

## An Adolescent Mini-Neptune in the Kepler Field

L. G. BOUMA,<sup>1</sup> J. L. CURTIS,<sup>2,3</sup> AND K. MASUDA<sup>4</sup>

<sup>1</sup>*Department of Astrophysical Sciences, Princeton University, 4 Ivy Lane, Princeton, NJ 08540, USA*

<sup>2</sup>*Department of Astronomy, Columbia University, 550 West 120th Street, New York, NY 10027, USA*

<sup>3</sup>*Department of Astrophysics, American Museum of Natural History, New York, NY 10024, USA*

<sup>4</sup>*Department of Earth and Space Science, Osaka University, Osaka 560-0043, Japan*

(Received —; Revised —; Accepted —)

Submitted to AAS Journals

### ABSTRACT

Kepler 1627A is a G8V star previously known to host a  $3.6 R_{\oplus}$  mini-Neptune on a 7.2 day orbit. The star was observed by the Kepler satellite because it is nearby ( $d \approx 333$  pc) and it resembles the Sun. Combining Gaia kinematics with rotation periods from TESS, we show that Kepler 1627 is a member of the  $35^{+6}_{-5}$  Myr old  $\delta$  Lyr cluster. To our knowledge, Kepler 1627Ab is therefore the youngest planet with a precise age yet found by the main Kepler mission. The Kepler photometry seems to show two anomalies. First, unresolved starspot-crossing events distort the average transit shape, and induce transit timing variations that imply the planet is likely on a prograde orbit. Second, although the flare arrival time distribution is probably Poissonian, there may be an excess of flares separated by the orbital and synodic periods, and linear combinations thereof. The  $\delta$  Lyr cluster is one of many stellar groups whose properties have been clarified using the Gaia data; many of the other new clusters also overlap with confirmed and candidate exoplanets. Further studies that combine Gaia kinematics and TESS rotation periods are well-poised to expand the census of age-dated planets, and our understanding of planetary evolution.

**Keywords:** exoplanet evolution (491), stellar associations (1582), 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 molten rocky cores (Kite et al. 2020) and extended atmospheric envelopes of hydrogen and helium, are expected to shrink in size by factors of several over their first  $10^8$  years (Owen & Wu 2016; Owen 2020). Specifically, they start with sizes of  $4\text{--}12 R_{\oplus}$  near the time of disk dispersal ( $\lesssim 10^7$  years), and shrink to sizes of  $2\text{--}4 R_{\oplus}$  by  $10^8$

years. This change is driven by a combination of stellar irradiation and internal heat that powers an outflow, which can eventually deplete or entirely strip the envelope from the rocky core (Owen & Wu 2013; Ginzburg et al. 2018). 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).

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). The Kepler mission however has not yielded any planets with precise ages below one gigayear (Meibom et al. 2013). The reason is that at the time of the main Kepler mission (2009–2013), only four open clusters were known in the Kepler field: NGC 6866,

Corresponding author: L. G. Bouma  
bouma.luke@gmail.com

NGC 6811, NGC 6819, and NGC 6791, with ages spanning 0.7 Gyr to 9 Gyr (Meibom et al. 2011). Since that time, analyses of the kinematic, photometric, and astrometric Gaia data have expanded our knowledge of open cluster and moving group 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).

Comparing our young star list against the Kepler field yielded two discoveries. The first, to be discussed in an upcoming analysis by J. Curtis et al., is that Kepler observed the  $\approx 350$  Myr open cluster Theia 520 (UBC 1). At least six Kepler planets are known in the cluster from the Kepler-52 and Kepler-968 systems (Rowe et al. 2014; Jontof-Hutter et al. 2021). The second discovery, and the focus of this work, is that Kepler observed a portion of the  $\delta$  Lyr cluster (Stephenson-1; Theia 73; Stephenson 1959). Figure 1 shows the overlap between the Kepler field and the cluster.

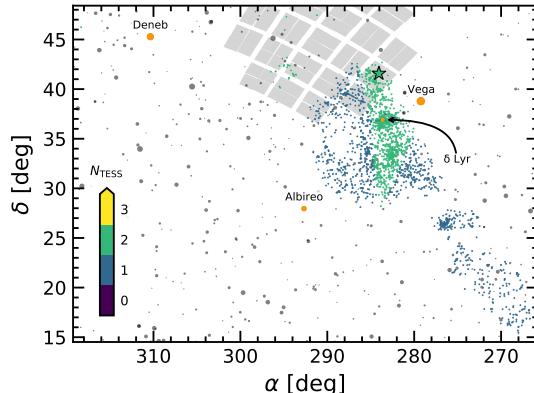
A previous clustering analysis of the Gaia data reported that Kepler 1627 (KIC 6184894; KOI 5245; TIC 120105470) is a member of the  $\delta$  Lyr cluster (Kounkel & Covey 2019). Given the statistical validation of its close-in mini-Neptune (Tenenbaum et al. 2012; Morton et al. 2016; Thompson et al. 2018), we examined the age of the star and its putative host cluster most closely (Section 2). We find that the  $\delta$  Lyr cluster is  $35^{+6}_{-5}$  Myr old, and that Kepler 1627 is indeed a member of the cluster. Focusing on the planet itself (Section 3), we confirm that despite a previously unreported binary companion, it is indeed a mini-Neptune orbiting the G-dwarf primary star. The presence of a correlation between transit timing variations and the local light curve slope renders any other explanation implausible. We conclude by highlighting some broader implications for our ability to age-date a statistical sample of planets (Section 4).

## 2. THE CLUSTER

To assess the age of the cluster and our star of interest, we used a combination of the isochronal, gyrochronal, and lithium-based age-dating.

### 2.1. Color-Magnitude Diagram

We measured the isochrone age using an empirical approach. The left panel of Figure 2 shows the color-magnitude diagram (CMD) of stars in the  $\delta$  Lyr cluster, IC 2602, the Pleiades, and the field. The stars shown in the  $\delta$  Lyr cluster are selected based on their positions and kinematics, as discussed in Appendix B. Those from the Pleiades

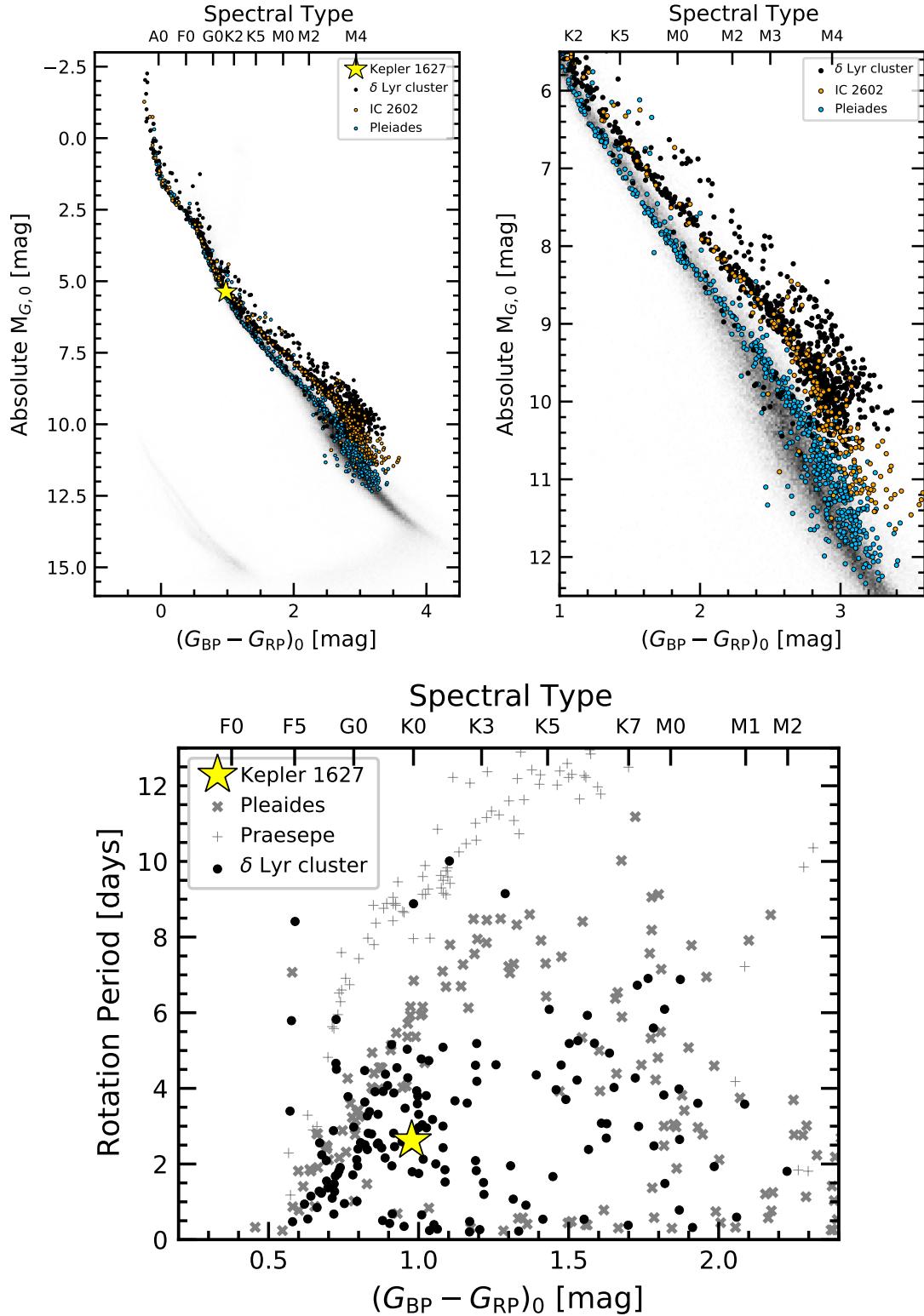


**Figure 1. Kepler and TESS views of the  $\delta$  Lyr cluster.** Colored points are kinematically selected members of the  $\delta$  Lyr cluster. Both Kepler (gray squares) and TESS observed portions of the cluster. Naked-eye stars ( $m_V < 6.5$ ) are shown in gray; four of them (orange points) have their names annotated. Kepler 1627 (green star) was observed during the entirety of the Kepler mission, and has been observed by TESS for two lunar months to date. The colors of the other cluster members indicate the number of TESS sectors during which they have been observed.

and IC 2602 are adopted from Cantat-Gaudin et al. (2018), and the field stars are from the Gaia EDR3 Catalog of Nearby Stars (Gaia Collaboration et al. 2021). We cleaned these following identical data filtering criteria discussed in Appendix B of Gaia Collaboration et al. (2018a), except that we weakened the parallax precision requirement to  $\varpi/\sigma_\varpi > 5$ . These filters were designed to include genuine binaries while omitting instrumental artifacts. We then corrected for extinction by querying the 3-dimensional maps of Capitanio et al. (2017) and Lallement et al. (2018)<sup>1</sup>, and by then applying the extinction coefficients  $k_X \equiv A_X/A_0$  computed by Gaia Collaboration et al. (2018a) while assuming that  $A_0 = 3.1E(B-V)$ . For IC 2602, the Pleiades, and the  $\delta$  Lyr cluster, this procedure yielded a respective mean and standard deviation for the reddening of  $E(B-V) = \{0.020 \pm 0.003, 0.045 \pm 0.008, 0.032 \pm 0.006\}$ . These values agree reasonably well with previously reported values from the literature (e.g., Gaia Collaboration et al. 2018a; Kounkel & Covey 2019; Bossini et al. 2019).

The immediate impression from Figure 2 is that the  $\delta$  Lyr cluster and IC 2602 overlap, and therefore are approximately the same isochronal age. In our exploration, we also compared against  $\mu$ -Tau ( $\approx 60$  Myr; Gagné et al. 2020) and the Upper-Centaurus-Lupus component of the Sco OB2 asso-

<sup>1</sup> <https://stilism.obspm.fr/>



**Figure 2. The  $\delta$  Lyr cluster is  $35 \pm 15$  Myr old.** Top: HR diagram of the  $\delta$  Lyr cluster, IC 2602 ( $\approx 35$  Myr), the Pleiades ( $\approx 115$  Myr), and the field. The right panel provides a closer view of stars on the pre-main-sequence. The  $\delta$  Lyr cluster and IC 2602 are approximately the same isochronal age. Bottom: TESS and Kepler rotation periods and dereddened Gaia colors, with the Pleiades (125 Myr) and Praesepe (650 Myr) shown for reference (Rebull et al. 2016; Douglas et al. 2017). Kepler 1627 is shown with a yellow star. **todo: caption improve..**

ciation (UCL;  $\approx 10\text{--}15$  Myr; Damiani et al. 2019). The pre-main-sequence M dwarfs of the  $\delta$  Lyr cluster were intermediate between these latter two clusters. To quantify this into an age measurement with an uncertainty, we adopted the empirical procedure developed by Gagné et al. (2020) in their Section 6. 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, we adopted members of UCL, IC 2602, and the Pleiades from Damiani et al. (2019) and Cantat-Gaudin et al. (2018) respectively. 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 by manually excluding stars that were obvious photometric binaries in the CMD. We then passed a moving box average and standard deviation across the CMD in 0.10 mag bins, fitted a univariate spline to the binned values, and assembled a piecewise grid of hybrid isochrones spanning the ages between UCL to the Pleiades following Equations 6 and 7 from Gagné et al. (2020).

The ages returned by this procedure depend on the ages assumed for each reference cluster. In our case, the IC 2602 cluster is most important. As noted by Bouma et al. (2020), plausible ages for IC 2602 span 30 Myr to 46 Myr, with slightly older ages being preferred by the lithium-depletion-boundary (LDB) measurements and younger ages by the main-sequence turn-off (Stauffer et al. 1997; Dobbie et al. 2010; David & Hillenbrand 2015; Randich et al. 2018; 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 case of 46 Myr, the  $\delta$  Lyr cluster would be  $44^{+8}_{-7}$  Myr old. We adopt an intermediate 35 Myr age as our reference, yielding an age for the  $\delta$  Lyr cluster of  $35^{+6}_{-5}$  Myr, though it should be understood based on the systematic uncertainties that an age of up to  $\approx 50$  Myr remains plausible. Detailed analyses of the LDB or main-sequence turn-off in the  $\delta$  Lyr cluster could help clarify the issue, but both are beyond our scope.

## 2.2. Stellar Rotation Periods

Of the 3,071 candidate  $\delta$  Lyr cluster members reported by Kounkel & Covey (2019), TESS data were available for 924 stars. We extracted light curves using the nearest pixel to each star, and regressed them against systematics using 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.<sup>0</sup> radius of the target star (1 TESS pixel). We recorded the number of stars within the aperture with greater brightness than the target star, and with brightness within 1.25 and 2.5 TESS magnitudes of the target star. **We then supplemented these measurements with XXX Kepler rotation periods (CITEP MCQUILLAN2014).**

We cleaned the candidate TESS rotation period measurements through a combination of automated and manual steps. As a matter of scope, we restricted our attention to the 391 stars in the spatial and kinematic proximity of Kepler 1627, as discussed in Appendix B. We also only focused on stars for which no companions were known with a brightness exceeding one-tenth of the target star within a 21.<sup>0</sup> radius: there were 192 stars that met these requirements, and that had TESS data available. For plotting purposes we then imposed a selection based on the strength of the rotation period signal itself: we required the Lomb Scargle power to exceed 0.1, and the period to be below 15 days.

The lower panel of Figure 2 shows the resulting 157 stars. The majority of these stars fall below the “slow sequence” of the Pleiades, consistent with a gyrochronological age for the  $\delta$  Lyr cluster below 100 Myr. Approximately 10 stars appear as outliers above the “slow sequence”. Assuming that they are all false positives (*i.e.*, field interlopers), the rotation period detection fraction would be  $147/192 \approx 77\%$ .

Binarity is known to be an important factor in the density of stars falling underneath the “slow sequence” (Meibom et al. 2007; Gillen et al. 2020; Bouma et al. 2021). Appendix C discusses this topic, and finds evidence for similar effects to those that have been previously discussed in the literature.

## 2.3. Age of Kepler 1627

To summarize, the age of the  $\delta$  Lyr cluster cluster based on the empirical fitting procedure described in Section 2.1 is  $35^{+6}_{-5}$  Myr. This procedure hinges on the age adopted for IC 2602, since it is the closest reference cluster in the color-magnitude diagram. Gyrochronologically, while most of the G and K dwarfs in the  $\delta$  Lyr cluster have not yet converged to form a slow sequence, their rotation distribution and that of the F dwarfs is consistent with the same age. The rotation periods of other 30 Myr to 50 Myr clusters (*e.g.*, IC 2602, Collinder 135) are indistinguishable (S. Douglas, priv. comm.). The most precise age for Kepler 1627 and its planet are therefore also  $35^{+6}_{-5}$  Myr, based on the star’s spatial, kinematic, and rotational connection to this larger ensemble, and the assumption that Kepler 1627Ab formed very quickly after the star itself.

Beyond the kinematic and rotational lines of evidence for the youth of Kepler 1627, there is a final consistency check: the abundance of Li I as inferred from the 6708 Å doublet. This connection of this element to age measurements for young stars has been reviewed by Soderblom et al. (2014). Following the equivalent width measurement procedure described by Bouma et al. (2021), we find  $\text{EW}_{\text{Li}} = 232^{+23}_{-13} \text{ mÅ}$ , which does not correct for the Fe I blend at 6707.44 Å. Given  $T_{\text{eff}} \approx 5500 \text{ K}$ , this measurement is also in agreement with an age for the star between 30 Myr and 50 Myr (see e.g., Figure 9 of Bouma et al. 2021, or the data from Randich et al. 2018).

### 3. THE PLANET

If the Kepler 1627Ab transit signal is created by a bona fide planet, then to our knowledge its age would make it the youngest planet yet found by the main Kepler mission<sup>2</sup>. A portion of the light curve is shown with the phase-folded transit in Figure 3. The dominant signal in the PDCSAP photometry is a quasi-period starspot signal with a peak-to-peak amplitude that varies between roughly 2% and 8%, depending on the degree of asymmetry in spot coverage between the stellar hemispheres. The starspot signal repeats every  $2.642 \pm 0.042$  days, where we have adopted an uncertainty based on the scatter observed between different portions (quarters) of the time-series data. The details of how we model this starspot signal to derive the parameters of the planet are discussed in Appendix D. The results, given in Table 1, are consistent with a mini-Neptune sized planet ( $3.56 \pm 0.04 R_{\oplus}$ ) on a circular 7.2 day orbit ( $e < 0.25$ ), around a G8V star ( $0.91 \pm 0.05 R_{\odot}$ ). **TODO quote the eccentricity limit**

Could the transit signal be produced by anything other than a planet orbiting this near-solar analog? Morton et al. (2016) considered the false positive scenarios of a foreground eclipsing binary, a hierarchical triple-star binary system, and a background eclipsing binary. Based on the transit shape, Morton et al. (2016) found the latter scenario to be the only possible case, with a (model-dependent) probability of  $\approx 10^{-5}$ . While this statistical validation is reassuring, “validated” planets have been shown to be false positives (Shporer et al. 2017). Most failures of the statistical validation approach arise when attempting to separate the smallest stars from identically-sized gas giant planets. With a size of  $\approx 3.6 R_{\oplus}$ , Kepler 1627b does not meet this concern. The fact that we detect a low-mass stel-

lar companion in archival Keck-NIRC2 Kp-band ( $2.12 \mu\text{m}$ ) imaging does not change the conclusion: the companion star contributes 1% of the flux observed by Kepler. Combined with the observed transit depth and shape (Appendix F), a planet transiting the primary star is the only plausible interpretation of the data.

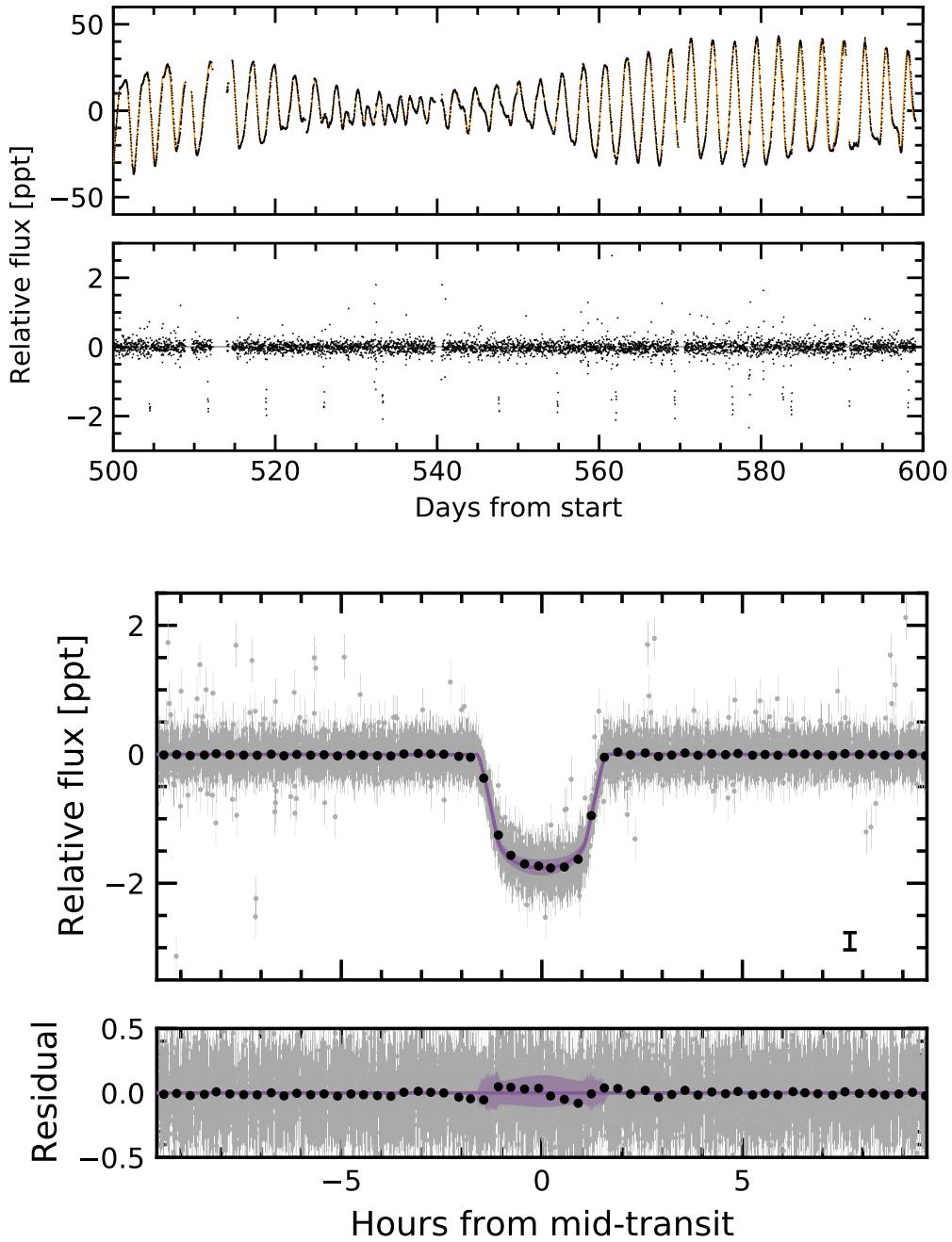
**todo: LGB to check numbers on full-dataset transit ecc and b** Two additional new lines of evidence confirm the planetary interpretation. First, the  $\approx 100$  days of 1-minute cadenced Kepler observations reduce the cadence-induced smearing, enabling a more precise measurement of the transit impact parameter (Kipping 2010). The planet transits with  $b < 0.YY$  ( $3\sigma$ ) from the 1-minute cadenced data, compared to  $< 0.XX$  ( $3\sigma$ ) from the 30-minute data.

**todo: Kento to check numbers on per-transit analysis** The second line of evidence comes from our transit timing analysis. We 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 second-order polynomial and transit, and *ii*) a local linear trend. We let the mid-time of each transit float, and then calculated the residual between the measured mid-time and that of a perfectly periodic orbit. This residual, the “transit timing variation” (TTV), is plotted in Figure 4 against the local linear slope. Only a handful of Kepler Objects of Interest have shown this correlation (Holczer et al. 2015), which is most readily interpreted as a TTV induced by unresolved starspot crossings (Mazeh et al. 2015). This is of course only possible if the planet is transiting the primary star, which excludes the any of the aforementioned false positive scenarios. It also implies that the planet’s orbit is prograde. The latter interpretation assumes that the dominant photometric variability is induced by dark spots, rather than bright faculae. Given the observed transition of Sun-like stellar variability from spot to faculae-dominated regimes, we expect this inference to be reasonably secure (Shapiro et al. 2016; Montet et al. 2017; Reinhold & Hekker 2020).

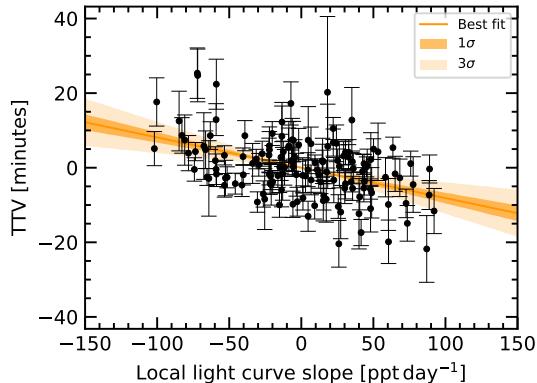
### 4. DISCUSSION & CONCLUSIONS

**LGB TODO: ask James Owen what he thinks the mass loss rate is** Kepler 1627Ab therefore provides a new extremum in the ages of the Kepler planets, and opens multiple avenues for study from the ground. Observations of its Doppler transit are expected to induce a Rossiter-McLaughlin anomaly with an amplitude of  $\approx 20 \text{ m s}^{-1}$ , and could yield a quantification of the stellar obliquity, and a means for verifying our transit-timing based suggestion of a prograde orbit. Similar transit spectroscopy aimed at detecting either hydrogen or helium absorption could yield insight into the planet’s

<sup>2</sup> The re-purposed K2 mission found two slightly younger systems: K2-33b (David et al. 2016; Mann et al. 2016) and V1298 Tau (David et al. 2019)



**Figure 3. The light curve of Kepler 1627.** *Top:* The full Kepler dataset spans 1,437 days (3.9 years), sampled at 30 minute cadence. A one hundred day segment is shown here. 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 model for the stellar variability (orange line) is subtracted below, revealing the transits of Kepler 1627Ab, as well as other deviations from the stellar variability model. The Figure Set available online shows the entire 3.9 years of available data. *Bottom:* Phase-folded transit of Kepler 1627b, with stellar variability removed. The  $2\sigma$  model uncertainties and the maximum *a posteriori* model are shown as the faint purple band, and the dark purple line. Gray points are individual PDCSAP flux measurements; black points are binned to 20 minutes, with a representative error bar shown. The asymmetric residual visible during the transit is robust against detrending methods; we believe it is caused by starspot crossings.



**Figure 4. Evidence for a prograde orbit from transit timing variations.**

**transit timing variations.** The time of each transit was measured, along with the local slope of the light curve. The two quantities appear anti-correlated, which is most easily explained 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 are most easily understood by considering that the flux from the star changes by  $\sim 60$  ppt per half rotation period ( $\sim 1.3$  days).

atmosphere, which is expected to be outflowing at a rate of  $X \cdot X M_{\oplus} \text{ yr}^{-1}$  (CITE). A more challenging measurement would be a measurement of the planet's mass, through either detailed timing analyses of the Kepler data, or through a high-cadence multi-color radial velocity campaign. The mass could yield constraints on both the planet's composition as well as its initial entropy (Owen 2020).

The precision of the Kepler data itself may enable more detailed analyses without needing to acquire more data. While our transit timing study did not yield the detection of any clearly significant planet-planet induced TTVs, the star's flaring may be suggestive of star-planet interactions. Appendix E highlights an odd flare timing coincidence that may be worthy of further examination.

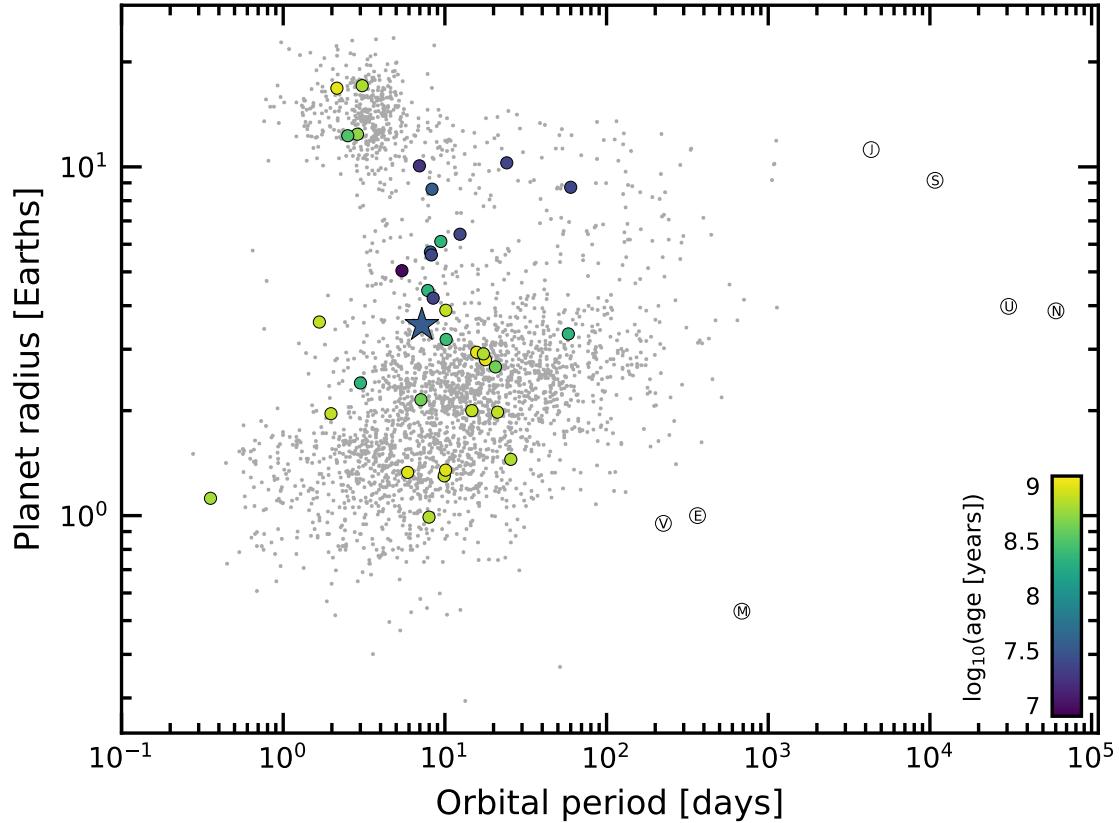
In the context of the known transiting planets, Kepler 1627Ab is among the youngest known (Figure 5). Comparable systems with precisely measured ages spanning 10 to 100 Myr include K2-33b, DS Tuc Ab, HIP 67522b, TOI 837b, AU Mic b, and the four-planet V1298 Tau system (Mann et al. 2016; David et al. 2016; Newton et al. 2019; David et al. 2019; Bouma et al. 2020; Plavchan et al. 2020). Of these systems, Kepler 1627Ab inhabits a similar range of planetary size and orbital period. It is the smallest though ( $3.6 R_{\oplus}$ ), which may hint at selection effects in the sample of young planets (Zhou et al. 2021). Assuming that these planets have masses between  $\approx 10$  and  $20 M_{\oplus}$ , their large sizes and presumed low densities are in accord with the expectation that mini-Neptunes start their lives with

large primordial atmospheres that are lost over the first gigayear (Owen & Wu 2013; Fulton et al. 2017; Ginzburg et al. 2018). The observation from starspot-induced TTVs that the planetary orbit could be aligned with the stellar spin differs from a number of other close-in mini-Neptunes (Sanchis-Ojeda & Winn 2011; Albrecht et al. 2012; Dalal et al. 2019; Rubenzahl et al. 2021). Overall however, it is consistent with a quiescent disk-driven migration history for the planet, while the in-situ formation scenario is disfavored because the necessary disk mass would always drive planets to migrate (Inamdar & Schlichting 2015; Ogihara et al. 2015).

Ultimately, our age measurement for Kepler 1627Ab was enabled by identifying its connection to the  $\delta$  Lyr cluster using the Gaia data. The Supplementary Table 2 enables similar cross-matches to already known planets, as well as forthcoming planets from the TESS mission (Ricker et al. 2015; Guerrero et al. 2021). An alternative approach would be to identify kinematic associations around exoplanet host stars using the Gaia data, and to then verify these associations with ancillary rotation periods and spectroscopy (Toflemire et al. 2021).

## ACKNOWLEDGMENTS

L.G.B. and J.L.C. are grateful to T. David for help with the transit fitting, to R. Kerr for kindly providing us with the Kerr et al. (2021) membership list prior to its publication. L.G.B. acknowledges support by the TESS GI Program, programs G011103 and G022117, through NASA grants 80NSSC19K0386 and 80NSSC19K1728. L.G.B. was also supported by a Charlotte Elizabeth Procter Fellowship from Princeton University. This study was based in part on observations at Cerro Tololo Inter-American Observatory at NSF's NOIRLab (NOIRLab Prop. ID 2020A-0146; 2020B-0029 PI: Bouma), which is managed by the Association of Universities for Research in Astronomy (AURA) under a cooperative agreement with the National Science Foundation. 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.



**Figure 5. Radii, orbital periods, and ages of known transiting exoplanets.** Planets older than  $10^9$  years are shown with gray markers. Planets younger than a gigayear, with  $t/\sigma_t > 3$ , are shown in color. Kepler 1627Ab is shown with a star. The abnormally large radii of the youngest transiting planets could be explained by their primordial atmospheres not yet having evaporated. Selection effects could also be important. Parameters were acquired from the NASA Exoplanet Archive on July 2, 2021.

*Software:* 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; Kipping 2013; Luger et al. 2019; Theano Development Team 2016), IPython (Pérez & Granger 2007), matplotlib (Hunter 2007), numpy (Walt et al. 2011), pandas (McKinney 2010), PyMC3 (Salvatier et al. 2016), scipy (Jones et al. 2001), TESS-point (Burke et al. 2020), wotan (Hipke et al. 2019).

*Facilities:* Astrometry: Gaia (Gaia Collaboration et al. 2018b, 2020). Imaging: Second Generation Digitized Sky Survey. Spectroscopy: CTIO1.5m (CHIRON; Tokovinin et al. 2013), AAT (HERMES; Lewis et al. 2002; Sheinis et al. 2015), VLT:Kueyen (FLAMES; Pasquini et al. 2002). Photometry: Kepler (Borucki et al. 2010), TESS (Ricker et al. 2015).

**Table 1.** Priors and posteriors for the transit and stellar variability model fitted to the long-cadence Kepler 1627b photometric timeseries.

Param.	Unit	Prior	Median	Mean	Std. Dev.	3%	97%	ESS	$\hat{R}-1$
<i>Sampled</i>									
$P$	d	$\mathcal{N}(7.20281; 0.01000)$	7.2028035	7.2028033	0.0000073	7.2027893	7.2028171	1910.5903732	0.0035905
$t_0^{(1)}$	d	$\mathcal{N}(120.79053; 0.02000)$	120.790504	120.790505	0.0009438	120.7886867	120.7922431	1564.1105056	0.0003213
$\log R_p/R_\star$	–	$\mathcal{U}(-4.605; 0.000)$	-3.33523	-3.33569	0.06618	-3.45772	-3.21617	1173.56574	0.00310
$b$	–	$\mathcal{U}(0; 1+R_p/R_\star)$	0.3971	0.3886	0.2070	0.0204	0.7289	378.9528	0.0177
$u_1$	–	Kipping (2013)	0.28	0.30	0.179	0.002	0.603	1161.95	0.005
$u_2$	–	Kipping (2013)	0.425	0.381	0.314	-0.197	0.912	900.884	0.002
$R_\star$	$R_\odot$	$\mathcal{T}(0.910; 0.052)$	0.911	0.910	0.051	0.814	1.004	2265.830	-0.001
$\log g$	cgs	$\mathcal{N}(4.600; 0.100)$	4.604	4.601	0.094	4.417	4.769	923.943	0.
$\langle f \rangle$	–	$\mathcal{N}(0.500; 0.100)$	0.4999	0.4999	0.0003	0.4993	0.5005	2964.6676	0.0013
$e^{(2)}$	–	Van Eylen et al. (2019)	0.127	0.168	0.147	0.	0.446	518.492	0.004
$\omega$	rad	$\mathcal{U}(0.000; 6.283)$	-0.235	-0.170	1.867	-2.879	3.132	1212.260	0.006
$\log \sigma_f$	–	$\mathcal{N}(\log \langle \sigma_f \rangle; 2.000)$	-8.016	-8.016	0.008	-8.031	-8.001	2224.427	-0.
$\rho$	d	$\mathcal{U}(1.000; 10.000)$	2.953	2.955	0.096	2.777	3.131	1936.492	0.
$\sigma$	$d^{-1}$	InvGamma(1.000; 5.000)	0.013	0.013	0.001	0.012	0.014	2155.687	-0.
$\sigma_{\text{rot}}$	$d^{-1}$	InvGamma(1.000; 5.000)	0.897	0.933	0.222	0.552	1.316	2147.324	0.003
$\log P_{\text{rot}}$	log(d)	$\mathcal{N}(0.958; 0.020)$	0.964	0.964	0.001	0.963	0.966	2486.029	0.
$\log Q_0$	–	$\mathcal{N}(0.000; 2.000)$	12.935	12.960	0.454	12.157	13.857	2126.581	0.004
$\log dQ$	–	$\mathcal{N}(0.000; 2.000)$	0.029	0.021	2.032	-3.785	3.780	1755.941	0.
$f$	–	$\mathcal{U}(0.100; 1.000)$	0.111	0.113	0.010	0.1	0.130	1253.358	-0.0
<i>Derived</i>									
$R_p/R_\star$	–	–	0.036	0.036	0.002	0.031	0.040	1173.566	0.003
$\rho_\star$	$\text{g cm}^{-3}$	–	2.27	2.312	0.516	1.408	3.272	910.746	0.001
$R_p$	$R_{\text{Jup}}$	–	0.316	0.317	0.037	0.251	0.386	1692.736	0.001
$a/R_\star$	–	–	18.396	18.407	1.375	15.690	20.780	910.717	0.001
$\cos i$	–	–	0.022	0.021	0.011	0.002	0.038	439.303	0.011
$T_{14}$	hr	–	2.822	2.823	0.057	2.710	2.917	1086.008	0.002
$T_{13}$	hr	–	2.578	2.566	0.083	2.430	2.714	590.630	0.011

NOTE— ESS refers to the number of effective samples.  $\hat{R}$  is the Gelman-Rubin convergence diagnostic. Logarithms through this table are in base- $e$ .  $\mathcal{U}$  denotes a uniform distribution,  $\mathcal{N}$  a normal distribution, and  $\mathcal{T}$  a truncated normal bounded between zero and an upper limit much larger than the mean. (1) The ephemeris is in units of BJDTDB - 2454833. (2) The eccentricity vectors are sampled in the  $(e \cos \omega, e \sin \omega)$  basis.

## REFERENCES

- Agol, E., Luger, R., & Foreman-Mackey, D. 2020, *AJ*, **159**, 123
- Akeson, R. L., Chen, X., Ciardi, D., et al. 2013, *PASP*, **125**, 989
- Albrecht, S., Winn, J. N., Johnson, J. A., et al. 2012, *ApJ*, **757**, 18
- Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, *AJ*, **156**, 123
- Beaumont, C., Robitaille, T., Borkin, M., & Goodman, A. 2014, glueviz v0.4: multidimensional data exploration
- Bhatti, W., Bouma, L. G., & Wallace, J. 2018, astrobase, <https://doi.org/10.5281/zenodo.1469822>
- Boisse, I., Pepe, F., Perrier, C., et al. 2012, *A&A*, **545**, A55
- Borucki, W. J., Koch, D., Basri, G., et al. 2010, *Science*, **327**, 977
- Bossini, D., Vallenari, A., Bragaglia, A., et al. 2019, *A&A*, **623**, A108
- Bouma, L. G., Curtis, J. L., Hartman, J. D., Winn, J. N., & Bakos, G. A. 2021, arXiv:2107.08050 [astro-ph], arXiv: 2107.08050
- Bouma, L. G., Hartman, J. D., Bhatti, W., Winn, J. N., & Bakos, G. Á. 2019, *ApJS*, **245**, 13
- Bouma, L. G., Hartman, J. D., Brahm, R., et al. 2020, *AJ*, **160**, 239
- Burke, C. J., Levine, A., Fausnaugh, M., et al. 2020, TESS-Point: High precision TESS pointing tool, Astrophysics Source Code Library, [ascl:2003.001](https://ascl.net/2003.001)
- Cantat-Gaudin, T., & Anders, F. 2020, *A&A*, **633**, A99
- Cantat-Gaudin, T., Jordi, C., Vallenari, A., et al. 2018, *A&A*, **618**, A93
- Cantat-Gaudin, T., Jordi, C., Vallenari, A., et al. 2018, *A&A*, **618**, A93
- Cantat-Gaudin, T., Jordi, C., Wright, N. J., et al. 2019, *A&A*, **626**, A17
- Cantat-Gaudin, T., Anders, F., Castro-Ginard, A., et al. 2020, *A&A*, **640**, A1
- Capitanio, L., Lallement, R., Vergely, J. L., Elyajouri, M., & Monreal-Ibero, A. 2017, *A&A*, **606**, A65
- Castro-Ginard, A., Jordi, C., Luri, X., et al. 2020, *A&A*, **635**, A45
- Cotten, T. H., & Song, I. 2016, *ApJS*, **225**, 15
- Dalal, S., Hébrard, G., Lecavelier des Étangs, A., et al. 2019, *A&A*, **631**, A28
- Damiani, F., Prisinzano, L., Pillitteri, I., Micela, G., & Sciortino, S. 2019, *A&A*, **623**, A112, publisher: EDP Sciences
- Damiani, F., Prisinzano, L., Pillitteri, I., Micela, G., & Sciortino, S. 2019, *A&A*, **623**, A112
- David, T., Hillenbrand, L., & Petigura, E. 2016, *Nature*, **534**, 658
- David, T. J., & Hillenbrand, L. A. 2015, *ApJ*, **804**, 146
- David, T. J., Petigura, E. A., Luger, R., et al. 2019, *ApJL*, **885**, L12
- Dias, W. S., Monteiro, H., Caetano, T. C., et al. 2014, *Astronomy and Astrophysics*, **564**, A79
- Dobbie, P. D., Lodieu, N., & Sharp, R. G. 2010, *MNRAS*, **409**, 1002
- Douglas, S. T., Agüeros, M. A., Covey, K. R., & Kraus, A. 2017, *ApJ*, **842**, 83
- Eggen, O. J. 1968, *The Astrophysical Journal*, **152**, 77
- Esplin, T. L., & Luhman, K. L. 2019, *AJ*, **158**, 54
- Foreman-Mackey, D. 2016, *Journal of Open Source Software*, **1**, 24
- Foreman-Mackey, D., Czekala, I., Luger, R., et al. 2020, exoplanet-dev/exoplanet v0.2.6
- Fulton, B. J., Petigura, E. A., Howard, A. W., et al. 2017, *AJ*, **154**, 109
- Fürnkranz, V., Meingast, S., & Alves, J. 2019, *A&A*, **624**, L11
- Gagné, J., David, T. J., Mamajek, E. E., et al. 2020, *ApJ*, **903**, 96
- Gagné, J., & Faherty, J. K. 2018, *ApJ*, **862**, 138
- Gagné, J., Roy-Loubier, O., Faherty, J. K., Doyon, R., & Malo, L. 2018a, *ApJ*, **860**, 43
- Gagné, J., Mamajek, E. E., Malo, L., et al. 2018b, *ApJ*, **856**, 23
- Gagné, J., David, T. J., Mamajek, E. E., et al. 2020, *ApJ*, **903**, 96
- Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2020, arXiv e-prints, arXiv:2012.01533
- Gaia Collaboration, Babusiaux, C., van Leeuwen, F., et al. 2018a, *A&A*, **616**, A10
- Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2018b, *A&A*, **616**, A1
- Gaia Collaboration, Smart, R. L., Sarro, L. M., et al. 2021, *A&A*, **649**, A6
- Gillen, E., Briegal, J. T., Hodgkin, S. T., et al. 2020, *MNRAS*, **492**, 1008
- Ginsburg, A., Sipocz, B., Madhura Parikh, et al. 2018, Astropy/Astroquery: V0.3.7 Release
- Ginzburg, S., Schlichting, H. E., & Sari, R. 2018, *MNRAS*, **476**, 759
- Goldman, B., Röser, S., Schilbach, E., Moór, A. C., & Henning, T. 2018, *ApJ*, **868**, 32

- Guerrero, N. M., Seager, S., Huang, C. X., et al. 2021, [arXiv:2103.12538 \[astro-ph\]](#), arXiv: 2103.12538
- Hattori, S., Foreman-Mackey, D., Hogg, D. W., et al. 2021, arXiv e-prints, arXiv:2106.15063
- Hippke, M., David, T. J., Mulders, G. D., & Heller, R. 2019, [AJ, 158, 143](#)
- Holczer, T., Shporer, A., Mazeh, T., et al. 2015, [ApJ, 807, 170](#)
- Holczer, T., Mazeh, T., Nachmani, G., et al. 2016, [ApJS, 225, 9](#)
- Hunter, J. D. 2007, Computing in Science & Engineering, 9, 90
- Inamdar, N. K., & Schlichting, H. E. 2015, [MNRAS, 448, 1751](#)
- Jones, E., Oliphant, T., Peterson, P., et al. 2001, Open source scientific tools for Python
- Jontof-Hutter, D., Wolfgang, A., Ford, E. B., et al. 2021, [AJ, 161, 246](#)
- Kerr, R., Rizzuto, A. C., Kraus, A. L., & Offner, S. S. R. 2021, [arXiv:2105.09338 \[astro-ph\]](#), arXiv: 2105.09338
- Kharchenko, N. V., Piskunov, A. E., Schilbach, E., Röser, S., & Scholz, R.-D. 2013, [A&A, 558, A53](#)
- Kipping, D. M. 2010, [MNRAS, 408, 1758](#)
- Kipping, D. M. 2013, [MNRAS, 435, 2152](#)
- Kite, E. S., Fegley, Jr., B., Schaefer, L., & Ford, E. B. 2020, [ApJ, 891, 111](#)
- Kounkel, M., & Covey, K. 2019, [AJ, 158, 122](#)
- Kounkel, M., & Covey, K. 2019, [AJ, 158, 122](#)
- Kounkel, M., Covey, K., & Stassun, K. G. 2020, [AJ, 160, 279](#)
- Kounkel, M., Covey, K., Suárez, G., et al. 2018, [AJ, 156, 84](#)
- Kounkel, M., Covey, K., Suárez, G., et al. 2018, [AJ, 156, 84](#)
- Kraus, A. L., Shkolnik, E. L., Allers, K. N., & Liu, M. C. 2014, [AJ, 147, 146](#)
- Lallement, R., Capitanio, L., Ruiz-Dern, L., et al. 2018, [A&A, 616, A132](#)
- Lewis, I. J., Cannon, R. D., Taylor, K., et al. 2002, [MNRAS, 333, 279](#)
- Lomb, N. R. 1976, [Astrophysics and Space Science, 39, 447](#)
- Luger, R., Agol, E., Foreman-Mackey, D., et al. 2019, [AJ, 157, 64](#)
- Mann, A. W., Newton, E. R., Rizzuto, A. C., et al. 2016, [AJ, 152, 61](#)
- Mann, A. W., Johnson, M. C., Vanderburg, A., et al. 2020, [AJ, 160, 179](#)
- Mazeh, T., Holczer, T., & Shporer, A. 2015, [ApJ, 800, 142](#)
- McKinney, W. 2010, in Proceedings of the 9th Python in Science Conference, ed. S. van der Walt & J. Millman, 51
- Meibom, S., Mathieu, R. D., & Stassun, K. G. 2007, [ApJL, 665, L155](#)
- Meibom, S., Barnes, S. A., Latham, D. W., et al. 2011, [The Astrophysical Journal Letters, 733, L9](#)
- Meibom, S., Torres, G., Fressin, F., et al. 2013, [Nature, 499, 55](#)
- Meingast, S., & Alves, J. 2019, [A&A, 621, L3](#)
- Meingast, S., Alves, J., & Rottensteiner, A. 2021, [A&A, 645, A84](#)
- Montet, B. T., Tovar, G., & Foreman-Mackey, D. 2017, [ApJ, 851, 116](#)
- Morris, B. M. 2020, [ApJ, 893, 67](#)
- Morton, T. D., Bryson, S. T., Coughlin, J. L., et al. 2016, [ApJ, 822, 86](#)
- Newton, E. R., Mann, A. W., Tofflemire, B. M., et al. 2019, [ApJ, 880, L17](#)
- Ogihara, M., Morbidelli, A., & Guillot, T. 2015, [A&A, 578, A36](#)
- Oh, S., Price-Whelan, A. M., Hogg, D. W., Morton, T. D., & Spergel, D. N. 2017, [AJ, 153, 257](#)
- Owen, J. E. 2020, [arXiv:2009.03919 \[astro-ph\]](#), arXiv: 2009.03919
- Owen, J. E., & Wu, Y. 2013, [ApJ, 775, 105](#)  
—. 2016, [ApJ, 817, 107](#)
- Pasquini, L., Avila, G., Blecha, A., et al. 2002, [The Messenger, 110, 1](#)
- Pavlidou, T., Scholz, A., & Teixeira, P. S. 2021, [MNRAS, 503, 3232](#)
- Pecaut, M. J., & Mamajek, E. E. 2013, [ApJS, 208, 9](#)
- Pérez, F., & Granger, B. E. 2007, [Computing in Science and Engineering, 9, 21](#)
- Plavchan, P., Barclay, T., Gagné, J., et al. 2020, [Nature, 582, 497](#)
- Randich, S., Tognelli, E., Jackson, R., et al. 2018, [A&A, 612, A99](#)
- Ratzenböck, S., Meingast, S., Alves, J., Möller, T., & Bomze, I. 2020, [A&A, 639, A64](#)
- Rebull, L. M., Stauffer, J. R., Bouvier, J., et al. 2016, [AJ, 152, 113](#)
- Reinhold, T., & Hekker, S. 2020, [A&A, 635, A43](#)
- Ricker, G. R., Winn, J. N., Vanderspek, R., et al. 2015, [Journal of Astronomical Telescopes, Instruments, and Systems, 1, 014003](#)
- Rizzuto, A. C., Mann, A. W., Vanderburg, A., Kraus, A. L., & Covey, K. R. 2017, [AJ, 154, 224](#)
- Rizzuto, A. C., Newton, E. R., Mann, A. W., et al. 2020, [arXiv:2005.00013 \[astro-ph\]](#)

- Roccatagliata, V., Franciosini, E., Sacco, G. G., Randich, S., & Sicilia-Aguilar, A. 2020, [A&A](#), 638, A85
- Röser, S., & Schilbach, E. 2020, [A&A](#), 638, A9
- Rowe, J. F., Bryson, S. T., Marcy, G. W., et al. 2014, [ApJ](#), 784, 45
- Rubenzahl, R. A., Dai, F., Howard, A. W., et al. 2021, [AJ](#), 161, 119
- Salvatier, J., Wieckiâ, T. V., & Fonnesbeck, C. 2016, PyMC3: Python probabilistic programming framework
- Sanchis-Ojeda, R., & Winn, J. N. 2011, [ApJ](#), 743, 61
- Scargle, J. D. 1982, [ApJ](#), 263, 835
- Shapiro, A. I., Solanki, S. K., Krivova, N. A., Yeo, K. L., & Schmutz, W. K. 2016, [A&A](#), 589, A46
- Sheinis, A., Anguiano, B., Asplund, M., et al. 2015, *Journal of Astronomical Telescopes, Instruments, and Systems*, 1, 035002
- Shporer, A., Zhou, G., Vanderburg, A., et al. 2017, [ApJL](#), 847, L18
- Soderblom, D. R., Hillenbrand, L. A., Jeffries, R. D., Mamajek, E. E., & Naylor, T. 2014, *Protostars and Planets VI*, 219
- Stauffer, J., Rebull, L., Bouvier, J., et al. 2016, [AJ](#), 152, 115
- Stauffer, J. R., Hartmann, L. W., Prosser, C. F., et al. 1997, [ApJ](#), 479, 776
- Stephenson, C. B. 1959, [PASP](#), 71, 145
- Tenenbaum, P., Christiansen, J. L., Jenkins, J. M., et al. 2012, [ApJS](#), 199, 24
- Theano Development Team. 2016, [arXiv e-prints](#), [abs/1605.02688](#)
- Thompson, S. E., Coughlin, J. L., Hoffman, K., et al. 2018, [ApJS](#), 235, 38
- Tian, H.-J. 2020, [ApJ](#), 904, 196
- Tofflemire, B. M., Rizzuto, A. C., Newton, E. R., et al. 2021, [AJ](#), 161, 171
- Tokovinin, A., Fischer, D. A., Bonati, M., et al. 2013, [PASP](#), 125, 1336
- Ujjwal, K., Kartha, S. S., Mathew, B., Manoj, P., & Narang, M. 2020, [AJ](#), 159, 166
- Van Eylen, V., Albrecht, S., Huang, X., et al. 2019, [AJ](#), 157, 61
- Villa Vélez, J. A., Brown, A. G. A., & Kenworthy, M. A. 2018, *Research Notes of the American Astronomical Society*, 2, 58
- Walt, S. v. d., Colbert, S. C., & Varoquaux, G. 2011, *Computing in Science & Engineering*, 13, 22
- Wenger, M., Ochsenbein, F., Egret, D., et al. 2000, [A&AS](#), 143, 9
- Zari, E., Hashemi, H., Brown, A. G. A., Jardine, K., & de Zeeuw, P. T. 2018, [A&A](#), 620, A172
- Zari, E., Hashemi, H., Brown, A. G. A., Jardine, K., & de Zeeuw, P. T. 2018, [A&A](#), 620, A172
- Zhou, G., Quinn, S. N., Irwin, J., et al. 2021, [AJ](#), 161, 2

**Table 2.** Young, Age-dated, and Age-dateable Stars Within the Nearest Few Kiloparsecs (v0.5 of the CDIPS Target List).

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	Uma,IR_excess,NASAExoArchive_ps_20210506	Comma-separated cluster or group name.
age	nan,nan,9.48	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	Ujjwal2020,CottenSong2016,NASAExoArchive_ps_20210506	Comma-separated provenance of group membership.
reference_bibcode	2020AJ....159..166U,2016ApJS..225...15C,2013PASP..125..989A	ADS bibcode corresponding to reference_id.

NOTE— Table 2 is published in its entirety in a machine-readable format. This table is a concatenation of the studies listed in Table 3. 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.

## APPENDIX

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

The v0.5 CDIPS target catalog (Table 2) includes some important updates from previous versions. As in Bouma et al. (2019), we collected membership information for young, age-dated, or age-dateable stars from across the literature. Table 3 gives a list of the sources included, and some brief summary statistics.

The first major important change is that the extent of analyses performed on the Gaia data at the time of our compilation was wide and deep enough that we opted to neglect pre-Gaia analyses, except in cases for which spectroscopically confirmed samples of stars had been collected. The membership lists for instance of Kharchenko et al. (2013) and Dias et al. (2014) (MWSC and DAML) were no longer required, especially given their relatively high field-star contamination rates compared to Gaia-derived membership catalogs.

For any of the catalogs for which Gaia DR2 identifiers were not immediately available, we either followed the spatial (plus proper-motion) crossmatching procedures described in Bouma et al. (2019), or else we pulled the Gaia DR2 source identifiers associated with the catalog from SIMBAD. We consequently opted to drop the ext\_catalog\_name and dist columns maintained in Bouma et al. (2019), as these were only populated for a small number of stars.

The most crucial parameters of a given star for our purposes are the Gaia DR2 source identifier

(source\_id), the cluster name (cluster), and the (age). Given the hierarchical nature of many stellar associations, we do not attempt to resolve the cluster names to a single unique string. The Orion complex for instance, can be divided into almost one hundred kinematic subgroups (Kounkel et al. 2018). Similar complexity applies to the problem of determining homogeneous ages, which we do not attempt to resolve. Instead, we simply merged the cluster names and ages reported by various authors together.

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

The catalogs we included for which ages were not immediately available were those of Cotten & Song (2016), Oh et al. (2017), Zari et al. (2018), Gagné et al. (2018b), Gagné et al. (2018a), Gagné

**Table 3.** 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
Wenger et al. (2000) Y*O	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 3 describes the provenances for the young and age-dateable stars in Table 2.  $N_{\text{Gaia}}$ : 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 appear multiple times when multiple tables from the analysis were included in the concatenation. **todo: add Kerr2021, fix the > < signs**

& Faherty (2018), and Ujjwal et al. (2020). While in principle the moving group members discussed by Gagné et al. (2018b,a); Gagné & Faherty (2018) and Ujjwal et al. (2020) have easily associated ages, our SIMBAD cross-matching lost the moving group association from those studies, which should

therefore be recovered tools such as BANYAN  $\Sigma$ .<sup>3</sup>. We also included the SIMBAD object identifiers TT\*, Y\*O,Y\*?, TT?, and pMS\*. Finally, we also included every star in the NASA Exoplanet Archive ps table that had a Gaia identifier available (Akeson et al. 2013). If the age had finite un-

<sup>3</sup> <http://www.exoplanetes.umontreal.ca/banyan/banyansigma.php>

certainties, we also included it, since stellar ages determined through the combination of isochrone-fitting and transit-derived stellar densities typically have higher precision than from isochrones alone.

The technical manipulations for the merging, cleaning, and joining were performed using pandas (McKinney 2010). The eventual crossmatch (using the Gaia DR2 source\_id) against the Gaia DR2 archive was performed asynchronously on the Gaia archive website<sup>4</sup>.

## B. KINEMATIC SELECTION OF $\delta$ LYR CLUSTER MEMBERS

Figure 6 shows stars reported by Kounkel & Covey (2019) to be in the group “Theia 73”, which was cross-matched by Kounkel & Covey as being “Stephenson 1”. Kounkel & Covey (2019) reported 3,071 stars to be present in this cluster. For our Figure, galactic positions were calculated and plotted only for stars with parallax signal-to-noise exceeding 20. The location of the Sun is shown on the plots. The smattering of reported cluster members (gray and black points) has a significantly different structure than the cluster initially identified by Stephenson (1959) and corroborated by Eggen (1968). While the non-uniform “clumps” might be part of a real structure, they could also be an artifact of the data processing steps performed by Kounkel & Covey (2019). We therefore opted to only consider stars in the immediate kinematic and spatial vicinity of Kepler 1627. The tangential velocities relative to Kepler 1627 are shown in the bottom right panel. These are computed by assuming that every star has the same three-dimensional spatial velocity as Kepler 1627, where we assume a systemic radial velocity of  $-16.7 \pm 0.2 \text{ km s}^{-1}$  based on the reconnaissance spectra obtained by A. Howard on HIRES and D. Latham on TRES. The relevant projection effects are then taken into account, as discussed by *e.g.*, Meingast et al. (2021) and Bouma et al. (2021). We performed the actual selection by then manually drawing lassos with the interactive `glue` visualization tool (Beaumont et al. 2014) in the four projections shown in Figure 6. While bona fide members likely exist outside of our selection region (and our selection also includes some field star interlopers), our aim is to verify the existence of the cluster in the vicinity of Kepler 1627, and to measure its age. The procedure we have adopted enables both tasks.

## C. BINARITY AND STELLAR ROTATION FOR $\delta$ LYR CLUSTER MEMBERS

Figure 7 shows the rotation-color diagram for  $\delta$  Lyr cluster members, with the points colored ac-

cording to indicators of binarity. The binaries tend to be redder and have shorter rotation periods. We know of two possible explanations based on selection effects, and two based on physics.

Possible selection effects include *i*) binaries have a component that contributes additional red light to the system, which could skew the color measurement of the primary; and *ii*) the unresolved binary companions could contaminate the rotation period measurement (*e.g.*, Stauffer et al. 2016, Section 5.1).

The two possible physical effects that could be relevant are tidal locking and pre-main-sequence disk locking. Tidal locking has been argued to be an unlikely explanation for the frequency of rapid rotators due to the population statistics; a more likely scenario would be that the presence of the binary leads to faster disk dispersal, enabling the primary to contract to more rapid rotation periods than possible for single stars (Meibom et al. 2007; Bouma et al. 2020).

## D. TRANSIT AND STELLAR VARIABILITY MODEL

### D.1. Long Cadence Data

We assumed a quadratic limb-darkening law, with the uninformative prior advocated by CITET Kipping2013.

Our default stellar variability model (RotGPtransit) allows for a Rot Term GP kernel, with a logjitter term to inflate the uncertainties to account for excess white noise. (This results in an inflation of the uncertainties by a factor of about three).

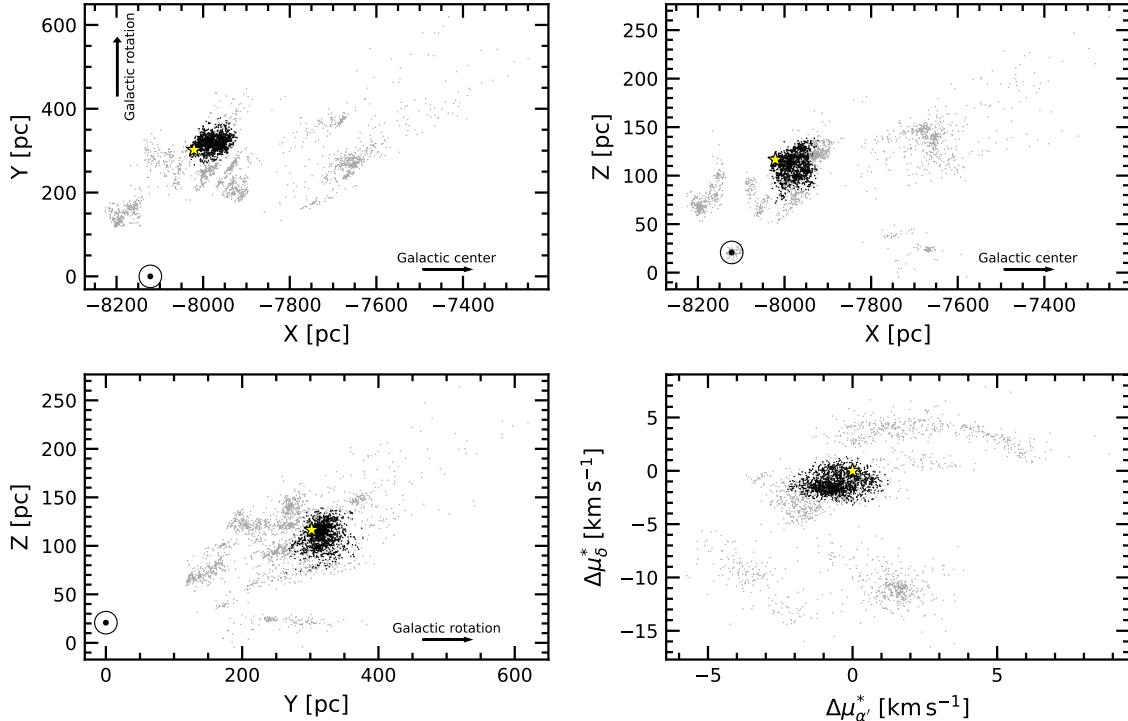
We considered including an additive SHOTerm kernel to account for stochastic noise (RotStochG-Ptransit). This didn’t seem to affect the results much, so we opted for the simpler model.

Figure ?? shows X, Y, Z. The residuals during the transit may hint at some small degree of unfitted signal – in the sense that the observations are systematically high in the first half of transit, and low in the second half. This is the phase of maximal scatter during each orbital period (TODO: in vs out of transit scatter). The amplitude of the anomaly is  $\approx 30 \text{ ppm}$ .

To explore the origin of the anomaly, we binned the Kepler data over quarters (Figure 8) and years (Figure 9). In Figure 8 Quarter 6 shows the strongest asymmetry out of any of the quarters: a deviation of about 3 ppt from expectation. Quarter 7 shows an anomaly at roughly the same transit phase. Year 2 correspondingly shows the strongest anomaly out of any year in Figure 9.

We considered three possible explanations for the anomaly: gravity darkening, transit timing variations, and spot-crossing events.

<sup>4</sup> <https://gea.esac.esa.int/archive/>



**Figure 6. Galactic position and tangential velocities of the  $\delta$  Lyr cluster.** Points are candidate cluster members with  $\bar{\omega}/\sigma_{\bar{\omega}} > 20$ , reported to be in the group by Kounkel & Covey (2019). We focus on stars in a small region (black points) in the kinematic vicinity of Kepler 1627 (yellow star). The other candidate cluster members (gray points) may or may not share the ages of the selected kinematic group. The location of the Sun is ( $\odot$ ) is shown.

Gravity darkening (e.g., CITE Masuda 2015) is based on the premise that the rapidly rotating star becomes oblate, and brighter near the poles than the equator. The fractional shape change due to gravity darkening 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 the parameters from Table 1, this yields an expected 1.6% distortion of the  $\approx 1.8$  ppt transit depth: *i.e.*, an absolute deviation of  $\approx 3$  ppm. This is smaller than the observed anomaly by an order of magnitude, and therefore seems unlikely.

The scenario of transit timing variations (TTVs) producing the asymmetry seems unlikely, since the analysis by Holczer et al. (2016) implies that any such variations in Kepler 1627 need to be less than of order minutes. Figures 8 and 9 also provide little evidence in support of this possibility.

The final possibility is that of starspot crossings. Given the high expected spot-covering fractions for young stars (e.g., Morris 2020, Plavchan et al. 2020), Kepler 1627Ab may cross spot groups on the stellar surface in projection. Spot-crossing anomalies often reach amplitudes exceeding 100 ppm (e.g., CITE Dai 2018, maybe Southworth’s recent stuff). For our system,  $P_{\text{orb}}/P_{\text{rot}} \approx 2.76$ . This means that every 4 transits (and 11 stellar rotations), the planet crosses over roughly the

same stellar longitude. Over a given quarter, this could occur at most three times, assuming the spot groups are persistent.

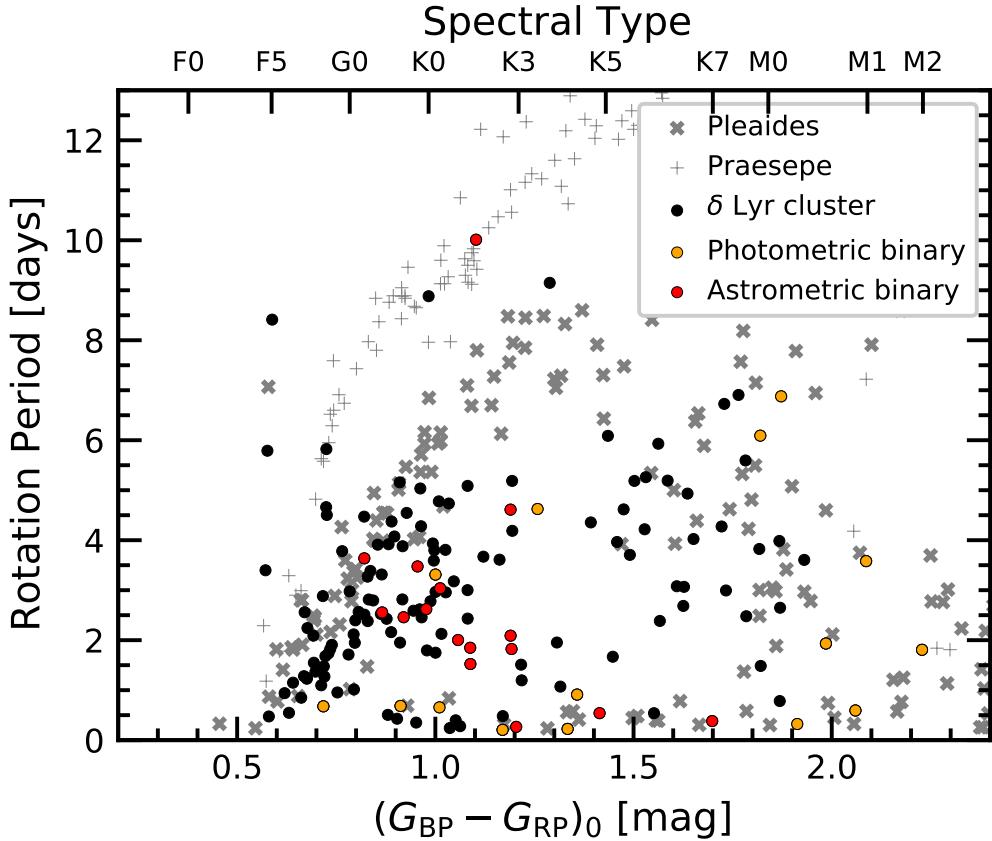
This could provide a plausible path to creating the distorted transit signal, based on the amplitudes involved. However, the typical S/N per Kepler transit is  $\approx 8$ , making individual spot-crossing unresolved, and this scenario challenging to test. *A priori*, one would also expect the spot-crossing events to not have a preferred orbital phase, so that they would average out over the  $\approx 200$  Kepler transits. Nonetheless, given the available data, it is our best explanation for what we believe is real additional photometric scatter observed during the transits.

## D.2. Short Cadence Data

Figure 10 shows the result from analyzing 100 days of short cadence (1-minute) data acquired during Quarter 15.

## E. FLARE ANALYSIS

The short cadence (1-minute) Kepler observations span 97.7 days. The light curve shows a significant number of flares (Figure 11). To quantify clean the light curve and identify the flares, we performed the following procedure. For cleaning, we performed the following iterative detrending procedure.



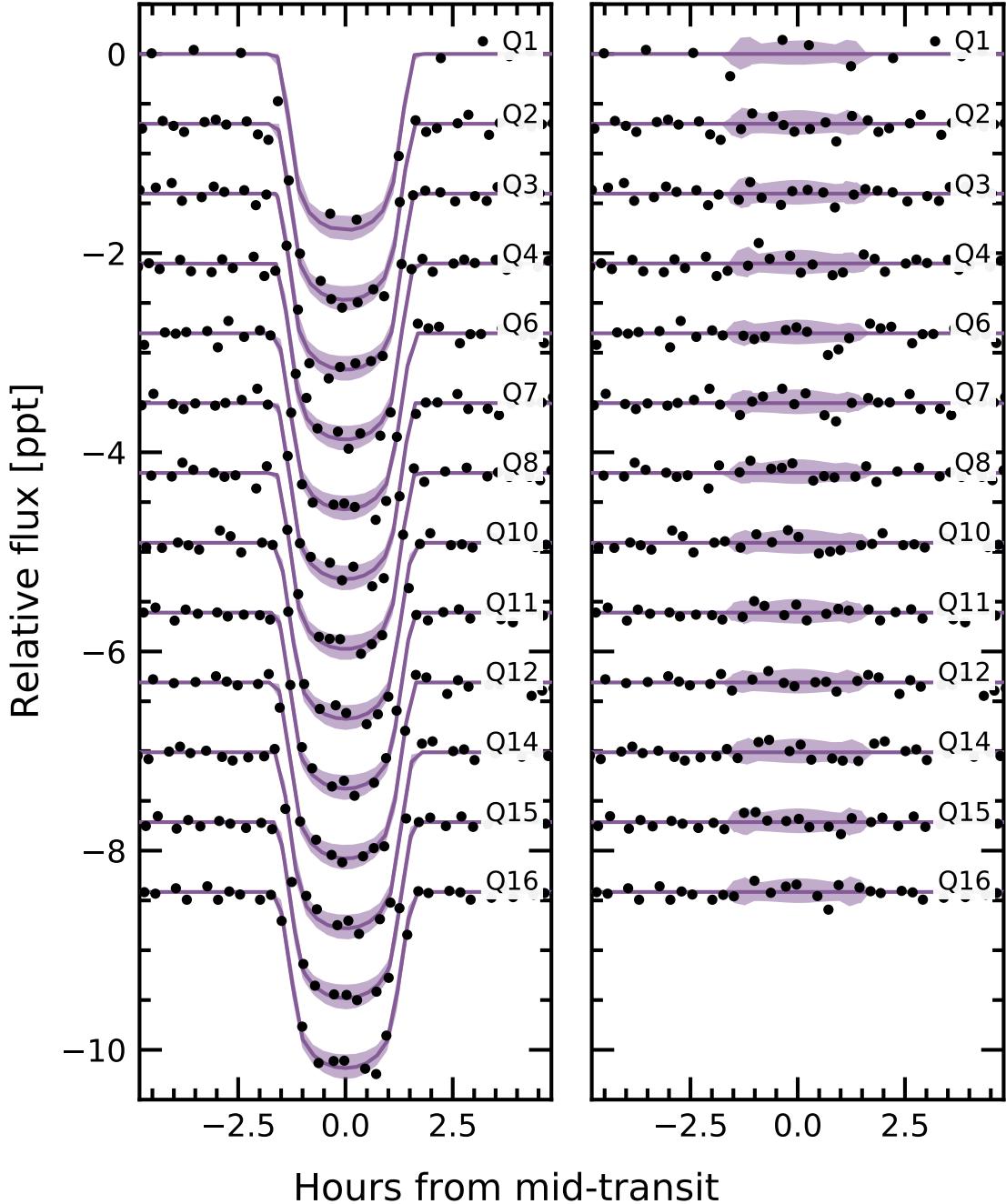
**Figure 7. Binarity indicators for rotators in  $\delta$  Lyr cluster.** Data are as in the lower panel of Figure 2. Astrometric binaries ( $\text{RUWE} > 1.2$ ) appear in red; photometric binaries ( $>0.3$  mag above an empirical isochrone; 20% of stars) are orange. Over  $0.6 < (G_{\text{BP}} - G_{\text{RP}})_0 < 1.5$ , the detected rotation periods in binaries tend to fall below the slow sequence.

- Step 1: Build a two-term mixed SHOTerm GP model with quasi-periodic kernels at  $\text{Prot}$  and  $0.5 \times \text{Prot}$ . Fit the model to the time, flux, and flux uncertainty.
- Step 2: Select points more than twice the median absolute deviation from the residual, and exclude them from the light curve. Repeat Step 1.
- Step 3: On 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.

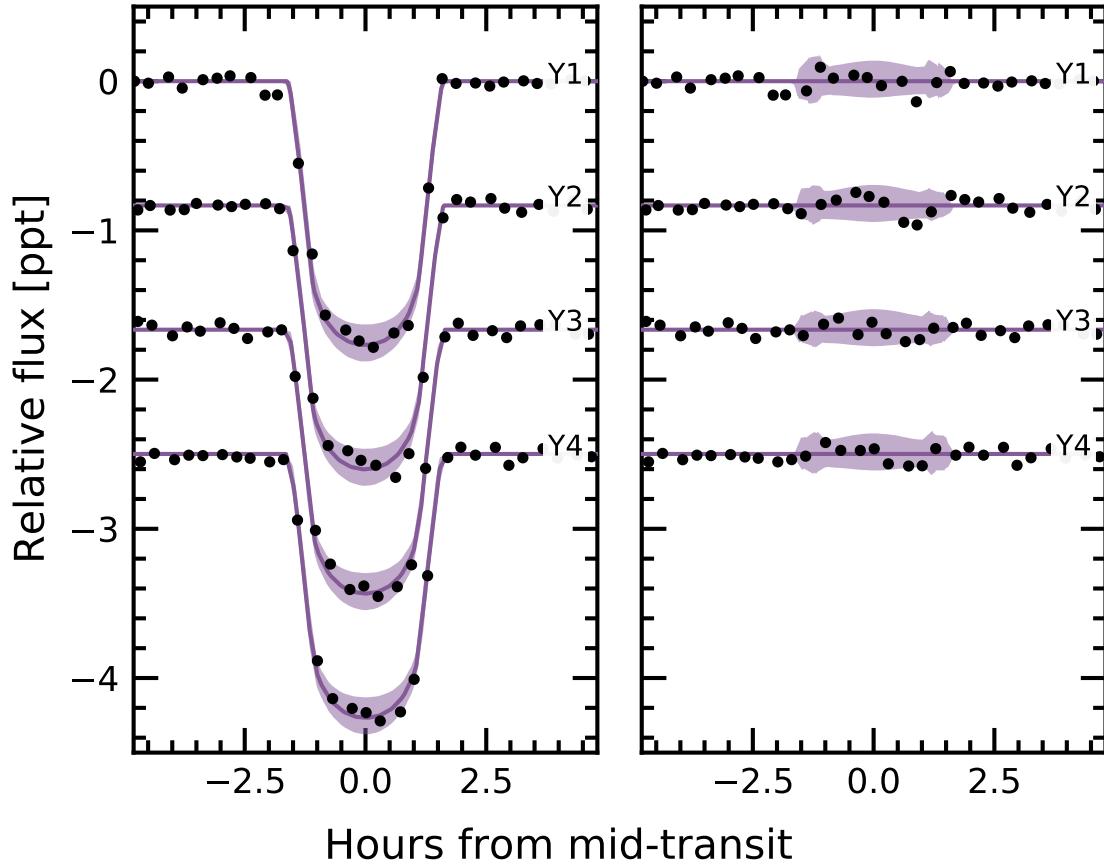
The flares were then identified and fitted using ALTAIPONY (CITE CITE). The fitted model is that of CITE XXX, which parametrizes the flare with a start time, a lag time, and an amplitude (CHECK). Figure 12 shows the resulting flares, amplitudes, and phases.

There were  $N_f = 24$  flares exceeding 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). For better or worse, 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 11. The timing error is good to a  $\approx 0.2\%$  difference from the orbital period, which 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, when one does not have a firm prediction, it is difficult to assess the statistical significance of a surprise. 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 from a Poisson distribution. Specifically, we drew  $N_f = 24$  samples from a Pois-



**Figure 8. Transit model residuals through time.** *Left:* Phase-folded transit of Kepler 1627b, with stellar variability removed, binned by Kepler quarter. Black points are binned to 20 minute intervals. The  $1-\sigma$  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. Quarters 6 and 7 show a consistent deviation in the second half of the transit.



**Figure 9. Transit model residuals through time (Part 2).** *Left:* Phase-folded transit of Kepler 1627b, with stellar variability removed, binned by year of observation. Points and models are as in Figure 8. *Right:* As on the left, with the transit removed.

son distribution with  $\lambda = \Delta t / N_f$ , for  $\Delta t = 97.7$  days the duration of the observations, and repeated the draw  $10^3$  times with unique random seeds.

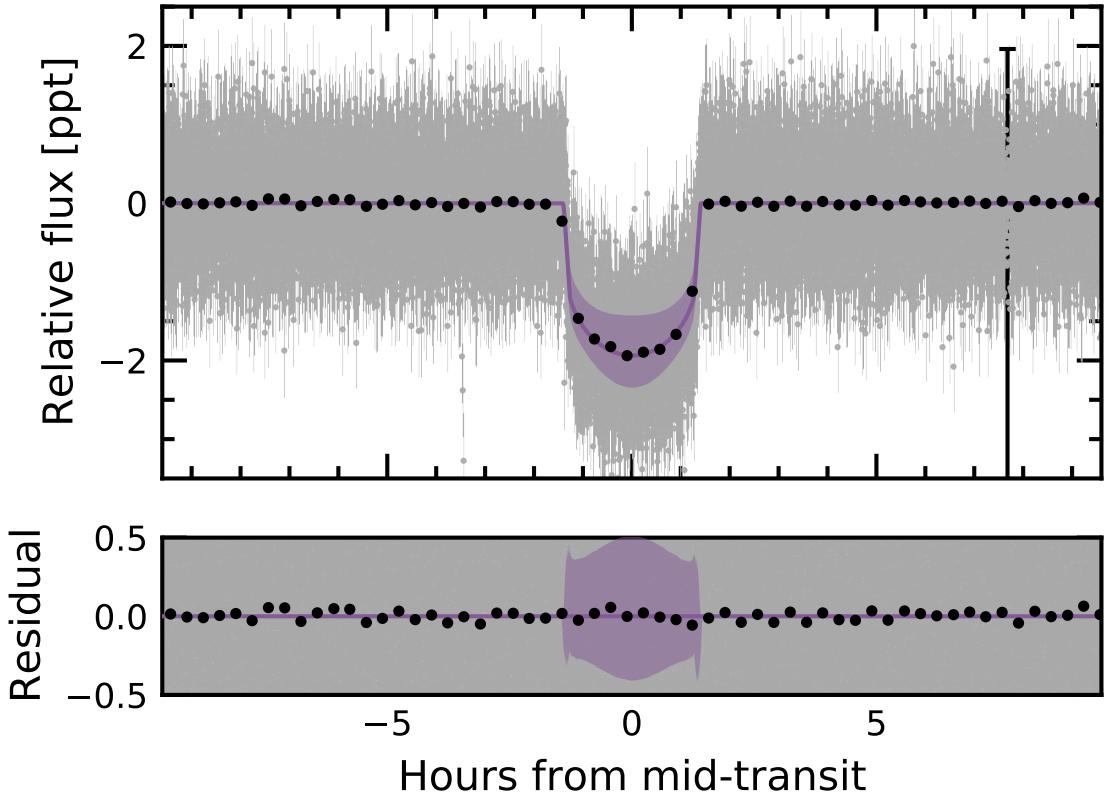
Figure 13 shows the results. The vertical lines in the figure show the planetary orbital period, the synodic period  $P_{\text{syn}} = (P_{\text{rot}}^{-1} - P_{\text{orb}}^{-1})^{-1}$ , and linear combinations thereof. Note that the tidal period (half the synodic period) is not shown. The bins are logarithmically 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 a rather 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 particularly strong. 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 more versed in the topic than ourselves could very well yield new insights.

## F. COMPANION STAR AND FALSE POSITIVE ASSESSMENT

Our analysis of archival Keck-NIRC2 Kp-band ( $2.12 \mu\text{m}$ ) imaging revealed the existence of a previously unreported stellar neighbor, unresolved in the Gaia source catalog **FIXME: uncertainties**. The NIRC2 images yield a projected separation  $\rho = 0.^{\circ}17$ , with  $\Delta \text{Kp} = 2.5$ . Using the measured Gaia EDR3 parallax for the system, this implies a projected separation of 54 AU. The presence of this star is consistent with the excess noise in the Gaia astrometric time-series (see Section F). Given the low chance of a star imaged within **X.X arcseconds** to be a chance companion along the line of sight, we proceed under the assumption that it is bound, and that the Kepler 1627 system is binary.

Unfortunately, we do not have any reliable color information about Kepler 1627B. Based on the tabulation<sup>5</sup> by [Pecaut & Mamajek \(2013\)](#), the measured NIR-contrast for a main sequence G8V star

<sup>5</sup> [http://www.pas.rochester.edu/~emamajek/EEM\\_dwarf\\_UBVIJHK\\_colors\\_Teff.txt](http://www.pas.rochester.edu/~emamajek/EEM_dwarf_UBVIJHK_colors_Teff.txt), version 2021/03/02.



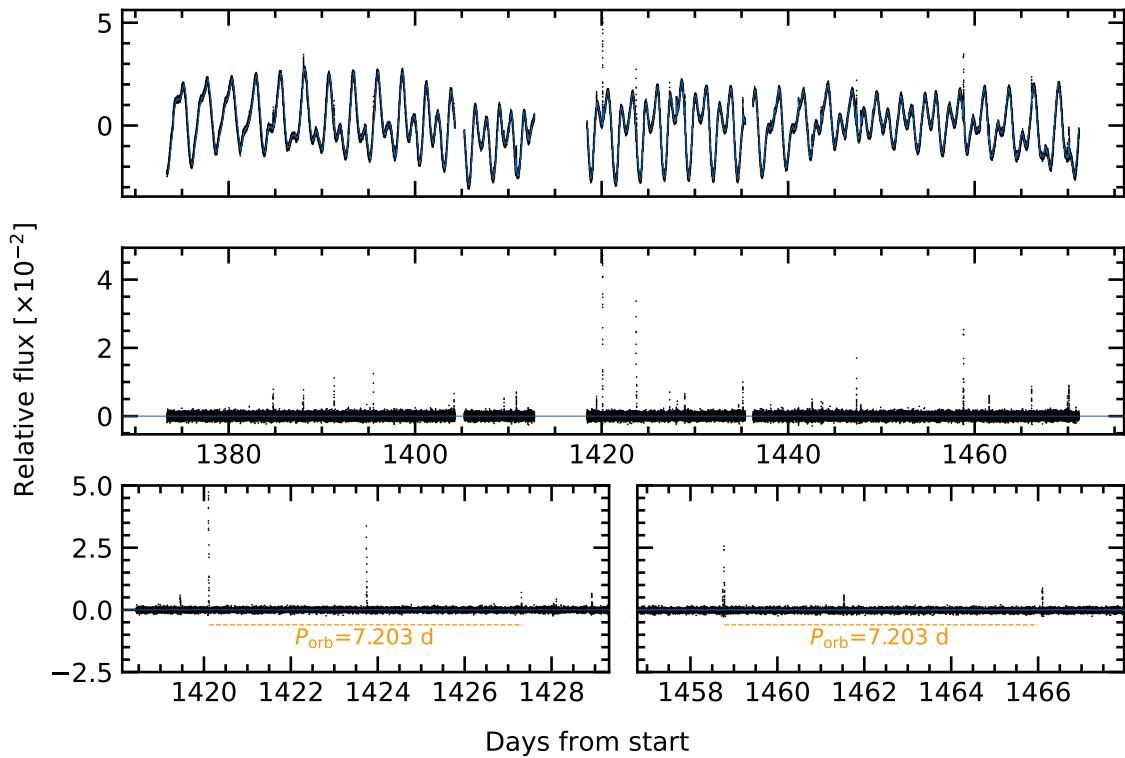
**Figure 10. Phase-folded short cadence transit of Kepler 1627 Ab.** The precision with which the impact parameter can be measured is higher than from the long cadence data.

corresponds to a spectral type for the companion of  $\approx$ M2V ( $M_* \approx 0.44 M_\odot$ ). However, the companion should have a longer pre-main-sequence contraction phase than the primary, which would imply that this mass is overestimated by  $\sim 10$  to 20%.

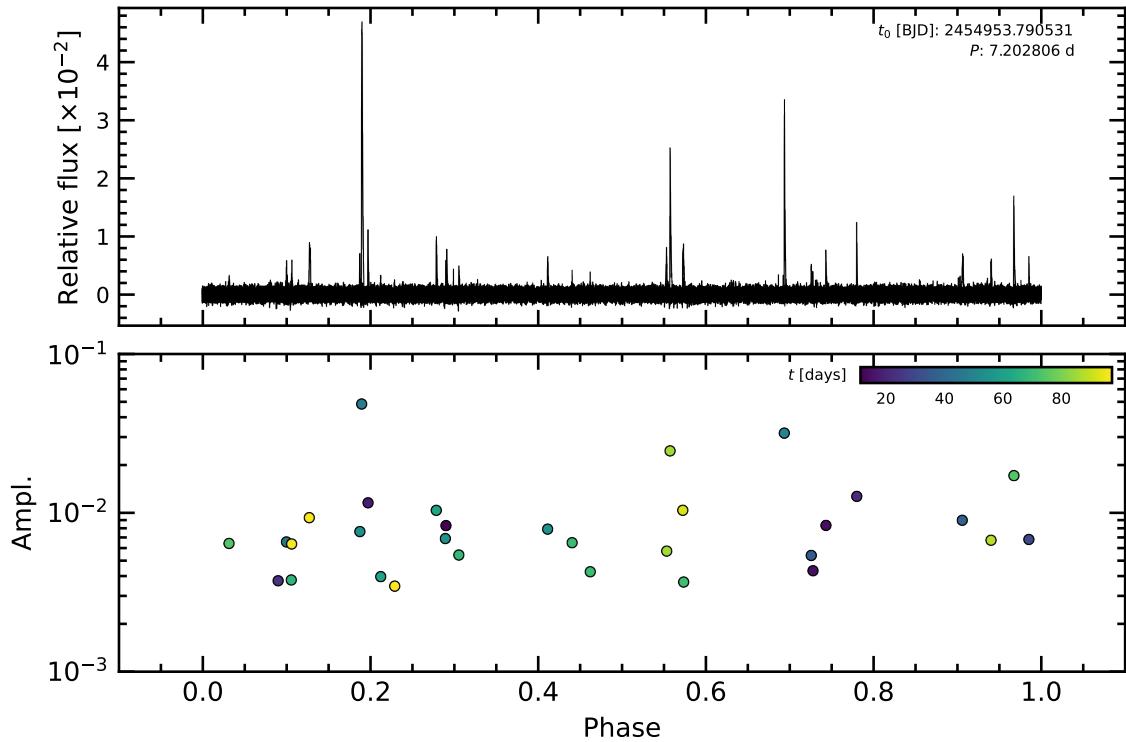
Could the companion be creating a false positive signal? The companion star contributes  $\approx 1\%$  of the total flux observed in the Kepler aperture (0.7-2.0% V-band to G-band dmag difference from EEM table; todo is fix using model spectra). The observed transit has a depth of 0.18%. An 18% deep eclipse of the secondary star would therefore be needed to produce a deep enough signal. The shape of the observed signal requires allowed impact parameters to span 0.02 to 0.73 (Table 1); the body transiting the secondary would therefore need to be non-grazing with  $R_3/R_2 \approx 0.42$ . Assuming a  $\approx 0.42 R_\odot$  radius of the imaged secondary, this would imply a tertiary stellar radius of  $\approx 0.2 R_\odot$ . This scenario ultimately yields a contradiction, because it would require an ingress and egress phase that each span  $\approx 40\%$  of the transit duration ( $\approx 65$  minutes). The actual measured ingress and egress duration is  $\approx 15$  minutes),  $4.4 \times$  shorter.

*Stellar density implied by transit duration*—The duration of the transit ( $2.823 \pm 0.057$  hr) and the implied stellar density ( $2.31 \pm 0.52$  g cm $^{-3}$ ) could in theory

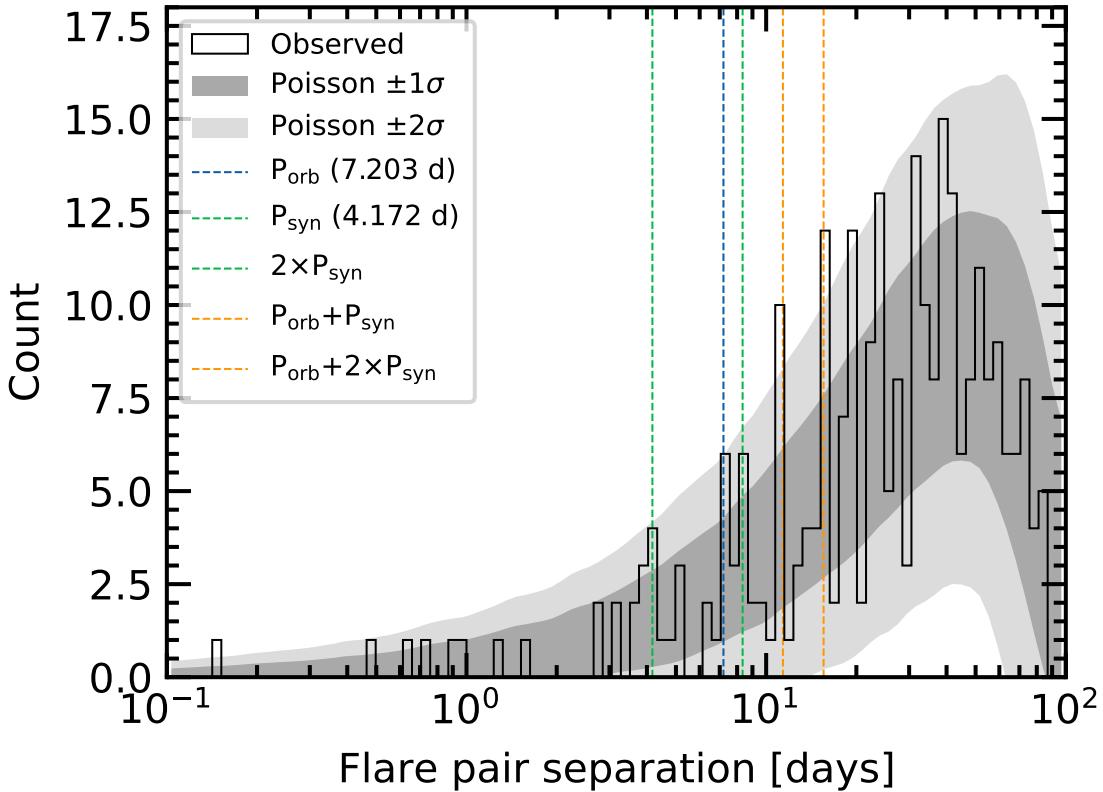
help rule between whether the transiting body orbits the primary or secondary star. Ultimately, our stellar density measurement is not precise enough to render the blend scenario implausible. At 30 Myr, a  $0.40 M_\odot$  solar-metallicity dwarf is  $\approx 26\%$  larger than when it is fully contracted on the main sequence ( $0.40 R_\odot$  vs.  $0.50 R_\odot$ ; CITEALT: Choi et al, MIST grids). The theoretically implied companion density of  $3.2$  g cm $^{-3}$  is indeed larger than the primary star's density of  $1.80$  g cm $^{-3}$  (measured through the HIRES reconnaissance spectroscopy). However the stellar density measured from the transit fitting ( $2.31 \pm 0.52$  g cm $^{-3}$ ) is only discrepant at the  $\approx 2\sigma$  level from the M-dwarf blend scenario. It is instead the combination of the flux contrast, transit depth, and ingress duration that rule out this scenario.



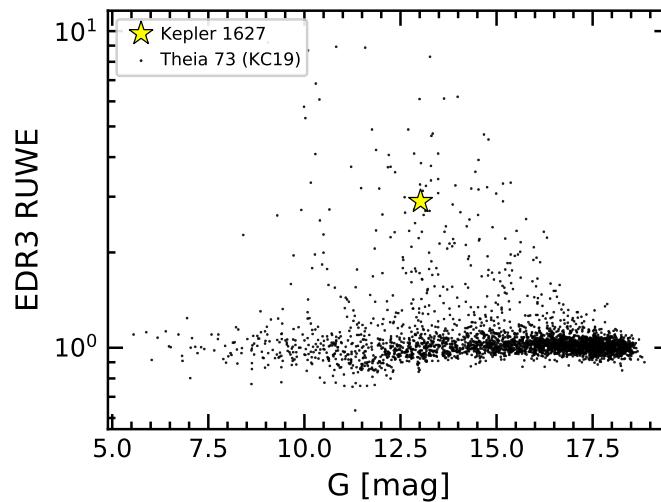
**Figure 11. Flares in Kepler 1627.** *Top:* The full short-cadence Kepler dataset, acquired at 1-minute sampling (black points) is shown with a stellar variability model (blue line). *Middle:* Residual after subtracting the stellar variability model. Flares appear as spikes. *Bottom:* Zooms of the brightest, and third-brightest flares. A timing coincidence – that both flares have “successors” approximately one orbital period after the initial event – is emphasized.



**Figure 12. Phase-folded flares in Kepler 1627.** *Top:* As in the middle of Figure 11, phase-folded at the planet's ephemeris. *Bottom:* Fitted flare amplitudes and orbital phases – each point represents one flare. Colors indicate relative time, from the beginning of the 98-day short-cadence Quarter 15 dataset (dark blue) to the end (light yellow). Lower amplitude flares likely exist in the data, and were not examined.



**Figure 13. 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 the locations of these “special periods”, the statistical evidence for a non-Poissonian process driving the flares does not reach the  $5\sigma$  threshold.



**Figure 14. Gaia EDR3 renormalized unit weight error (RUWE) point estimates for Kepler 1627A and other members of the  $\delta$  Lyr cluster.** Since other members of the cluster with similar colors have comparable degrees of photometric variability, the high RUWE estimate suggests that Kepler 1627A is a binary.



**Figure 15.** Keck NIRC2 AO image of Kepler 1627A and Kepler 1627B. SCALEBAR DENOTES... North is up, East is left.