (3.9 years), sampled at 30 minute cadence; a 100 day segment is shown. The top panel shows the PDCSAP median-subtracted flux in units of parts-per-thousand ( $\times 10^{-3}$ ). The dominant signal is induced by starspots. The stellar variability model (orange line) is subtracted below, revealing the transits of Kepler 1627Ab. The online Figure Set spans the entire 3.9 years of observations. *Bottom:* Phase-folded transit of Kepler 1627Ab with stellar variability removed. Windows over 20 hours (*left*) and the entire orbit (*right*) are shown, and the residual after subtracting the transit is in the bottom-most row. The  $2\sigma$  model uncertainties and the best-fit model are the light purple band and the dark purple line. Gray points are individual flux measurements; black points bin these to 15 minute intervals, and have a representative  $1\sigma$  error bar in the lower right of each panel. The asymmetric residual during transit is larger than the out-of-transit scatter.



**Figure 6.** (Replaced: Possible replaced with: Weak) evidence for a prograde orbit of Kepler 1627 Ab. The time of each Kepler transit was measured, along with the local slope of the light curve. The two quantities (Replaced: are weakly replaced with: might be) anti-correlated ( $\approx 2\sigma$ ), which (Replaced: might replaced with: could) be caused by starspot crossings during the first (second) half of transit inducing a positive (negative) TTV, provided that the orbit is prograde (Mazeh et al. 2015). The units along the abscissa can be understood by considering that the stellar flux changes by  $\sim 60$  ppt per half rotation period ( $\sim 1.3$  days).

that grows to account for excess white noise in the flux. The significance of the asymmetry is therefore best assessed in comparison against the intrinsic out-of-transit scatter in the data ( $\approx 16$  ppm), not the model uncertainties. The lower right panel of Figure 5 demonstrates that the scatter during transit is higher than during all other phases of the planet’s orbit.

To determine whether the asymmetry could be a systematic caused by our stellar variability model, we explored an alternative approach in which we isolated each transit window, locally fitted out polynomial trends, and then binned all the observed transits; the asymmetry was still present at a comparable amplitude. Appendix B describes a more detailed analysis, which finds that the asymmetry also seems to be robust to different methods of data binning in time and by local light curve slope. Possible astrophysical explanations are discussed in Section 5.

#### 4.1.3. Transit Timing and the Local Slope

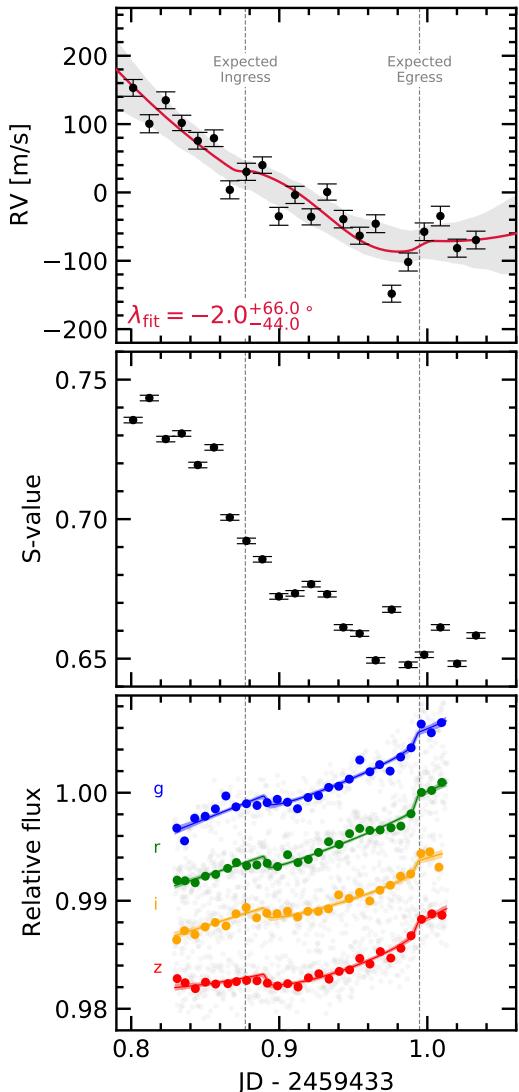
The previous analysis by Holczer et al. (2016) did not find any significant long-term transit timing or duration variations (TTVs or TDVs) for Kepler 1627. Quantitatively, the mean and standard deviation of the TTVs and TDVs they measured were  $-1.1 \pm 13.8$  min and  $-3.3 \pm 22.1$  min. In an earlier analysis however, Holczer et al. (2015) studied correlations between TTVs and local light

curve slopes, and for Kepler 1627 found a weak correlation of  $-29 \pm 13$  min day $^{-1}$  between the two quantities. Given the possible connection between such correlations and the unresolved starspot crossings that we expect to be present in the Kepler 1627 light curve (Mazeh et al. 2015), we opted to re-examine the individual transit times.

We therefore isolated each of the 144 observed transits to within  $\pm 4.5$  hr of each transit, and fitted each window with both *i*) a local (Added: second or fourth-order) polynomial baseline plus the transit, and *ii*) a local linear trend(Added: plus the transit). (Deleted: We considered the results both for a second and fourth-order time-dependence in the local baseline.) We let the mid-time of each transit float, and then calculated the residual between the measured mid-time and that of a periodic orbit. This residual, the transit timing variation, is plotted in Figure 6 against the local linear slope for the fourth-order polynomial baseline. The slope of  $-21 \pm 10$  min day $^{-1}$  is similar to that found by Holczer et al. (2015). (Added: The best-fit line yields  $\chi^2 = 306.1$ , with  $n = 140$  data points. An alternative model of just a flat line yields  $\chi^2 = 315.6$ . The difference in the Bayesian information criterion between the two models is  $BIC_{\text{flat}} - BIC_{\text{line}} = 4.5$ , which corresponds to a Bayes factor of  $\approx 9.4$ . According to the usual Kass & Raftery (1995) criteria, this is “positive” evidence for the model with a finite slope. We view it as suggestive at best, particularly given the poor reduced  $\chi^2$ .)

(Replaced: One replaced with: A separate) concern we had in this analysis was whether our transit fitting procedure might induce spurious correlations between the slope and transit time. In particular, using the second-order polynomial baseline yielded a larger anti-correlation between the TTVs and local slopes, of  $-79 \pm 14$  min day $^{-1}$ . We therefore performed an injection-recovery procedure in which we injected transits at different phases in the Kepler 1627 light curve and repeated the TTV analysis. This was done at  $\approx 50$  phases, each separated by  $0.02 P_{\text{orb}}$  while omitting the phases in transit. For the second-order polynomial baseline, this procedure yielded a similar anti-correlation in the injected transits as that present in the real transit; using this baseline would therefore bias the result. However for the fourth-order baseline, the correlation present in the data was stronger than in all but one of the injected transits. Possible interpretations are discussed below. Given the (Added: lack of) statistical significance, this analysis should be interpreted as suggestive at best.

#### 4.2. Planet Confirmation



**Figure 7. Keck/HIRES radial velocities and MuSCAT3 photometry from the transit of 2021 Aug 7.** Top: The radial velocity jitter across the 15 minute exposures ( $\sigma_{\text{RV}} \approx 30 \text{ m s}^{-1}$ ) prevented us from detecting the RM effect; a model including the RM anomaly and a quadratic trend in time to fit the spot-induced  $\approx 250 \text{ m s}^{-1}$  trend is shown (see Appendix C). Shaded bands show  $2-\sigma$  model uncertainties. Middle: The RV variations are strongly correlated with varying emission in the Ca H and K lines. Bottom: The photometric transit depths are consistent across the *griz* bandpasses. The photometry is binned at 10 minute intervals.

If the Kepler 1627Ab transit signal is created by a genuine planet, then to our knowledge it would be the youngest planet yet found by the (Replaced:

main replaced with: prime) Kepler mission.<sup>5</sup> Could the transit be produced by anything other than a planet orbiting this near-solar analog? Morton et al. (2016) validated the planet based on the transit shape, arguing that the most probable false positive scenario was that of a background eclipsing binary, which had a model-dependent probability of  $\approx 10^{-5}$ . However, this calculation was performed without knowledge of the low-mass stellar companion ( $M_B \approx 0.33 M_\odot$ ). Validated planets have also previously been refuted (e.g., Shporer et al. 2017). We therefore reassessed false positive scenarios in some detail.

As an initial plausibility check, Kepler 1627B contributes 1% to 2% of the total flux observed in the Kepler aperture. For the sake of argument, assume the former value. The observed transit has a depth of  $\approx 0.18\%$ . A 18% deep eclipse of Kepler 1627B would therefore be needed to produce a signal with the appropriate depth. The shape of the transit signal however requires the impact parameter to be below 0.77 ( $2-\sigma$ ); the tertiary transiting the secondary would therefore need to be non-grazing with  $R_3/R_2 \approx 0.4$ . This yields a contradiction: this scenario requires an ingress and egress phase that each span  $\approx 40\%$  of the transit duration ( $\approx 68$  min). The actual measured ingress and egress duration is  $\approx 17$  min, a factor of four times too short. The combination of Kepler 1627B's brightness, the transit depth, and the ingress duration therefore disfavor the scenario that Kepler 1627B might host the transit signal.

Beyond this simple test, a line of evidence that effectively confirms the planetary interpretation is that the stellar density implied by the transit duration and orbital period is inconsistent with an eclipsing body around the M-dwarf companion. We find  $\rho_* = 2.00 \pm 0.24 \text{ g cm}^{-3}$ , while the theoretically expected density for Kepler 1627B (Added: given its nominal age of 38 Myr and mass of  $0.30 M_\odot$ ) is  $\approx 4.6 \text{ g cm}^{-3}$  (Baraffe et al. 2015). The transit duration is therefore too long to be explained by a star eclipsing the M dwarf secondary at  $10-\sigma$ . (Added: Adversarially assuming a younger age (32 Myr) and a larger companion mass ( $0.35 M_\odot$ ), the companion's density could be as low as  $\approx 3.5 \text{ g cm}^{-3}$ . This is still in tension with the data.) While the planet might hypothetically still orbit a hidden close and bright companion, this possibility is implausible given *i*) the lack of secondary lines in the HIRES spectra, *ii*) the lack of secondary rotation signals in the Kepler photom-

<sup>5</sup> The re-purposed K2 mission however has found two younger systems containing five planets: K2-33b ( $9 \pm 1$  Myr; Mann et al. 2016; David et al. 2016) and V1298 Tau ( $23 \pm 4$  Myr; David et al. 2019).

try, and *iii*) the proximity of Kepler 1627 to the  $\delta$  Lyr cluster locus on the Gaia CAMD (Figure 3).

The correlation noted in Section 4.1.3 between the TTVs and the local light curve slope might be an additional line of evidence in support of the planetary interpretation. Unless it is a statistical fluke (a  $\approx 5\%$  possibility), then the most likely cause of the correlation is unresolved starspot crossings (Mazeh et al. 2015). These would only be possible if the planet transits the primary star, which excludes a background eclipsing binary scenario. The correlation would also suggest that the planet’s orbit is prograde. The latter point assumes that the dominant photometric variability is induced by dark spots, and not bright faculae. Given the observed transition of Sun-like stellar variability from spot to faculae-dominated regimes between young and old ages, we expect this latter assumption to be reasonably secure (Shapiro et al. 2016; Montet et al. 2017; Reinhold & Hekker 2020).

A third supporting line of evidence for the planetary interpretation also exists. We observed a transit of Kepler 1627Ab on the night of 2021 Aug 7 (**Replaced: simultaneously with Keck/HIRES and MuSCAT3 replaced with: spectroscopically with HIRES at the Keck-I telescope and photometrically in *griz* bands with MuSCAT3 at Haleakalā Observatory**). (**Added: Details of the observation sequence are discussed in Appendix C; Figure 7 shows the results.**) (**Deleted: .—We scheduled the observations using the ephemeris of Holezer et al. (2016).**) Although we did not detect the Rossiter-McLaughlin (RM) anomaly, the multi-band MuSCAT3 light curves show that the transit is achromatic (**Deleted: (Figure 7)**). Quantitatively, when we fitted the MuSCAT3 photometry with a model that lets the transit depths vary across each bandpass, we found *griz* depths consistent with the Kepler depth at 0.6, 0.3, 0.3, and  $1.1\sigma$  respectively. (**Added: The achromatic transits strongly favor Kepler 1627A as the transit host, since Kepler 1627B is a much redder star.**) Conditioned on the ephemeris and transit depth from the Kepler data, the MuSCAT3 observations also suggested a transit duration  $17.3 \pm 4.3$  min shorter than the Kepler transits. However, given both the lack of TDVs in the Kepler data and the relatively low signal-to-noise of the MuSCAT3 transit, further photometric follow-up would be necessary to determine whether the transit duration is (**Replaced: indeed replaced with: actually**) changing.

For our RM analysis, the details are discussed in Appendix C. While the velocities are marginally more consistent with a prograde or polar orbit than a retrograde orbit, the spot-corrected exposure-to-

exposure scatter ( $\sigma_{RV} \approx 30 \text{ m s}^{-1}$ ) is comparable to the expected RM anomaly assuming an aligned orbit ( $\Delta v_{RM} \approx 20 \text{ m s}^{-1}$ ). We are therefore not in a position to claim a spectroscopic detection of the RM effect, nor to quantify the stellar obliquity.

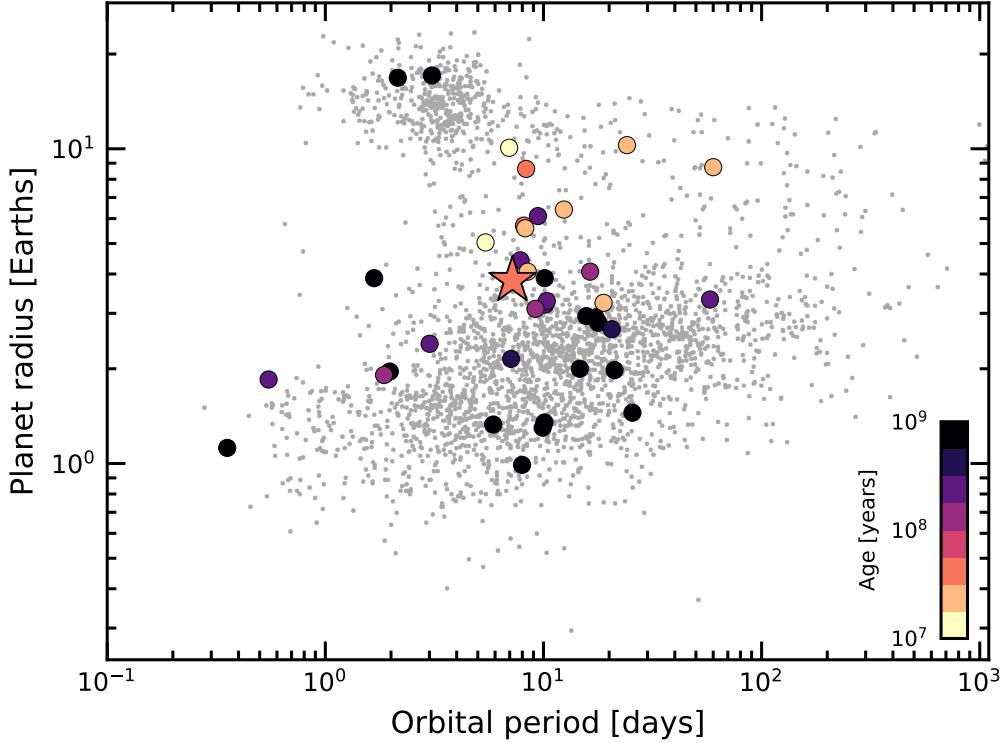
## 5. DISCUSSION & CONCLUSIONS

Kepler 1627Ab provides a new extremum in the ages of the Kepler planets, and opens multiple avenues for further study. Observations of spectroscopic transits at greater (**Replaced: precision replaced with: signal-to-noise**) should yield a measurement of the stellar obliquity, which would confirm or refute the prograde orbital geometry suggested by the TTV-local slope correlation. Separately, transit spectroscopy aimed at detecting atmospheric outflows could yield insight into the evolutionary state of the atmosphere (e.g., Ehrenreich et al. 2015; Spake et al. 2018; Vissapragada et al. 2020). Observations that quantify the amount of high-energy irradiation incident on the planet would complement these efforts, by helping to clarify the expected outflow rate (e.g., Poppenhaeger et al. 2021). Finally, a challenging but informative quantity to measure would be the planet’s mass. Measured at sufficient precision, for instance through a multi-wavelength radial velocity campaign, the combination of the size, mass, and age would yield constraints on both the planet’s composition and its initial entropy (Owen 2020).

More immediately, the Kepler data may yet contain additional information. For instance, one possible explanation for the transit asymmetry shown in Figure 5 is that of a dusty asymmetric outflow. Dusty outflows are theoretically expected for young mini-Neptunes, and the amplitude of the observed asymmetry is consistent with predictions (Wang & Dai 2019). A second possibility is that the planetary orbit is slightly misaligned from the stellar spin axis, and tends to transit starspot groups at favored stellar latitudes. This geometry would be necessary in order to explain how the starspot crossings could add up coherently (**Added: , given that the planetary orbital period (7.203 days) and the stellar rotation period (2.642 days) are not a rational combination**). Other possibilities including gravity darkening or TTVs causing the asymmetry are disfavored (see Appendix B).

Beyond the asymmetric transits, Appendix D highlights an additional abnormality in the short-cadence Kepler data, in the arrival time distribution of stellar flares. We encourage its exploration by investigators more versed in the topic than ourselves.

In the context of the transiting planet population, Kepler 1627Ab is among the youngest known (Figure 8). Comparable systems with precise ages include K2-33 (Mann et al. 2016; David et al. 2016),



**Figure 8. Radii, orbital periods, and ages of transiting exoplanets.** Planets younger than a gigayear with  $\tau/\sigma_\tau > 3$  are emphasized, where  $\tau$  is the age and  $\sigma_\tau$  is its uncertainty. Kepler 1627Ab is shown with a star. The large sizes of the youngest transiting planets could be explained by their primordial atmospheres not yet having evaporated; direct measurements of the atmospheric outflows or planetary masses would help to confirm this expectation. Selection effects may also be important. Parameters are from the NASA Exoplanet Archive (2021 Sept 15).

DS Tuc (Benatti et al. 2019; Newton et al. 2019), HIP 67522 (Rizzuto et al. 2020), TOI 837 (Bouma et al. 2020), the two-planet AU Mic (Plavchan et al. 2020; Martioli et al. 2021) and the four-planet V1298 Tau (David et al. 2019). Kepler 1627Ab is one of the smaller planets in this sample ( $3.82 \pm 0.16 R_\oplus$ ), which could be linked to the selection effects imposed by spot-induced photometric variability at very young ages (e.g., Zhou et al. 2021). (Replaced: However, replaced with: However, it seems that) smaller planets could have been detected: (Added: based on the per-target detection contours,) the Kepler pipeline's median completeness extended to  $1.6 R_\oplus$  at 10 day orbital periods, and  $3.3 R_\oplus$  at 100 days (Burke & Catanzarite 2021). (Added: These limits account for the spot-induced variability in the system through a correction based on the Combined Differential Photometric Precision in the Kepler 1627 light curve over the relevant transit timescales (Burke & Catanzarite 2017).) The large size of Kepler 1627Ab relative to most Kepler mini-Neptunes might therefore support a picture in which the typical  $5 M_\oplus$  (Deleted: to  $10 M_\oplus$ ) mini-Neptune (Wu 2019) loses a significant fraction of its primordial atmosphere over its first gi-

gayear (Owen & Wu 2013; Ginzburg et al. 2018). (Added: It could also be consistent with a scenario in which an earlier “boil-off” of the planet’s atmosphere during disk dispersal decreases the entropy of the planetary interior, leading to a  $\sim 10^8$  year Kelvin-Helmholtz contraction timescale (Owen 2020). Confirming either of these scenarios would require a measurement of the planetary mass; otherwise, alternative explanations for its large size could include that it is abnormally massive, or that it has an abnormally large envelope to core mass ratio.) Ultimately, the main advance of this work is a precise measurement of the age of Kepler 1627Ab. This measurement was enabled by identifying the connection of the star to the  $\delta$  Lyr cluster using Gaia kinematics, and by then using the Gaia color-absolute magnitude diagram and TESS stellar rotation periods to verify the cluster’s existence. Table 3 enables similar cross-matches for both known and forthcoming exoplanet systems (e.g., Guerrero et al. 2021). Confirming these candidate associations using independent age indicators is essential because their false positive rates are not known. A related path is to identify new kinematic associations around known exoplanet host stars using po-

sitions and tangential velocities from Gaia, and to verify these associations with stellar rotation periods and spectroscopy (*e.g.*, Tofflemire et al. 2021). Each approach seems likely to expand the census of planets with precisely measured ages over the coming years, which will help in deciphering the early stages of exoplanet evolution.

## ACKNOWLEDGMENTS

The authors are grateful to J. Winn, J. Spake, A. Howard, and T. David for illuminating discussions and suggestions, and to R. Kerr for providing us with the Kerr et al. (2021) membership list prior to its publication. The authors are also grateful to K. Collins for helping resolve the scheduling conflict that would have otherwise prevented the MuSCAT3 observations. L.G.B. acknowledges support from a Charlotte Elizabeth Procter Fellowship from Princeton University, as well as from the TESS GI Program (NASA grants 80NSSC19K0386 and 80NSSC19K1728) and the Heising-Simons Foundation (51 Pegasi b Fellowship). Keck/NIRC2 imaging was acquired by program 2015A/N301N2L (PI: A. Kraus). In addition, this paper is based in part on observations made with the MuSCAT3 instrument, developed by the Astrobiology Center and under financial support by JSPS KAKENHI (JP18H05439) and JST PRESTO (JPMJPR1775), at Faulkes Telescope North on Maui, HI, operated by the Las Cumbres Observatory. This work is partly supported by JSPS KAKENHI Grant Numbers 22000005, JP15H02063, JP17H04574, JP18H05439, JP18H05442, JST PRESTO Grant Number JPMJPR1775, the Astrobiology Center of National Institutes of Natural Sciences (NINS) (Grant Number AB031010). This paper also includes data collected by the TESS mission, which are publicly available from the Mikulski Archive for Space Telescopes (MAST). Funding for the TESS mission is provided by NASA's Science Mission directorate. We thank the TESS Architects (G. Ricker, R. Vanderspek, D. Latham, S. Seager, J. Jenkins) and the many TESS team members for their efforts to make the mission a continued success. Finally, this research has made use of the Keck Observatory Archive (KOA), which is operated by the W. M. Keck Observatory and the NASA Exoplanet Science Institute (NExSci), under contract with the National Aeronautics and Space Administration. We also thank the Keck Observatory staff for their support of HIRES and remote observing. We recognize the importance that the summit of Maunakea has within the indigenous Hawaiian community, and are deeply grateful to have the opportunity to conduct observations from this mountain.

**Software:** `altaipony` (Ilin et al. 2021), `astrobase` (Bhatti et al. 2018), `astropy` (Astropy Collaboration et al. 2018), `astroquery` (Ginsburg et al. 2018), `corner` (Foreman-Mackey 2016), `exoplanet` (Foreman-Mackey et al. 2020), and its dependencies (Agol et al. 2020; Kip-

ping 2013; Luger et al. 2019; Theano Development Team 2016), `PyMC3` (Salvatier et al. 2016), `scipy` (Jones et al. 2001), `TESS-point` (Burke et al. 2020), `wotan` (Hippke et al. 2019).

**Facilities:** *Astrometry:* Gaia (Gaia Collaboration et al. 2018b, 2021a). *Imaging:* Second Generation Digitized Sky Survey. Keck:II (NIRC2; [www2.keck.hawaii.edu/inst/nirc2](http://www2.keck.hawaii.edu/inst/nirc2)). Gemini:North ('Alopeke; Scott et al. 2018, 2021). *Spectroscopy:* Keck:I (HIRES; Vogt et al. 1994). *Photometry:* Kepler (Borucki et al. 2010), MuSCAT3 (Narita et al. 2020), TESS (Ricker et al. 2015).

**Table 1.** Literature and Measured Properties for Kepler 1627

Primary Star			
TIC 120105470 GAIADR2 <sup>†</sup> 2103737241426734336			
Parameter	Description	Value	Source
$\alpha_{J2015.5}$	Right Ascension (hh:mm:ss) .....	18:56:13.6	1
$\delta_{J2015.5}$	Declination (dd:mm:ss) .....	+41:34:36.22	1
V .....	Johnson V mag. ....	13.11 $\pm$ 0.08	2
G .....	Gaia G mag. ....	13.02 $\pm$ 0.02	1
$G_{BP}$ .....	Gaia BP mag. ....	13.43 $\pm$ 0.02	1
$G_{RP}$ .....	Gaia RP mag. ....	12.44 $\pm$ 0.02	1
T .....	TESS T mag. ....	12.53 $\pm$ 0.02	2
J .....	2MASS J mag. ....	11.69 $\pm$ 0.02	3
H .....	2MASS H mag. ....	11.30 $\pm$ 0.02	3
K <sub>S</sub> .....	2MASS K <sub>S</sub> mag. ....	11.19 $\pm$ 0.02	3
$\pi$ .....	Gaia EDR3 parallax (mas) .....	3.009 $\pm$ 0.032	1
$d$ .....	Distance (pc) .....	329.5 $\pm$ 3.5	1, 4
$\mu_\alpha$ .....	Gaia EDR3 proper motion .....	1.716 $\pm$ 0.034	1
	in RA (mas yr <sup>-1</sup> )		
$\mu_\delta$ .....	Gaia EDR3 proper motion .....	-1.315 $\pm$ 0.034	1
	in DEC (mas yr <sup>-1</sup> )		
RUWE .....	Gaia EDR3 renormalized .....	2.899	1
	unit weight error		
RV .....	Systemic radial velocity (km s <sup>-1</sup> ) .....	-16.7 $\pm$ 1.0	5
Spec. Type .....	Spectral Type .....	G8V	5
$v \sin i_*$ .....	Rotational velocity* (km s <sup>-1</sup> ) .....	18.9 $\pm$ 1.0	5
Li EW .....	6708Å Equiv. Width (mÅ) .....	235 <sup>+5</sup> <sub>-7</sub>	5
$T_{\text{eff}}$ .....	Effective Temperature (K) .....	5505 $\pm$ 60	6
$\log g_*$ .....	Surface Gravity (cgs) .....	4.53 $\pm$ 0.05	6
$R_*$ .....	Stellar radius ( $R_\odot$ ) .....	0.881 $\pm$ 0.018	6
$M_*$ .....	Stellar mass ( $R_\odot$ ) .....	0.953 $\pm$ 0.019	6
Av .....	Interstellar reddening (mag) .....	0.2 $\pm$ 0.1	6
[Fe/H] .....	Metallicity .....	0.1 $\pm$ 0.1	6
$P_{\text{rot}}$ .....	Rotation period (d) .....	2.642 $\pm$ 0.042	7
Age .....	Adopted stellar age (Myr) .....	38 <sup>+6</sup> <sub>-5</sub>	8
$\Delta m_{832}$	Mag difference ('Alopeke 832 nm)	3.14 $\pm$ 0.15	9
$\theta_B$	Position angle (deg) .....	92 $\pm$ 1	9
$\rho_B$	Apparent separation of .....	0.164 $\pm$ 0.002	9
	primary and secondary (as)		
$\rho_B$	Apparent separation of .....	53 $\pm$ 4	1, 4, 9
	primary and secondary (AU)		
$\Delta m_{K'}$	Mag difference (NIRC2 $K'$ ) .....	2.37 $\pm$ 0.02	10
$\theta_B$	Position angle (deg) .....	95.9 $\pm$ 0.5	10
$\rho_B$	Apparent separation of .....	0.1739 $\pm$ 0.0010	10
	primary and secondary (as)		

NOTE—<sup>†</sup> The GAIADR2 and GAIAEDR3 identifiers for Kepler 1627A are identical. The secondary is not resolved in the Gaia point source catalog. \* Given only  $v \sin i$  and  $2\pi R_*/P_{\text{rot}}$ ,  $\cos i = 0.11_{-0.08}^{+0.11}$ . Provenances are: <sup>1</sup>Gaia Collaboration et al. (2021a), <sup>2</sup>Stassun et al. (2019), <sup>3</sup>Skrutskie et al. (2006), <sup>4</sup>Lindgren et al. (2021), <sup>5</sup>HIREs spectra and Yee et al. (2017), <sup>6</sup>Cluster isochrone (MIST adopted; PARSEC compared for quoted uncertainty), <sup>7</sup>Kepler light curve, <sup>8</sup>Pre-main-sequence CAMD interpolation (Section 2.2.1), <sup>9</sup>'Alopeke imaging 2021 June 24 (Scott et al. 2021), <sup>10</sup>NIRC2 imaging 2015 July 22, using the Yelda et al. (2010) optical distortion solution to convert pixel-space relative positions to on-sky relative astrometry. The “discrepancy” between the two imaging epochs likely indicates orbital motion.

**Table 2.** Priors and Posteriors for Model Fitted to the Long Cadence Kepler 1627Ab Light Curve.

Param.	Unit	Prior	Median	Mean	Std. Dev.	3%	97%	ESS	$\hat{R} - 1$
<i>Sampled</i>									
$P$	d	$\mathcal{N}(7.20281; 0.01000)$	7.2028038	7.2028038	0.0000073	7.2027895	7.2028168	7464	3.9e-04
$t_0^{(1)}$	d	$\mathcal{N}(120.79053; 0.02000)$	120.7904317	120.7904254	0.0009570	120.7886377	120.7921911	3880	2.0e-03
$\log \delta$	—	$\mathcal{N}(-6.3200; 2.0000)$	-6.3430	-6.3434	0.0354	-6.4094	-6.2767	6457	3.0e-04
$b^{(2)}$	—	$\mathcal{U}(0.000; 1.000)$	0.4669	0.4442	0.2025	0.0662	0.8133	1154	1.6e-03
$u_1$	—	Kipping (2013)	0.271	0.294	0.190	0.000	0.628	3604	1.5e-03
$u_2$	—	Kipping (2013)	0.414	0.377	0.326	-0.240	0.902	3209	1.4e-03
$R_*$	$R_{\odot}$	$\mathcal{N}(0.881; 0.018)$	0.881	0.881	0.018	0.847	0.915	8977	3.1e-04
$\log g$	cgs	$\mathcal{N}(4.530; 0.050)$	4.532	4.533	0.051	4.435	4.627	6844	1.6e-03
$\langle f \rangle$	—	$\mathcal{N}(0.000; 0.100)$	-0.0003	-0.0003	0.0001	-0.0005	-0.0000	8328	1.1e-03
$e^{(3)}$	—	Van Eylen et al. (2019)	0.154	0.186	0.152	0.000	0.459	1867	2.0e-03
$\omega$	rad	$\mathcal{U}(0.000; 6.283)$	0.055	0.029	1.845	-3.139	2.850	3557	8.6e-05
$\log \sigma_f$	—	$\mathcal{N}(\log \langle \sigma_f \rangle; 2.000)$	-8.035	-8.035	0.008	-8.049	-8.021	9590	3.9e-04
$\sigma_{\text{rot}}$	$d^{-1}$	InvGamma(1.000; 5.000)	0.070	0.070	0.001	0.068	0.072	9419	1.4e-03
$\log P_{\text{rot}}$	log(d)	$\mathcal{N}(0.958; 0.020)$	0.978	0.978	0.001	0.975	0.980	8320	2.2e-04
$\log Q_0$	—	$\mathcal{N}(0.000; 2.000)$	-0.327	-0.326	0.043	-0.407	-0.246	9659	2.7e-04
$\log dQ$	—	$\mathcal{N}(0.000; 2.000)$	7.697	7.698	0.103	7.511	7.899	5824	3.7e-04
$f$	—	$\mathcal{U}(0.010; 1.000)$	0.01006	0.01009	0.00009	0.01000	0.01025	4645	4.0e-04
<i>Derived</i>									
$\delta$	—	—	0.001759	0.001759	0.000062	0.001641	0.001875	6457	3.0e-04
$R_p/R_*$	—	—	0.039	0.039	0.001	0.037	0.042	1811	1.1e-03
$\rho_*$	$\text{g cm}^{-3}$	—	1.990	2.004	0.240	1.570	2.461	6905	2.1e-03
$R_p^{(4)}$	$R_{\text{Jup}}$	—	0.337	0.338	0.014	0.314	0.367	2311	1.0e-03
$R_p^{(4)}$	$R_{\text{Earth}}$	—	3.777	3.789	0.157	3.52	4.114	2311	1.0e-03
$a/R_*$	—	—	17.606	17.619	0.702	16.277	18.906	6905	2.1e-03
$\cos i$	—	—	0.027	0.025	0.010	0.004	0.040	1312	1.2e-03
$T_{14}$	hr	—	2.841	2.843	0.060	2.734	2.958	3199	3.6e-04
$T_{13}$	hr	—	2.555	2.539	0.094	2.360	2.692	1960	1.4e-03

NOTE— ESS refers to the number of effective samples.  $\hat{R}$  is the Gelman-Rubin convergence diagnostic. Logarithms in this table are base-e.  $\mathcal{U}$  denotes a uniform distribution, and  $\mathcal{N}$  a normal distribution. (1) The ephemeris is in units of BJDTDB - 2454833. (2) Although  $\mathcal{U}(0, 1 + R_p/R_*)$  is formally correct, for this model we assumed a non-grazing transit to enable sampling in  $\log \delta$ . (3) The eccentricity vectors are sampled in the  $(e \cos \omega, e \sin \omega)$  plane. (4) The true planet size is a factor of  $((F_1 + F_2)/F_1)^{1/2}$  larger than that from the fit because of dilution from Kepler 1627B, where  $F_1$  is the flux from the primary, and  $F_2$  is that from the secondary; the mean and standard deviation of  $R_p = 3.817 \pm 0.158 R_{\oplus}$  quoted in the text includes this correction, assuming  $(F_1 + F_2)/F_1 \approx 1.015$ .

**Table 3.** Young, Age-dated, and Age-dateable Stars Within the Nearest Few Kiloparsecs.

Parameter	Example Value	Description
source_id	1709456705329541504	Gaia DR2 source identifier.
ra	247.826	Gaia DR2 right ascension [deg].
dec	79.789	Gaia DR2 declination [deg].
parallax	35.345	Gaia DR2 parallax [mas].
parallax_error	0.028	Gaia DR2 parallax uncertainty [mas].
pmra	94.884	Gaia DR2 proper motion $\mu_{\alpha} \cos \delta$ [mas yr $^{-1}$ ].
pmdec	-86.971	Gaia DR2 proper motion $\mu_{\delta}$ [mas yr $^{-1}$ ].
phot_g_mean_mag	6.85	Gaia DR2 $G$ magnitude.
phot_bp_mean_mag	6.409	Gaia DR2 $G_{\text{BP}}$ magnitude.
phot_rp_mean_mag	7.189	Gaia DR2 $G_{\text{RP}}$ magnitude.
cluster	NASAExoArchive_ps_20210506.Uma,IR_excess	Comma-separated cluster or group name.
age	9.48,nan,nan	Comma-separated logarithm (base-10) of reported <sup>a</sup> age in years.
mean_age	9.48	Mean (ignoring NaNs) of age column.
reference_id	NASAExoArchive_ps_20210506,Ujjwal2020,CottenSong2016	Comma-separated provenance of group membership.
reference_bibcode	2013PASP..125..989A,2020AJ....159..166U,2016ApJS..225..15C	ADS bibcode corresponding to reference_id.

NOTE— Table 3 is published in its entirety in a machine-readable format. This table is a concatenation of the studies listed in Table 4. One entry is shown for guidance regarding form and content. In this particular example, the star has a cold Jupiter on a 16 year orbit, HD 150706b (Boisse et al. 2012). An infrared excess has been reported (Cotten & Song 2016), and the star was identified by Ujjwal et al. (2020) as a candidate UMa moving group member ( $\approx 400$  Myr; Mann et al. 2020). The star's RV activity and TESS rotation period corroborate its youth.

**Table 4.** Provenances of Young and Age-dateable Stars.

Reference	$N_{\text{Gaia}}$	$N_{\text{Age}}$	$N_{G_{\text{RP}} < 16}$
Kounkel et al. (2020)	987376	987376	775363
Cantat-Gaudin & Anders (2020)	433669	412671	269566
Cantat-Gaudin et al. (2018)	399654	381837	246067
Kounkel & Covey (2019)	288370	288370	229506
Cantat-Gaudin et al. (2020)	233369	227370	183974
Zari et al. (2018) UMS	86102	0	86102
Wenger et al. (2000) Y*?	61432	0	45076
Zari et al. (2018) PMS	43719	0	38435
Gaia Collaboration et al. (2018a) $d > 250 \text{ pc}$	35506	31182	18830
Castro-Ginard et al. (2020)	33635	24834	31662
Kerr et al. (2021)	30518	25324	27307
Wenger et al. (2000) Y*○	28406	0	16205
Villa Vélez et al. (2018)	14459	14459	13866
Cantat-Gaudin et al. (2019)	11843	11843	9246
Damiani et al. (2019) PMS	10839	10839	9901
Oh et al. (2017)	10379	0	10370
Meingast et al. (2021)	7925	7925	5878
Wenger et al. (2000) pMS*	5901	0	3006
Gaia Collaboration et al. (2018a) $d < 250 \text{ pc}$	5378	817	3968
Kounkel et al. (2018)	5207	3740	5207
Ratzenböck et al. (2020)	4269	4269	2662
Wenger et al. (2000) TT*	4022	0	3344
Damiani et al. (2019) UMS	3598	3598	3598
Rizzuto et al. (2017)	3294	3294	2757
Akeson et al. (2013)	3107	868	3098
Tian (2020)	1989	1989	1394
Goldman et al. (2018)	1844	1844	1783
Cotten & Song (2016)	1695	0	1693
Gagné et al. (2018b)	1429	0	1389
Röser & Schilbach (2020) Psc-Eri	1387	1387	1107
Röser & Schilbach (2020) Pleiades	1245	1245	1019
Wenger et al. (2000) TT?	1198	0	853
Gagné & Faherty (2018)	914	0	913
Pavlidou et al. (2021)	913	913	504
Gagné et al. (2018a)	692	0	692
Ujjwal et al. (2020)	563	0	563
Gagné et al. (2020)	566	566	351
Esplin & Luhman (2019)	377	443	296
Roccatagliata et al. (2020)	283	283	232
Meingast & Alves (2019)	238	238	238
Fürnkranz et al. (2019) Coma-Ber	214	214	213
Fürnkranz et al. (2019) Neighbor Group	177	177	167
Kraus et al. (2014)	145	145	145

NOTE— Table 4 describes the provenances for the young and age-dateable stars in Table 3.  $N_{\text{Gaia}}$ : number of Gaia stars we parsed from the literature source.  $N_{\text{Age}}$ : number of stars in the literature source with ages reported.  $N_{G_{\text{RP}} < 16}$ : number of Gaia stars we parsed from the literature source with either  $G_{\text{RP}} < 16$ , or a parallax S/N exceeding 5 and a distance closer than 100 pc. The latter criterion included a few hundred white dwarfs that would have otherwise been neglected. Some studies are listed multiple times if they contain multiple tables. Wenger et al. (2000) refers to the SIMBAD database.

## REFERENCES

- 972 Agol, E., Luger, R., & Foreman-Mackey, D. 2020, *AJ*,  
973 159, 123
- 974 Akeson, R. L., Chen, X., Ciardi, D., et al. 2013, *PASP*,  
975 125, 989
- 976 Astropy Collaboration, Price-Whelan, A. M., Sipőcz,  
977 B. M., et al. 2018, *AJ*, 156, 123
- 978 Baraffe, I., Homeier, D., Allard, F., & Chabrier, G.  
979 2015, *A&A*, 577, A42
- 980 Beaumont, C., Robitaille, T., Borkin, M., & Goodman,  
981 A. 2014, glueviz v0.4: multidimensional data  
982 exploration
- 983 Benatti, S., Nardiello, D., Malavolta, L., et al. 2019,  
984 *A&A*, 630, A81
- 985 Berger, T. A., Howard, A. W., & Boesgaard, A. M.  
986 2018, *ApJ*, 855, 115
- 987 Berger, T. A., Huber, D., Gaidos, E., & van Saders, J. L.  
988 2018, *ApJ*, 866, 99
- 989 Bhatti, W., Bouma, L. G., & Wallace, J. 2018,  
990 astrobase,  
991 <https://doi.org/10.5281/zenodo.1469822>
- 992 Boisse, I., Pepe, F., Perrier, C., et al. 2012, *A&A*, 545,  
993 A55
- 994 Borucki, W. J., Koch, D., Basri, G., et al. 2010, *Science*,  
995 327, 977
- 996 Bossini, D., Vallenari, A., Bragaglia, A., et al. 2019,  
997 *A&A*, 623, A108
- 998 Bouma, L. G., Curtis, J. L., Hartman, J. D., Winn, J. N.,  
999 & Bakos, G. A. 2021, arXiv:2107.08050 [astro-ph]
- 1000 Bouma, L. G., Hartman, J. D., Bhatti, W., Winn, J. N.,  
1001 & Bakos, G. Á. 2019, *ApJS*, 245, 13
- 1002 Bouma, L. G., Hartman, J. D., Brahm, R., et al. 2020,  
1003 *AJ*, 160, 239
- 1004 Bressan, A., Marigo, P., Girardi, L., et al. 2012,  
1005 *MNRAS*, 427, 127
- 1006 Burke, C., & Catanzarite, J. 2021, KeplerPORTS:  
1007 Kepler Planet Occurrence Rate Tools
- 1008 Burke, C. J., & Catanzarite, J. 2017, Planet Detection  
1009 Metrics: Per-Target Detection Contours for Data  
1010 Release 25, Kepler Science Document  
1011 KSCI-19111-002
- 1012 Burke, C. J., Levine, A., Fausnaugh, M., et al. 2020,  
1013 TESS-Point: High precision TESS pointing tool,  
1014 Astrophysics Source Code Library, ascl:2003.001
- 1015 Cantat-Gaudin, T., & Anders, F. 2020, *A&A*, 633, A99
- 1016 Cantat-Gaudin, T., Jordi, C., Vallenari, A., et al. 2018,  
1017 *A&A*, 618, A93
- 1018 Cantat-Gaudin, T., Jordi, C., Wright, N. J., et al. 2019,  
1019 *A&A*, 626, A17
- 1020 Cantat-Gaudin, T., Anders, F., Castro-Ginard, A., et al.  
1021 2020, *A&A*, 640, A1
- 1022 Capitanio, L., Lallement, R., Vergely, J. L., Elyajouri,  
1023 M., & Monreal-Ibero, A. 2017, *A&A*, 606, A65
- 1024 Castro-Ginard, A., Jordi, C., Luri, X., et al. 2020, *A&A*,  
1025 635, A45
- 1026 Choi, J., Dotter, A., Conroy, C., et al. 2016, *ApJ*, 823,  
1027 102
- 1028 Claret, A., & Bloemen, S. 2011, *A&A*, 529, A75
- 1029 Cotten, T. H., & Song, I. 2016, *ApJS*, 225, 15
- 1030 Dahm, S. E. 2015, *ApJ*, 813, 108
- 1031 Dai, F., Winn, J. N., Berta-Thompson, Z.,  
1032 Sanchis-Ojeda, R., & Albrecht, S. 2018, *AJ*, 155, 177
- 1033 Damiani, F., Prisinzano, L., Pillitteri, I., Micela, G., &  
1034 Sciotino, S. 2019, *A&A*, 623, A112
- 1035 Davenport, J. R. A. 2016, *ApJ*, 829, 23
- 1036 Davenport, J. R. A., Hawley, S. L., Hebb, L., et al.  
1037 2014, *ApJ*, 797, 122
- 1038 David, T., Hillenbrand, L., & Petigura, E. 2016, *Nature*,  
1039 534, 658
- 1040 David, T. J., & Hillenbrand, L. A. 2015, *ApJ*, 804, 146
- 1041 David, T. J., Petigura, E. A., Luger, R., et al. 2019,  
1042 *ApJL*, 885, L12
- 1043 David, T. J., Contardo, G., Sandoval, A., et al. 2021, *AJ*,  
1044 161, 265
- 1045 Dias, W. S., Monteiro, H., Caetano, T. C., et al. 2014,  
1046 *A&A*, 564, A79
- 1047 Dobbie, P. D., Lodieu, N., & Sharp, R. G. 2010,  
1048 *MNRAS*, 409, 1002
- 1049 Douglas, S. T., Agüeros, M. A., Covey, K. R., & Kraus,  
1050 A. 2017, *ApJ*, 842, 83
- 1051 Douglas, S. T., Pérez Chávez, J., Cargile, P. A., et al.  
1052 2021, 10.5281/zenodo.5131306
- 1053 Eggen, O. J. 1968, *ApJ*, 152, 77
- 1054 Ehrenreich, D., Bourrier, V., Wheatley, P. J., et al. 2015,  
1055 *Nature*, 522, 459
- 1056 Esplin, T. L., & Luhman, K. L. 2019, *AJ*, 158, 54
- 1057 Feinstein, A. D., Montet, B. T., Johnson, M. C., et al.  
1058 2021, arXiv:2107.01213 [astro-ph]
- 1059 Foreman-Mackey, D. 2016, *Journal of Open Source  
1060 Software*, 1, 24
- 1061 Foreman-Mackey, D., Czekala, I., Luger, R., et al. 2020,  
1062 exoplanet-dev/exoplanet v0.2.6
- 1063 Fulton, B. J., Petigura, E. A., Howard, A. W., et al.  
1064 2017, *AJ*, 154, 109
- 1065 Fürnkranz, V., Meingast, S., & Alves, J. 2019, *A&A*,  
1066 624, L11
- 1067 Gagné, J., David, T. J., Mamajek, E. E., et al. 2020,  
1068 *ApJ*, 903, 96

- 1069 Gagné, J., & Faherty, J. K. 2018, *ApJ*, 862, 138  
 1070 Gagné, J., Roy-Loubier, O., Faherty, J. K., Doyon, R.,  
     & Malo, L. 2018a, *ApJ*, 860, 43  
 1072 Gagné, J., Mamajek, E. E., Malo, L., et al. 2018b, *ApJ*,  
     856, 23  
 1074 Gagné, J., David, T. J., Mamajek, E. E., et al. 2020,  
     *ApJ*, 903, 96  
 1076 Gaia Collaboration, Babusiaux, C., van Leeuwen, F.,  
     et al. 2018a, *A&A*, 616, A10  
 1078 Gaia Collaboration, Brown, A. G. A., Vallenari, A.,  
     et al. 2018b, *A&A*, 616, A1  
 1080 —. 2021a, *A&A*, 649, A1  
 1081 Gaia Collaboration, Smart, R. L., Sarro, L. M., et al.  
     2021b, *A&A*, 649, A6  
 1083 Gelman, A., & Rubin, D. B. 1992, *Statistical Science*, 7,  
     457, publisher: Institute of Mathematical Statistics  
 1085 Gibson, S. R., Howard, A. W., Marcy, G. W., et al. 2016,  
     in SPIE Conference Series, Vol. 9908, *Ground-based*  
     and *Airborne Instrumentation for Astronomy VI*, ed.  
     C. J. Evans, L. Simard, & H. Takami, 990870  
 1089 Gillen, E., Briegal, J. T., Hodgkin, S. T., et al. 2020,  
     *MNRAS*, 492, 1008  
 1091 Ginsburg, A., Sipocz, B., Madhura Parikh, et al. 2018,  
     Astropy/Astroquery: V0.3.7 Release  
 1093 Ginsburg, S., Schlichting, H. E., & Sari, R. 2018,  
     *MNRAS*, 476, 759  
 1095 Goldman, B., Röser, S., Schilbach, E., Moór, A. C., &  
     Henning, T. 2018, *ApJ*, 868, 32  
 1097 Guerrero, N. M., Seager, S., Huang, C. X., et al. 2021,  
     *ApJS*, 254, 39  
 1099 Günther, M. N., Zhan, Z., Seager, S., et al. 2020, *AJ*,  
     159, 60  
 1101 Gupta, A., & Schlichting, H. E. 2020, *MNRAS*, 493,  
     792  
 1103 Hattori, S., Foreman-Mackey, D., Hogg, D. W., et al.  
     2021, arXiv e-prints, arXiv:2106.15063  
 1105 Hippke, M., David, T. J., Mulders, G. D., & Heller, R.  
     2019, *AJ*, 158, 143  
 1107 Hirano, T., Suto, Y., Taruya, A., et al. 2010, *ApJ*, 709,  
     458  
 1109 Hirano, T., Suto, Y., Winn, J. N., et al. 2011, *ApJ*, 742,  
     69  
 1111 Hoffman, M. D., & Gelman, A. 2014, *Journal of*  
     *Machine Learning Research*, 15, 1593  
 1113 Holczer, T., Shporer, A., Mazeh, T., et al. 2015, *ApJ*,  
     807, 170  
 1115 Holczer, T., Mazeh, T., Nachmani, G., et al. 2016,  
     *ApJS*, 225, 9  
 1117 Howard, A. W., Johnson, J. A., Marcy, G. W., et al.  
     2010, *ApJ*, 721, 1467  
 1119 Howell, S. B., Everett, M. E., Sherry, W., Horch, E., &  
     Ciardi, D. R. 2011, *AJ*, 142, 19  
 1121 Ikoma, M., & Hori, Y. 2012, *ApJ*, 753, 66  
 1122 Ilin, E., Schmidt, S. J., Poppenhöger, K., et al. 2021,  
     *A&A*, 645, A42  
 1124 Jones, E., Oliphant, T., Peterson, P., et al. 2001, Open  
     source scientific tools for Python  
 1126 Kass, R. E., & Raftery, A. E. 1995, *Journal of the*  
     *American Statistical Association*, 90, 773  
 1128 Kerr, R. M. P., Rizzuto, A. C., Kraus, A. L., & Offner,  
     S. S. R. 2021, *ApJ*, 917, 23  
 1130 Kharchenko, N. V., Piskunov, A. E., Schilbach, E.,  
     Röser, S., & Scholz, R.-D. 2013, *A&A*, 558, A53  
 1132 King, G. W., & Wheatley, P. J. 2021, *MNRAS*, 501, L28  
 1133 Kipping, D. M. 2010, *MNRAS*, 408, 1758  
 1134 Kipping, D. M. 2013, *MNRAS*, 435, 2152  
 1135 Kite, E. S., Fegley, Jr., B., Schaefer, L., & Ford, E. B.  
     2020, *ApJ*, 891, 111  
 1137 Klein, B., & Donati, J.-F. 2020, *MNRAS*, 493, L92  
 1138 Kounkel, M., & Covey, K. 2019, *AJ*, 158, 122  
 1139 Kounkel, M., Covey, K., & Stassun, K. G. 2020, *AJ*,  
     160, 279  
 1141 Kounkel, M., Covey, K., Suárez, G., et al. 2018, *AJ*,  
     156, 84  
 1143 Kounkel, M., Covey, K., Suárez, G., et al. 2018, *AJ*,  
     156, 84  
 1145 Kraus, A. L., Ireland, M. J., Huber, D., Mann, A. W., &  
     Dupuy, T. J. 2016, *AJ*, 152, 8  
 1147 Kraus, A. L., Shkolnik, E. L., Allers, K. N., & Liu,  
     M. C. 2014, *AJ*, 147, 146  
 1149 Lallement, R., Capitanio, L., Ruiz-Dern, L., et al. 2018,  
     *A&A*, 616, A132  
 1151 Lindegren, L., Bastian, U., Biermann, M., et al. 2021,  
     *A&A*, 649, A4  
 1153 Lomb, N. R. 1976, *Astrophysics and Space Science*, 39,  
     447  
 1155 Lopez, E. D., Fortney, J. J., & Miller, N. 2012, *ApJ*,  
     761, 59  
 1157 Luger, R., Agol, E., Foreman-Mackey, D., et al. 2019,  
     *AJ*, 157, 64  
 1159 Mann, A. W., Newton, E. R., Rizzuto, A. C., et al. 2016,  
     *AJ*, 152, 61  
 1161 Mann, A. W., Johnson, M. C., Vanderburg, A., et al.  
     2020, *AJ*, 160, 179  
 1163 Martioli, E., Hébrard, G., Correia, A. C. M., Laskar, J.,  
     & Lecavelier des Etangs, A. 2021, *A&A*, 649, A177  
 1165 Masuda, K. 2015, *ApJ*, 805, 28  
 1166 Mazeh, T., Holczer, T., & Shporer, A. 2015, *ApJ*, 800,  
     142  
 1168 McCann, J., Murray-Clay, R. A., Kratter, K., &  
     Krumholz, M. R. 2019, *ApJ*, 873, 89  
 1169

- 1170 McKinney, W. 2010, in Proceedings of the 9th Python  
 1171 in Science Conference, ed. S. van der Walt &  
 1172 J. Millman, 51
- 1173 McQuillan, A., Mazeh, T., & Aigrain, S. 2014, *ApJS*,  
 1174 211, 24
- 1175 Meibom, S., Mathieu, R. D., & Stassun, K. G. 2007,  
 1176 *ApJL*, 665, L155
- 1177 Meibom, S., Barnes, S. A., Latham, D. W., et al. 2011,  
 1178 *ApJL*, 733, L9
- 1179 Meibom, S., Torres, G., Fressin, F., et al. 2013, *Nature*,  
 1180 499, 55
- 1181 Meingast, S., & Alves, J. 2019, *A&A*, 621, L3
- 1182 Meingast, S., Alves, J., & Rottensteiner, A. 2021, *A&A*,  
 1183 645, A84
- 1184 Montet, B. T., Tovar, G., & Foreman-Mackey, D. 2017,  
 1185 *ApJ*, 851, 116
- 1186 Montet, B. T., Feinstein, A. D., Luger, R., et al. 2020,  
 1187 *AJ*, 159, 112
- 1188 Morris, B. M. 2020, *ApJ*, 893, 67
- 1189 Morton, T. D., Bryson, S. T., Coughlin, J. L., et al.  
 1190 2016, *ApJ*, 822, 86
- 1191 Narita, N., Fukui, A., Yamamuro, T., et al. 2020, in  
 1192 *SPIE Conference Series*, 114475K
- 1193 Newton, E. R., Mann, A. W., Tofflemire, B. M., et al.  
 1194 2019, *ApJ*, 880, L17
- 1195 Oh, S., Price-Whelan, A. M., Hogg, D. W., Morton,  
 1196 T. D., & Spergel, D. N. 2017, *AJ*, 153, 257
- 1197 Owen, J. E. 2020, *MNRAS*, 498, 5030
- 1198 Owen, J. E., & Wu, Y. 2013, *ApJ*, 775, 105  
 1199 —. 2016, *ApJ*, 817, 107  
 1200 —. 2017, *ApJ*, 847, 29
- 1201 Palle, E., Oshagh, M., Casasayas-Barris, N., et al. 2020,  
 1202 *A&A*, 643, 25
- 1203 Pavlidou, T., Scholz, A., & Teixeira, P. S. 2021,  
 1204 *MNRAS*, 503, 3232
- 1205 Pecaut, M. J., & Mamajek, E. E. 2013, *ApJS*, 208, 9
- 1206 Pecaut, M. J., & Mamajek, E. E. 2016, *MNRAS*, 461,  
 1207 794
- 1208 Plavchan, P., Barclay, T., Gagné, J., et al. 2020, *Nature*,  
 1209 582, 497
- 1210 Poppenhaeger, K., Ketzer, L., & Mallonn, M. 2021,  
 1211 *MNRAS*, 500, 4560
- 1212 Randich, S., Tognelli, E., Jackson, R., et al. 2018, *A&A*,  
 1213 612, A99
- 1214 Ratzenböck, S., Meingast, S., Alves, J., Möller, T., &  
 1215 Bomze, I. 2020, *A&A*, 639, A64
- 1216 Rebull, L. M., Stauffer, J. R., Bouvier, J., et al. 2016,  
 1217 *AJ*, 152, 113
- 1218 Reinhold, T., & Hekker, S. 2020, *A&A*, 635, A43
- 1219 Ricker, G. R., Winn, J. N., Vanderspek, R., et al. 2015,  
 1220 *JATIS*, 1, 014003
- 1221 Rizzuto, A. C., Mann, A. W., Vanderburg, A., Kraus,  
 1222 A. L., & Covey, K. R. 2017, *AJ*, 154, 224
- 1223 Rizzuto, A. C., Newton, E. R., Mann, A. W., et al. 2020,  
 1224 *AJ*, 160, 33
- 1225 Roccatagliata, V., Franciosini, E., Sacco, G. G.,  
 1226 Randich, S., & Sicilia-Aguilar, A. 2020, *A&A*, 638,  
 1227 A85
- 1228 Roettenbacher, R. M., Monnier, J. D., Korhonen, H.,  
 1229 et al. 2017, *ApJ*, 849, 120
- 1230 Rogers, J. G., & Owen, J. E. 2021, *Monthly Notices of  
 1231 the Royal Astronomical Society*, 503, 1526, aDS  
 1232 Bibcode: 2021MNRAS.503.1526R
- 1233 Röser, S., & Schilbach, E. 2020, *A&A*, 638, A9
- 1234 Salvatier, J., Wieckiâ, T. V., & Fonnesbeck, C. 2016,  
 1235 PyMC3: Python probabilistic programming  
 1236 framework
- 1237 Scargle, J. D. 1982, *ApJ*, 263, 835
- 1238 Scott, N. J., Howell, S. B., Horch, E. P., & Everett,  
 1239 M. E. 2018, *PASP*, 130, 054502
- 1240 Scott, N. J., Howell, S. B., Gnilka, C. L., et al. 2021,  
 1241 *Frontiers in Astronomy and Space Sciences*, 8, 138
- 1242 Seifahrt, A., Stürmer, J., Bean, J. L., & Schwab, C.  
 1243 2018, in *SPIE Conference Series*, Vol. 10702,  
 1244 *Ground-based and Airborne Instrumentation for  
 1245 Astronomy VII*, ed. C. J. Evans, L. Simard, &  
 1246 H. Takami, 107026D
- 1247 Shapiro, A. I., Solanki, S. K., Krivova, N. A., Yeo,  
 1248 K. L., & Schmutz, W. K. 2016, *A&A*, 589, A46
- 1249 Shporer, A., Zhou, G., Vanderburg, A., et al. 2017,  
 1250 *ApJL*, 847, L18
- 1251 Skrutskie, M. F., Cutri, R. M., Stiening, R., et al. 2006,  
 1252 *AJ*, 131, 1163
- 1253 Smith, J. C., Morris, R. L., Jenkins, J. M., et al. 2016,  
 1254 *PASP*, 128, 124501
- 1255 Smith, J. C., Stumpe, M. C., Jenkins, J. M., et al. 2017,  
 1256 *Kepler Science Document*, 8
- 1257 Soderblom, D. R., Hillenbrand, L. A., Jeffries, R. D.,  
 1258 Mamajek, E. E., & Naylor, T. 2014, *Protostars and  
 1259 Planets VI*, 219
- 1260 Spake, J. J., Sing, D. K., Evans, T. M., et al. 2018,  
 1261 *Nature*, 557, 68
- 1262 Stassun, K. G., Oelkers, R. J., Paegert, M., et al. 2019,  
 1263 *AJ*, 158, 138
- 1264 Stauffer, J., Rebull, L., Bouvier, J., et al. 2016, *AJ*, 152,  
 1265 115
- 1266 Stauffer, J. R., Hartmann, L. W., Prosser, C. F., et al.  
 1267 1997, *ApJ*, 479, 776
- 1268 Stefansson, G., Mahadevan, S., Maney, M., et al. 2020,  
 1269 *AJ*, 160, 192
- 1270 Stephenson, C. B. 1959, *PASP*, 71, 145

- 1271 Strassmeier, K. G. 2009, *Astronomy and Astrophysics Review*, 17, 251
- 1273 Tenenbaum, P., Christiansen, J. L., Jenkins, J. M., et al. 2012, *ApJS*, 199, 24
- 1275 Theano Development Team. 2016, arXiv e-prints, [abs/1605.02688](https://arxiv.org/abs/1605.02688)
- 1277 Thompson, S. E., Coughlin, J. L., Hoffman, K., et al. 2018, *ApJS*, 235, 38
- 1279 Tian, H.-J. 2020, *ApJ*, 904, 196
- 1280 Tofflemire, B. M., Rizzuto, A. C., Newton, E. R., et al. 2021, *AJ*, 161, 171
- 1282 Ujjwal, K., Kartha, S. S., Mathew, B., Manoj, P., & Narang, M. 2020, *AJ*, 159, 166
- 1284 Van Eylen, V., Agentoft, C., Lundkvist, M. S., et al. 2018, *MNRAS*, 479, 4786
- 1286 Van Eylen, V., Albrecht, S., Huang, X., et al. 2019, *AJ*, 157, 61
- 1288 Villa Vélez, J. A., Brown, A. G. A., & Kenworthy, M. A. 2018, *RNAAS*, 2, 58
- 1290 Vissapragada, S., Knutson, H. A., Jovanovic, N., et al. 2020, *AJ*, 159, 278
- 1292 Vogt, S. S., Allen, S. L., Bigelow, B. C., et al. 1994, SPIE Conference Series, ed. D. L. Crawford & E. R. Craine, Vol. 2198
- 1295 Walkowicz, L. M., & Basri, G. S. 2013, *MNRAS*, 436, 1883, arXiv: 1309.2159
- 1297 Wang, L., & Dai, F. 2019, *ApJL*, 873, L1
- 1298 Wenger, M., Ochsenbein, F., Egret, D., et al. 2000, *A&AS*, 143, 9
- 1300 Wirth, C. P., Zhou, G., Quinn, S. N., et al. 2021, *ApJL*, 917, L34
- 1302 Wu, Y. 2019, *ApJ*, 874, 91
- 1303 Yee, S. W., Petigura, E. A., & von Braun, K. 2017, *ApJ*, 836, 77
- 1305 Yelda, S., Lu, J. R., Ghez, A. M., et al. 2010, *ApJ*, 725, 331
- 1307 Zari, E., Hashemi, H., Brown, A. G. A., Jardine, K., & de Zeeuw, P. T. 2018, *A&A*, 620, A172
- 1309 Zhou, G., Winn, J. N., Newton, E. R., et al. 2020, *ApJ*, 892, L21
- 1311 Zhou, G., Quinn, S. N., Irwin, J., et al. 2021, *AJ*, 161, 2

1312

## APPENDIX

1313

## A. YOUNG, AGE-DATED, AND AGE-DATEABLE STAR COMPILED

1314 The v0.6 CDIPS target catalog (Table 3) includes stars that are young, age-dated, and age-dateable. By  
 1315 “age-dateable”, we mean that the stellar age should be measurable at greater precision than that of a typical  
 1316 FGK field star, through either isochronal, gyrochronal, or spectroscopic techniques. As in Bouma et al.  
 1317 (2019), we collected stars that met these criteria from across the literature. Table 4 gives a list of the studies  
 1318 included, and brief summary statistics. The age measurement methodologies adopted by each study differ: in  
 1319 many, spatial and kinematic clustering has been performed on the Gaia data, and ensemble isochrone fitting  
 1320 of the resulting clusters has been performed (typically focusing on the turn-off). In other studies however,  
 1321 the claim of youth is based on the location of a single star in the color-absolute magnitude diagram, or on  
 1322 spectroscopic information.

1323 One major change in Table 3 relative to the earlier iteration from Bouma et al. (2019) is that the extent  
 1324 of Gaia-based analyses has now matured to the point that we can neglect pre-Gaia cluster memberships,  
 1325 except for a few cases with spectroscopically confirmed samples of age-dated stars. The membership lists for  
 1326 instance of Kharchenko et al. (2013) and Dias et al. (2014) (MWSC and DAML) are no longer required. This  
 1327 is helpful for various post-processing projects, since the field star contamination rates were typically much  
 1328 higher in these catalogs than in the newer Gaia-based catalogs.

1329 The most crucial parameters of a given star for our purposes are the Gaia DR2 source\_id, the cluster  
 1330 or group name (cluster), and the age. Given the hierarchical nature of many stellar associations, we do  
 1331 not attempt to resolve the cluster names to a single unique string. The Orion complex for instance can be  
 1332 divided into almost one hundred kinematic subgroups (Kounkel et al. 2018). (**Added: Based on Figure 1,**  
 1333 **the δ Lyr cluster may also be part of a similar hierarchical association.**) Similar complexity applies to the  
 1334 problem of determining homogeneous ages, which we do not attempt to resolve. Instead, we simply merged  
 1335 the cluster names and ages reported by various authors into a comma-separated string.

1336 This means that the age column can be null, for cases in which the original authors did not report an age,  
 1337 or for which a reference literature age was not readily available. Nonetheless, since we do prefer stars with  
 1338 known ages, we made a few additional efforts to populate this column. When available, the age provenance  
 1339 is from the original analysis of the cluster. In a few cases however we adopted other ages when string-  
 1340 based cross-matching on the cluster name was straightforward. In particular, we used the ages determined  
 1341 by Cantat-Gaudin et al. (2020) to assign ages to the clusters from Gaia Collaboration et al. (2018a), Cantat-  
 1342 Gaudin et al. (2018), Castro-Ginard et al. (2020), and Cantat-Gaudin & Anders (2020).

1343 The catalogs we included for which ages were not immediately available were those of Cotten & Song  
 1344 (2016), Oh et al. (2017), Zari et al. (2018), Gagné et al. (2018a), Gagné et al. (2018b), Gagné & Faherty  
 1345 (2018), and Ujjwal et al. (2020). While in principle the moving group members discussed by Gagné et al.  
 1346 (2018a,b); Gagné & Faherty (2018) and Ujjwal et al. (2020) have easily associated ages, our SIMBAD cross-  
 1347 match did not retain the moving group identifiers given by those studies, which should therefore be recovered  
 1348 using tools such as BANYAN Σ (Gagné et al. 2018b). We also included the SIMBAD object identifiers TT\*,  
 1349 Y\*O, Y\*?, TT?, and pMS\*. Finally, we included every star in the NASA Exoplanet Archive planetary system  
 1350 (ps) table that had a Gaia identifier available (Akeson et al. 2013). If the age had finite uncertainties, we also  
 1351 included it, since stellar ages determined through the combination of isochrone-fitting and transit-derived  
 1352 stellar densities typically have higher precision than from isochrones alone.

1353 For any of the catalogs for which Gaia DR2 identifiers were not available, we either followed the spatial  
 1354 (plus proper-motion) cross-matching procedures described in Bouma et al. (2019), or else we pulled the  
 1355 Gaia DR2 source identifiers associated with the catalog from SIMBAD. We consequently opted to drop the  
 1356 ext\_catalog\_name and dist columns maintained in Bouma et al. (2019), as these were only popu-  
 1357 lated for a small number of stars. The technical manipulations for the merging, cleaning, and joining were  
 1358 performed using pandas (McKinney 2010). The eventual cross-match (using the Gaia DR2 source\_id)  
 1359 against the Gaia DR2 archive was performed asynchronously on the Gaia archive website.

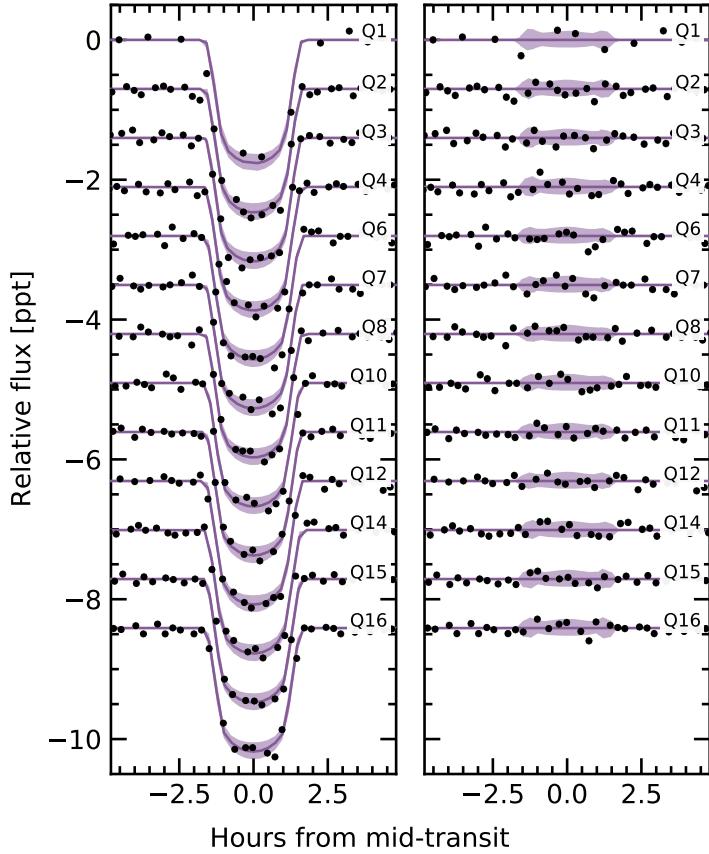
1360

## B. THE TRANSIT ASYMMETRY

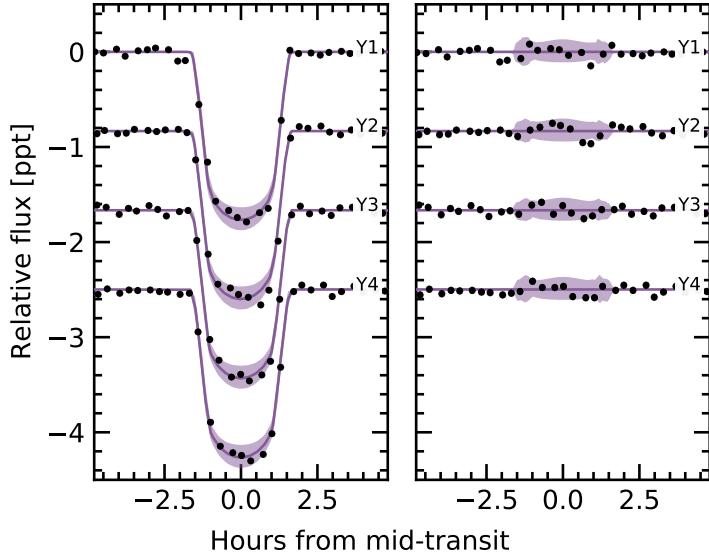
1361

## B.1. How Robust is the Asymmetric Transit?

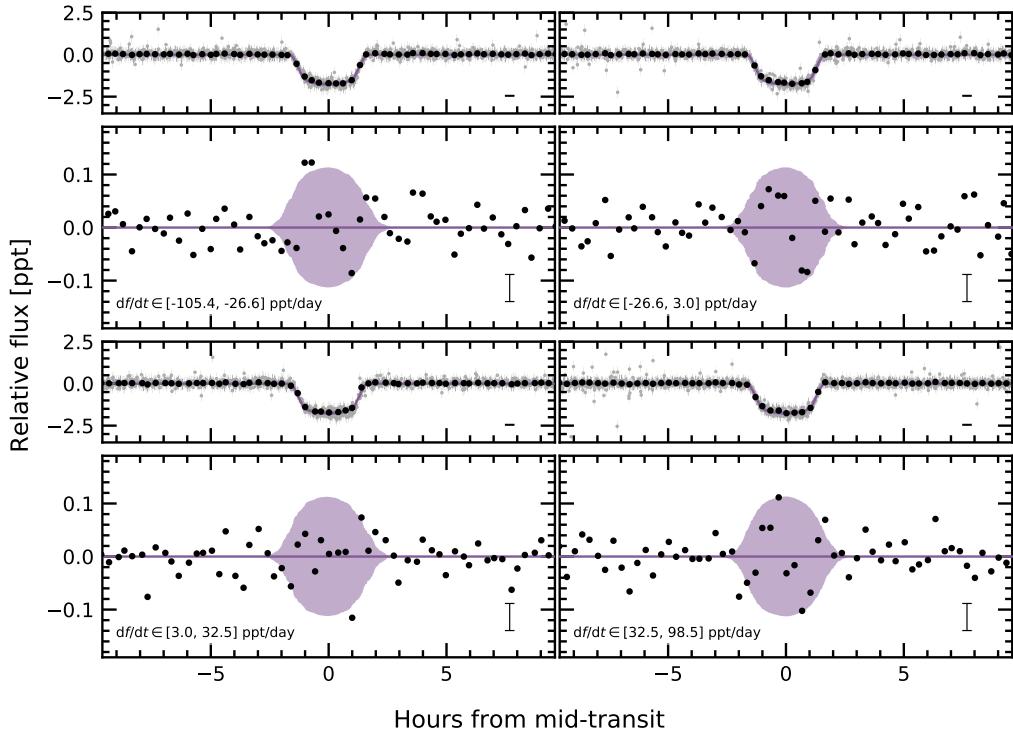
1362 As a means of exploring the robustness of the transit asymmetry, Figures 9, 10, and 11 show the Kepler data  
 1363 binned in three ways: over Kepler quarters, Julian years, and quartiles of local slope. Over Kepler quarters  
 1364 (Figure 9), Quarter 6 shows the strongest asymmetry: a deviation of about 3 ppt from expectation. Quarter



**Figure 9. Transit model residuals through time (binned by Kepler quarter).** *Left:* Phase-folded transit of Kepler 1627b, with stellar variability removed. Black points are binned to 20 minute intervals. The  $2\sigma$  model uncertainties and the maximum *a posteriori* model are shown as the faint purple band, and the dark purple line. *Right:* As on the left, with the transit removed.



**Figure 10. Transit model residuals through time (binned by year of observation).** *Left:* Phase-folded transit of Kepler 1627b, with stellar variability removed. Points and models are as in Figure 9. *Right:* As on the left, with the transit removed.



**Figure 11. Transit models and residuals, binned by quartiles in the local slope of the light curve.** Representative uncertainties for the black points (binned at 20 minute intervals) are shown in the lower right of each panel. A similar transit asymmetry to that shown in Figure 5 seems to be present in three of the four bins.

1365 7 shows an anomaly at roughly the same transit phase. Year 2 correspondingly shows the strongest anomaly  
 1366 out of any year in Figure 10; the asymmetry is visually apparent however in each of the four years.

1367 To bin by quartiles in local slope, we used our measurements of the local linear slopes in each of the  
 1368 observed transit windows (144 transits total). Four outlier transits were removed, leaving 140 transits. These  
 1369 were then divided into quartiles, so that each panel shows 35 transits binned together. The exact light curve  
 1370 slope intervals are listed in the lower left panels of Figure 11. Binned by local slope quartiles (Figure 11),  
 1371 the asymmetry is visually present in three of the four quartiles: the only bin in which it does not appear is  
 1372  $df/dt \in [3.0, 32.5] \text{ ppt day}^{-1}$ .

1373 Within the theory presented by Mazeh et al. (2015), unresolved starspot crossings cause the weak correlation  
 1374 between TTVs and the local light curve slope (Figure 6). In this model, we would expect the light curves  
 1375 with the most negative local slopes to have the largest positive TTVs, due to spot crossing events during the  
 1376 latter half of transit. The upper-left panel of Figure 11 agrees with this expectation. However, we would also  
 1377 expect the sign of the effect to reverse when considering the most positive local slopes (most negative TTVs).  
 1378 The lower-right panel of Figure 11 contradicts this expectation: the residual in both cases maintains the same  
 1379 parity! On the one hand, this shows that the residual is not dependent on the local light curve slope, which  
 1380 lowers the likelihood that it might be an artifact of our detrending methods. On the other, it raises the question  
 1381 of whether unresolved starspot crossings are indeed the root cause of the correlation shown in Figure 6. While  
 1382 we do not have a solution to this contradiction, the injection-recovery tests discussed in Section 4.1.3 provide  
 1383 some assurance that the TTV-slope correlation is not simply a systematic artifact.

1384

## B.2. Interpretation

1385 The transit asymmetry seems robust against most methods of binning the data, though with some caveats  
 1386 (e.g., the “middle quartile” in local flux,  $df/dt \in [3.0, 32.5] \text{ ppt day}^{-1}$ , where the asymmetry does not appear).  
 1387 Nonetheless, if the asymmetric were systematic we might expect its parity to reverse as a function of the sign  
 1388 of the local slope, and it does not. We therefore entertained four possible astrophysical explanations: gravity  
 1389 darkening, transit timing variations, spot-crossing events, and a persistent asymmetric dusty outflow.

1390 Gravity darkening is based on the premise that the rapidly rotating star is oblate, and brighter near the  
 1391 poles than the equator (e.g., Masuda 2015). The fractional transit shape change due to gravity darkening  
 1392 is on the order of  $(P_{\text{break}}/P_{\text{rot}})^2$ , for  $P_{\text{break}}$  the break-up rotation period, and  $P_{\text{rot}}$  the rotation period. Using

1393 the parameters from Table 2, this yields an expected 0.14% distortion of the  $\approx$ 1.8 ppt transit depth: *i.e.*, an  
 1394 absolute deviation of  $\approx$ 2.5 ppm. The observed residual has a semi-amplitude of  $\approx$ 50 ppm. Since the expected  
 1395 signal is smaller than the observed anomaly by over an order of magnitude, gravity darkening seems to be an  
 1396 unlikely explanation.

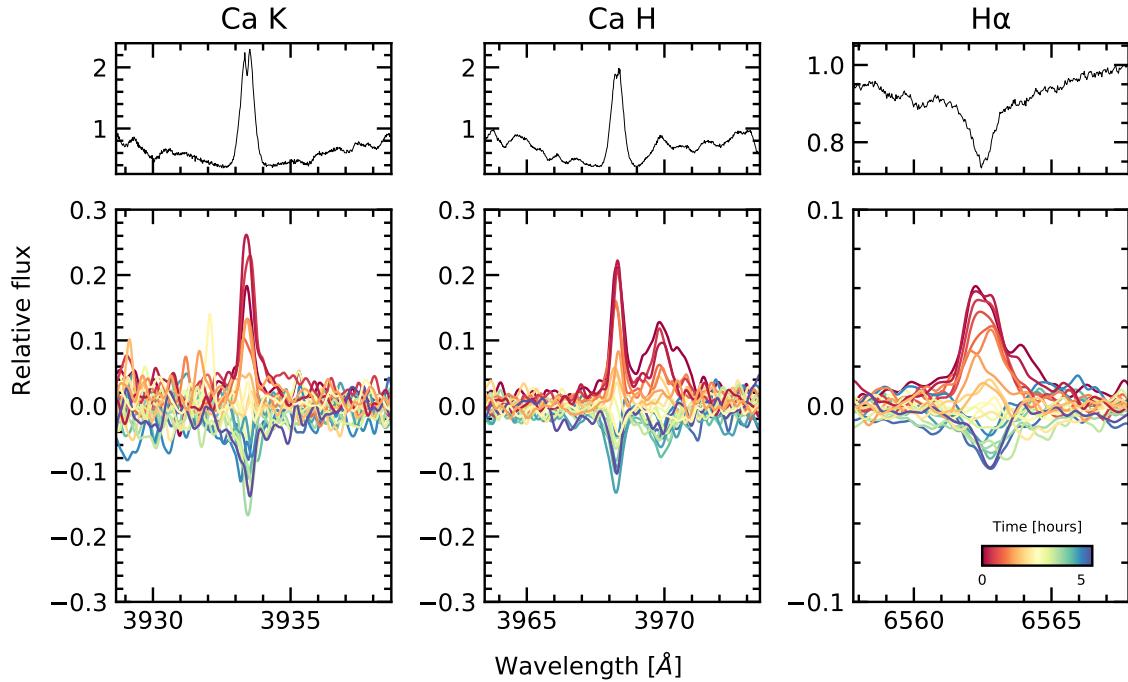
1397 The scenario of transit timing variations (TTVs) producing the asymmetry seems unlikely because the transit  
 1398 timing variations we do observe are correlated with the local light curve slope, which increases roughly as  
 1399 much as it decreases. From our analysis, the mean TTV and its standard deviation are  $0.66 \pm 5.53$  min;  
 1400 similarly the mean local slope and its standard deviation are  $0.59 \pm 45.50$  ppt day $^{-1}$ . There is therefore little  
 1401 expectation for TTVs to produce the asymmetry. A separate line of argument comes from Figure 11. If the  
 1402 local slope were essential to producing the transit asymmetry, we would expect that in the largest  $df/dt$  bin,  
 1403  $df/dt \in [3.0, 32.5]$  ppt day $^{-1}$ , the sign of the asymmetry would reverse. We do not see evidence for this being  
 1404 the case.

1405 The third and related possibility is that of starspot crossings. Young stars have higher spot-covering fractions  
 1406 than old stars (*e.g.*, Morris 2020). Young solar-type stars may also host dark starspots at high stellar  
 1407 latitudes (*e.g.*, EK Dra; Strassmeier 2009), though interferometric imaging of spotted giant stars has shown  
 1408 different starspot latitude distributions than those inferred from Doppler imaging (Roettenbacher et al. 2017).  
 1409 Regardless, for any spot-crossing anomalies to add coherently over the 140 Kepler transits, it seems likely  
 1410 that we would need either for spots to be persistent at a particular latitude (and for the planetary orbit to be  
 1411 somewhat misaligned), or for a “stroboscopic” longitudinal phasing (*e.g.*, Dai et al. 2018). For our system,  
 1412  $P_{\text{orb}}/P_{\text{rot}} \approx 2.76$ , which means that every 4 transits and 11 stellar rotations, the planet crosses over roughly the  
 1413 same stellar longitude, which might enable the necessary phasing if the spot-groups are large and long-lived.  
 1414 Unfortunately, the S/N per Kepler transit is  $\approx 8$ , which renders individual spot-crossing events unresolved.  
 1415 This explanation seems marginally plausible, mainly because the expected spot-crossing anomaly amplitudes  
 1416 ( $\approx 100$  ppm) resemble the observed amplitude of the asymmetry ( $\approx 50$  ppm). One issue with this explanation  
 1417 however is that there is no reason to expect starspot crossing events to last exactly half the transit duration.

1418 A persistent feature of the planet itself might therefore be needed to explain the transit asymmetry. An  
 1419 asymmetric outflow from the planet’s atmosphere could at least geometrically meet the requirements (*e.g.*,  
 1420 McCann et al. 2019). To explain the asymmetric transit, a small, dense component would lead the planet,  
 1421 and a long, more rarefied (and variable) component would trail it. This might also explain the slight flux  
 1422 decrement visible for  $\sim$ 1 hour pre-ingress (Figure 5). The amplitude of the asymmetry is roughly in line  
 1423 with theoretical expectations for dusty outflows (Wang & Dai 2019), and based on the planet’s size, its mass  
 1424 is likely in a regime where such outflows are possible. Out of the four explanations discussed, this one at  
 1425 least theoretically seems the most plausible. By composition, the expectation would be that the envelope  
 1426 is mostly hydrogen and helium gas, with a dust or haze component providing the broadband opacity in the  
 1427 Kepler bandpass. A natural path for testing this idea would be to observe additional transits of the planet in  
 1428 hydrogen absorption, metastable helium absorption, or across a broad wavelength range in the near-infrared.

### 1429 C. SPECTROSCOPY AND PHOTOMETRY DURING THE TRANSIT OF 2021 AUG 7

1430 (Added: We used the ephemeris of Holczer et al. (2016) to observe a transit of Kepler 1627Ab on the  
 1431 night of 2021 Aug 7 both spectroscopically and photometrically. We used the HIRES echelle spectro-  
 1432 graph at the Keck-I telescope and the MuSCAT3 photometer at Haleakalā Observatory on Maui, HI  
 1433 (Narita et al. 2020).) (Deleted: -We monitored Kepler 1627 with Keek/HIRES before, during, and after  
 1434 transit on the night of 2021 Aug 7. ) (Replaced: We used the iodine cell for wavelength calibration  
 1435 replaced with: For the HIRES wavelength calibration, we used the iodine cell), and extracted the 1-D  
 1436 spectra using the standard California Planet Survey pipeline (Howard et al. 2010). (Added: Given the faint-  
 1437 ness of the target ( $V = 13.1$ ), we observed using the C2 decker, which yielded an instrument resolution  
 1438 of  $\approx$ 50,000.) The airmass ranged between 1.1 and 2.2 from the start through the end of observations; the seeing  
 1439 ranged from  $1.^{\prime\prime}1$  at the beginning to  $1.^{\prime\prime}5$  at the end. (Added: The HIRES exposure time was set at  $\approx$ 15  
 1440 minutes in order to resolve the 2.8 hour transit event, which yielded a S/N per resolution element of  
 1441  $\approx 75$  (low for precision radial velocity standards).) (Replaced: We also simultaneously observed across  
 1442 griz bands using MuSCAT3 at Haleakalā Observatory on Maui, HI. replaced with: For the MuSCAT3  
 1443 observations, we observed simultaneously across the griz bands.) (Added: The exposure times in each  
 1444 bandpass ranged between 23 and 46 seconds, and were chosen in order to yield a S/N in the peak pixel  
 1445 that exceeded 130 while also preventing saturation.) Performing aperture photometry on the latter image  
 1446 stack yielded the data given in Table 5.



**Figure 12. Spectroscopic activity indicators during the transit of 2021 Aug 7.** The *top panels* show the median line profiles Ca K, Ca H, and H $\alpha$  line profiles from the HIRES spectra. The *lower panels* show the differences of each individual spectrum relative to the median spectrum. The bump in the red wing of Ca H is He. The spectra in the lower panels are smoothed for visualization purposes.

We considered two approaches to measuring the velocities: in the first, hereafter “Method 1”, we cross-correlated against a template found via spectral classification with SpecMatch-Emp (Yee et al. 2017). In “Method 2”, we used a high S/N template of V1298 Tau. Although V1298 Tau is cooler than Kepler 1627A by  $\approx 500$  K, it has a comparable amount of line-broadening ( $v \sin i = 23$  km s $^{-1}$ ), and a comparable level of stellar activity. The mean and standard deviation of the internal RV uncertainties averaged over all epochs were  $16.2 \pm 1.1$  m s $^{-1}$  from Method 1, and  $12.6 \pm 0.6$  m s $^{-1}$  from Method 2. The corresponding time-averaged reduced  $\chi^2$  from the template match was  $1.57 \pm 0.04$  (Method 1) and  $1.30 \pm 0.02$  (Method 2). Given these diagnostics, we adopted the velocities from the second approach, which are reported in Table 6.

Figure 7 shows the results. The MuSCAT3 photometry shows the expected starspot trend, along with the transit and what is likely a chromatic starspot crossing event at JD – 2459433 = 0.955. The radial velocities decrease by  $\approx 250$  m s $^{-1}$  over the six hour window. This decrease in RV is correlated with a decrease in the S-indices derived from the Ca HK lines. One outlying RV point is apparent shortly before egress; it is temporally coincident with an outlying value in the S-index time series.

Overall, we expect the dominant trends in both the photometry and radial velocities to be caused by starspots on the stellar photosphere rotating into and out of view. The plasma in the leading and receding limbs of the stellar disk has an apparent line-of-sight velocity of  $\pm 20$  km s $^{-1}$ . Over 10% of a rotation cycle ( $P_{\text{rot}} = 2.6$  days), spots near these limbs come into and out of view, modulate the stellar velocity profile, and can thereby produce the overall  $\approx 250$  m s $^{-1}$  trend. The Ca HK and H $\alpha$  emission profiles support this interpretation; Figure 12 shows that each line gradually decreases in intensity over the course of the six hour sequence.

The expectation however is for the starspot-induced signals to be smooth, at worst with contributions at  $0.5 P_{\text{rot}}$  or  $0.25 P_{\text{rot}}$  (Klein & Donati 2020). We therefore fitted the RVs using the Hirano et al. (2010, 2011) models for the Rossiter-McLaughlin (RM) effect, and allowed for an optional linear and quadratic trend in time to fit the  $\approx 250$  m s $^{-1}$  spot-induced trend. We followed the methodology developed by Stefansson et al. (2020). We allowed the sky-projected obliquity, the projected stellar equatorial velocity, and the Gaussian dispersion of the spectral lines to vary, and fixed the limb-darkening using the V-band tabulation from Claret & Bloemen (2011). We assumed a Gaussian prior on  $v \sin i$  and  $a/R_{\star}$  from Table 1, and also allowed for a white-noise jitter term to be added in quadrature to the measurement uncertainties. We used a 15 minute exposure time to numerically evaluate the model.

1475 The quadratic model with the RM effect is shown in Figure 7; the jitter term is incorporated in the model  
 1476 uncertainties, but not the plotted measurement uncertainties. The plotted measurement uncertainties are the  
 1477 internal uncertainties on the RVs ( $\approx 13 \text{ m s}^{-1}$ ), and are dominated by the  $v \sin i$  broadening. However, between  
 1478 exposures, the RVs show significant additional scatter that is not captured by the slow quadratic trend. The  
 1479 white-noise jitter for this particular model is  $\sigma_{\text{RV}} = 27^{+6}_{-5} \text{ m s}^{-1}$ , which is comparable to the expected RM  
 1480 anomaly of  $\Delta v_{\text{RM}} \approx f_{\text{LD}} \cdot \delta \cdot v \sin i \cdot \sqrt{1 - b^2} \approx 20 \text{ m s}^{-1}$ , assuming a perfectly aligned orbit.

The presence of this additional scatter prevents a convincing detection of the RM effect. The reason can be understood via model comparison. If we compare the model with a quadratic trend and the RM effect against a model with a linear trend and the RM effect, or even a model with no RM effect at all, then the respective Bayesian Information Criterion (BIC) values are as follows.

$$\begin{aligned} \text{BIC} &= 227.1 \quad (\text{Quadratic + RM}) \\ \text{BIC} &= 231.1 \quad (\text{Linear + RM}) \\ \text{BIC} &= 221.4 \quad (\text{Only Quadratic}). \end{aligned} \quad (\text{C1})$$

1481 There is therefore no evidence to prefer the model with the RM effect against a model that only accounts  
 1482 for the stellar variability. The “only quadratic” model does particularly well because it can inflate the jitter  
 1483 term to account for scatter during the transit (even if the scatter contains astrophysics!), and it has fewer free  
 1484 parameters. However, we cannot justify a physical prior on the jitter term, because we do not understand  
 1485 the origin of the exposure-to-exposure scatter. As noted above, the velocity deviations from starspots are  
 1486 expected to have contributions at the stellar rotation frequency, or harmonics thereof. This jitter is present on  
 1487 the exposure timescale (15 minutes), which is only 0.4% of the stellar rotation period; it is not obvious that  
 1488 starspots would be the culprit.

1489 The amplitude of both the spot-induced trend and the jitter are somewhat larger than recent comparable  
 1490 measurements in systems such as AU Mic (Palle et al. 2020), DS Tuc (Montet et al. 2020; Zhou et al. 2020)  
 1491 and TOI 942 (Wirth et al. 2021). One possible explanation for the jitter is that it is astrophysical in origin,  
 1492 and that it is caused by some novel process operating on the surface of Kepler 1627A. Another possibility  
 1493 is that our RV analysis underestimates our measurement uncertainties; in order to achieve the requisite time-  
 1494 sampling the S/N per resolution element in our spectra was 70 to 80, which is lower than desired for deriving  
 1495 high-precision velocities. In addition, the rapid rotation of the star could affect accuracy of the uncertainties  
 1496 from the velocity extraction. Observations at higher S/N are necessary to distinguish these two possibilities,  
 1497 and remain worthwhile in order to clarify the orbital geometry of Kepler 1627Ab. Useful next steps would  
 1498 include transit observations with a stabilized spectrograph in the optical (e.g., Gibson et al. 2016; Seifahrt  
 1499 et al. 2018), or in the near-infrared (e.g., Feinstein et al. 2021).

**Table 5.** MuSCAT3 photometry of Kepler 1627.

Time [BJD <sub>TDB</sub> ]	Rel. Flux	Rel. Flux Err.	Bandpass
2459433.829202	0.99719	0.00091	g
2459433.829324	0.99849	0.00112	r
2459433.829117	0.99611	0.00116	i
2459433.829406	0.99941	0.00136	z

NOTE— Table 5 is published in its entirety in a machine-readable format. Example entries are shown for guidance regarding form and content.

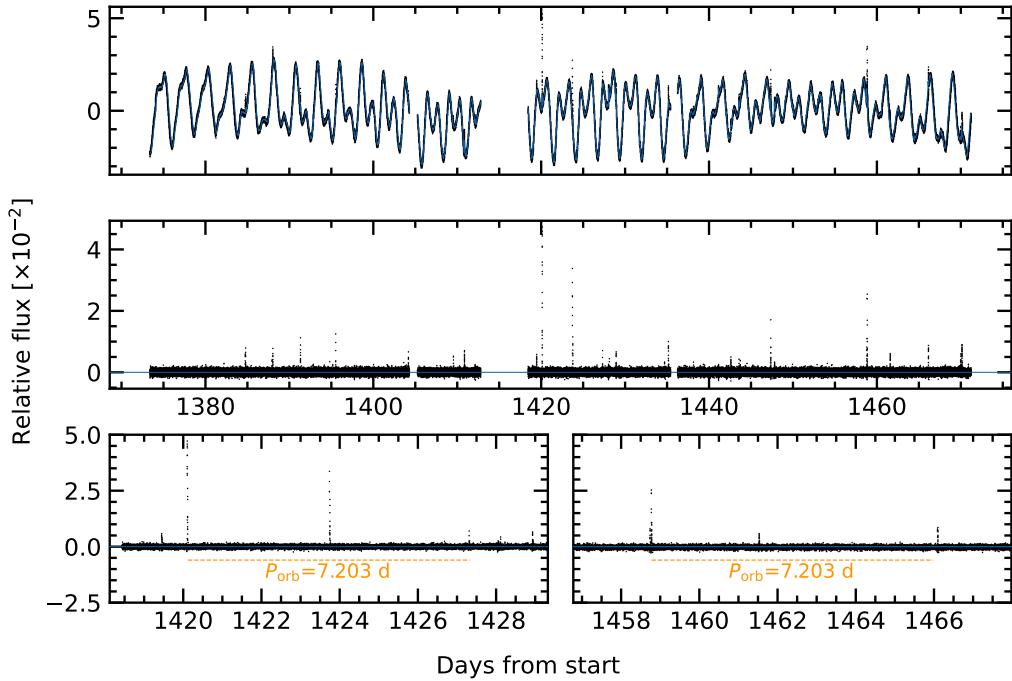
**Table 6.** Kepler 1627 radial velocities.

Time [BJD <sub>TDB</sub> ]	RV [ $\text{m s}^{-1}$ ]	$\sigma_{\text{RV}}$ [ $\text{m s}^{-1}$ ]	S-value
2459433.801306	152.97	12.29	0.7355
2459433.812255	100.5	13.23	0.7434

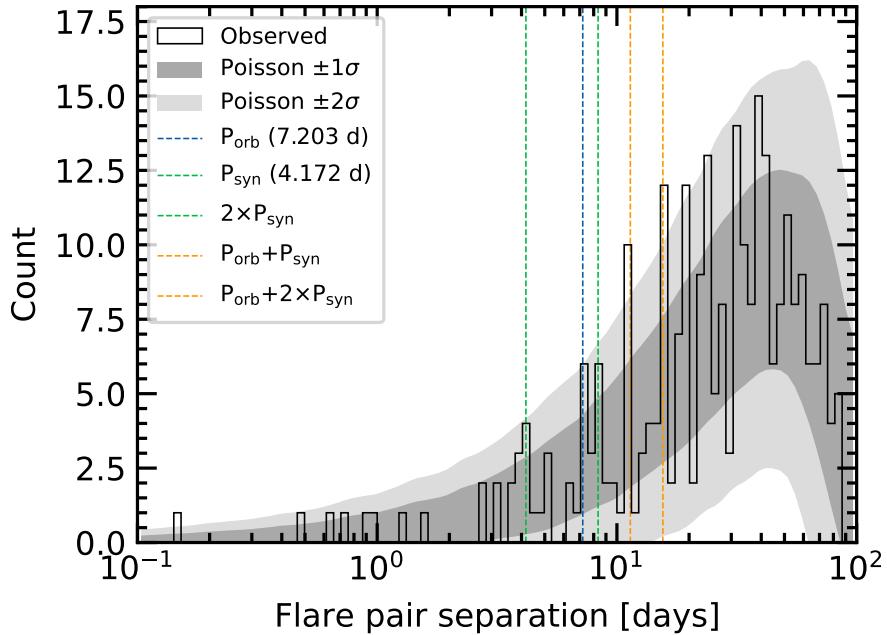
NOTE— Table 6 is published in its entirety in a machine-readable format. Example entries are shown for guidance regarding form and content.

## D. FLARE ANALYSIS

1500 In addition to the 3.9 years of long cadence data,  
 1501 short cadence (1-minute) Kepler observations were  
 1502 acquired over 97.7 days during Quarter 15. The



**Figure 13. Flares in Kepler 1627.** *Top:* The 1-minute cadence Kepler data (black points) is shown with a stellar variability model superposed (blue line). *Middle:* Residual after subtracting the stellar variability model. Flares appear as spikes. *Bottom:* Zooms of the brightest, and third-brightest flares. A timing coincidence – that both flares have “successors” approximately one (Added: planetary) orbital period after the initial event – is emphasized.



**Figure 14. Statistics of inter-flare arrival times.** 24 flares were recorded with amplitudes exceeding 0.5% over the 97.7 days of short cadence observations. The histogram of the time intervals between every possible pair of flares is shown in black. Some plausibly important timescales for star-planet interactions, namely the planetary orbital period and synodic period (the orbital period as seen from the rotating stellar frame) are shown along with their linear combinations. Monte Carlo draws from a Poisson distribution are shown with the gray bands. While peaks in the observed distribution do coincide with some of the “special periods”, the statistical evidence for a non-Poissonian process driving the flares does not clear the  $5\sigma$  threshold.

short cadence light curve shows a higher rate of flaring than visible in the long cadence data (Figure 13). We analyzed the short cadence light curve and its flares according to the following procedure.

1. Fit the starspot-induced variability using a Gaussian Process with a SHOTerm kernel, a white-noise jitter term, and the mean flux.
2. Select points more than twice the median absolute deviation from the residual, and exclude them from the light curve (these points include the flares). Repeat Step 1.
3. Using the residual from Step 2, identify all flares, requiring them to be at least 20 cadences apart, at least 7 median absolute deviations above the median baseline, and lasting at least 2 cadences in duration. Build the mask spanning these times, from 5 minutes before each flare begins to 2.5 minutes after the final flare cadence. Repeat Step 1 a final time.

The final step of flare identification and fitting was performed using `altaipony` (Davenport 2016; Ilin et al. 2021). The analytic flare model is from Davenport et al. (2014) and it parametrizes the flare with a start time, an exponential lag time, and an amplitude.

There were  $N_f = 24$  flares that exceeded 0.5% in relative flux during the short cadence observations. These 24 flares spanned a total of 6.5 hours ( $\sim 15$  minutes per flare). Inspecting the data, we noticed a coincidence in the flare arrival times. The coincidence is that despite the low flare duty cycle, one orbital period after the brightest flare, a second flare followed. This and a similar event are shown in Figure 13. The timing error is good to a  $\approx 0.2\%$  difference from the orbital period, which given the duty cycle seems *a priori* unlikely. If we consider flares falling within 2% of the planet’s orbital period after a previous flare, then 4 of the 24 flare events have candidate “successors”.

As with any coincidence, if one does not have a firm prediction, it is difficult to assess whether a surprise is statistically significant. Since our surprise was specifically at the inter-arrival time of certain flares coinciding with special time intervals, we performed the following analysis. First, we considered all unordered pairs of flares. For  $N$  flares there are  $\binom{n}{2}$  such pairs (for our case, 276 pairs). We then compared the distribution of the pair separations against that of a Poisson distribution. Specifically, we drew  $N_f = 24$  samples from a Poisson distribution with  $\lambda = \Delta t / N_f$ , for  $\Delta t = 97.7$  days the total duration of the observations, and repeated the draw  $10^3$  times with unique random seeds.

Figure 14 shows the results. The vertical lines in the figure show the planetary orbital period, the synodic period  $P_{\text{syn}} = (P_{\text{rot}}^{-1} - P_{\text{orb}}^{-1})^{-1}$ , and linear combinations thereof. The tidal period (half the synodic period) is not shown. The bins are logarithmically spaced to give 100 bins between the minimum and maximum ordinate values. The gray bands express the range of values observed from the Poissonian draws. While it does seem like an odd coincidence for peaks in the observed flare arrival time distribution to coincide with the locations of these “special intervals”, the statistical evidence for a non-Poissonian process driving the flares does not seem especially overwhelming. More quantitatively, the peaks observed at the orbital and synodic periods are within the  $\pm 2\sigma$  range of a Poissonian process, and those at  $P_{\text{orb}} + P_{\text{syn}}$  and  $P_{\text{orb}} + 2P_{\text{syn}}$  are only slightly above this range. With that said, future analyses of these data by investigators with more knowledge of this topic could very well yield more quantitative insights. Such analyses should keep in mind an important caveat: the amplitude distribution of M-dwarf flares extends up to many times the quiescent flux (see Figure 7 of Günther et al. 2020). A flare on Kepler 1627B producing double its quiescent white-light flux would yield a  $\approx 1\%$  apparent amplitude. Such flares could represent a significant fraction of those in the Kepler observations.