

An alternative interpretation of the exomoon candidate signal in the combined *Kepler* and *Hubble* data of Kepler-1625

René Heller¹, Kai Rodenbeck^{2,1}, and Giovanni Bruno³

¹ Max Planck Institute for Solar System Research, Justus-von-Liebig-Weg 3, 37077 Göttingen, Germany
e-mail: heller@mps.mpg.de; rodenbeck@mps.mpg.de

² Institute for Astrophysics, Georg August University Göttingen, Friedrich-Hund-Platz 1, 37077 Göttingen, Germany

³ INAF, Astrophysical Observatory of Catania, Via S. Sofia 78, 95123 Catania, Italy
e-mail: giovanni.bruno@inaf.it

Received 18 December 2018 / Accepted 25 February 2019

ABSTRACT

Context. *Kepler* and *Hubble* photometry of a total of four transits by the Jupiter-sized exoplanet Kepler-1625 b have recently been interpreted to show evidence of a Neptune-sized exomoon. The key arguments were an apparent drop in stellar brightness after the planet's October 2017 transit seen with *Hubble* and its 77.8 min early arrival compared to a strictly periodic orbit.

Aims. The profound implications of this first possible exomoon detection and the physical oddity of the proposed moon, i.e., its giant radius prompt us to examine the planet-only hypothesis for the data and to investigate the reliability of the Bayesian information criterion (BIC) used for detection.

Methods. We combined *Kepler*'s Pre-search Data Conditioning Simple Aperture Photometry (PDCSAP) with the previously published *Hubble* light curve. In an alternative approach, we performed a synchronous polynomial detrending and fitting of the *Kepler* data combined with our own extraction of the *Hubble* photometry. We generated five million parallel-tempering Markov chain Monte Carlo (PTMCMC) realizations of the data with both a planet-only model and a planet-moon model, and compute the BIC difference (Δ BIC) between the most likely models, respectively.

Results. The Δ BIC values of -44.5 (using previously published *Hubble* data) and -31.0 (using our own detrending) yield strong statistical evidence in favor of an exomoon. Most of our orbital realizations, however, are very different from the best-fit solutions, suggesting that the likelihood function that best describes the data is non-Gaussian. We measure a 73.7 min early arrival of Kepler-1625 b for its *Hubble* transit at the 3σ level. This deviation could be caused by a 1 d data gap near the first *Kepler* transit, stellar activity, or unknown systematics, all of which affect the detrending. The radial velocity amplitude of a possible unseen hot Jupiter causing the Kepler-1625 b transit timing variation could be approximately 100 m s^{-1} .

Conclusions. Although we find a similar solution to the planet-moon model to that previously proposed, careful consideration of its statistical evidence leads us to believe that this is not a secure exomoon detection. Unknown systematic errors in the *Kepler/Hubble* data make the Δ BIC an unreliable metric for an exomoon search around Kepler-1625 b, allowing for alternative interpretations of the signal.

Key words. planets and satellites: dynamical evolution and stability – methods: data analysis – planets and satellites: detection – eclipses – planets and satellites: individual: Kepler-1625 b – techniques: photometric

1. Introduction

The recent discovery of an exomoon candidate around the transiting Jupiter-sized object Kepler-1625 b orbiting a slightly evolved solar mass star (Teachey et al. 2018) came as a surprise to the exoplanet community. This Neptune-sized exomoon, if confirmed, would be unlike any moon in the solar system, it would have an estimated mass that exceeds the total mass of all moons and rocky planets of the solar system combined. It is currently unclear how such a giant moon could have formed (Heller 2018).

Rodenbeck et al. (2018) revisited the three transits obtained with the *Kepler* space telescope between 2009 and 2013 and found marginal statistical evidence for the proposed exomoon. Their transit injection-retrieval tests into the out-of-transit *Kepler* data of the host star also suggested that the exomoon could well be a false positive. A solution to the exomoon question was supposed to arrive with the new *Hubble* data of an October 2017 transit of Kepler-1625 b (Teachey & Kipping 2018).

The new evidence for the large exomoon by Teachey & Kipping (2018), however, remains controversial. On the one

hand, the *Hubble* transit light curve indeed shows a significant decrease in stellar brightness that can be attributed to the previously suggested moon. Perhaps more importantly, the transit of Kepler-1625 b occurred 77.8 min earlier than expected from a sequence of strictly periodic transits, which is in very good agreement with the proposed transit of the exomoon candidate, which occurred before the planetary transit. On the other hand, an upgrade of *Kepler*'s Science Operations Center pipeline from version 9.0 to version 9.3 caused the exomoon signal that was presented in the Simple Aperture Photometry (SAP) measurements in the discovery paper (Teachey et al. 2018) to essentially vanish in the SAP flux used in the new study of Teachey & Kipping (2018). This inconsistency, combined with the findings of Rodenbeck et al. (2018) that demonstrate that the characterization and statistical evidence for this exomoon candidate depend strongly on the methods used for data detrending, led us to revisit the exomoon interpretation in light of the new *Hubble* data.

Here we address two questions. How unique is the proposed orbital solution of the planet-moon system derived with the Bayesian information criterion (BIC)? What could be the

reason for the observed 77.8 min difference in the planetary transit timing other than an exomoon?

2. Methods

Our first goal was to fit the combined *Kepler* and *Hubble* data with our planet-moon transit model (Rodenbeck et al. 2018) and to derive the statistical likelihood for the data to represent the model. In brief, we first model the orbital dynamics of the star–planet–moon system using a nested approach, in which the planet-moon orbit is Keplerian and unperturbed by the stellar gravity. The transit model consists of two black circles, one for the planet and one for the moon, that pass in front of the limb-darkened stellar disk. The resulting variations in the stellar brightness are computed using Ian Crossfield’s python code of the Mandel & Agol (2002) analytic transit model¹. The entire model contains 16 free parameters and it features three major updates compared to Rodenbeck et al. (2018): (1) planet-moon occultations are now correctly simulated, (2) the planet’s motion around the local planet-moon barycenter is taken into account, and (3) inclinations between the circumstellar orbit of the planet-moon barycenter and the planet-moon orbit are now included.

We used the emcee code² of Foreman-Mackey et al. (2013) to generate Markov chain Monte Carlo (MCMC) realizations of our planet-only model (\mathcal{M}_0) and planet-moon model (\mathcal{M}_1) and to derive posterior probability distributions of the set of model parameters (θ). We tested both a standard MCMC sampling with 100 walkers and a parallel-tempering ensemble MCMC (PTMCMC) with five temperatures, each of which has 100 walkers. As we find a better convergence rate for the PTMCMC sampling, we use it in the following. Moreover, PTMCMC can sample both the parameter space at large and in regions with tight peaks of the likelihood function. The PTMCMC sampling is allowed to walk five million steps.

The resulting model light curves are referred to as $\mathcal{F}_i(t, \theta)$, where t are the time stamps of the data points from *Kepler* and *Hubble* (N measurements in total), for which time-uncorrelated standard deviations σ_j at times t_j are assumed, following the suggestion of Teachey & Kipping (2018). This simplifies the joint probability density of the observed (and detrended) flux measurements ($F(t)$) to the product of the individual probabilities for each data point,

$$p(F|\theta, \mathcal{M}_i) = \prod_{j=1}^N \frac{1}{\sqrt{2\pi\sigma_j^2}} \exp\left(-\frac{(F(t_j) - \mathcal{F}_i(t_j, \theta))^2}{2\sigma_j^2}\right). \quad (1)$$

We then determined the set of parameters (θ_{\max}) that maximizes the joint probability density function ($p(F|\theta_{\max}, \mathcal{M}_i)$) for a given light curve $F(t_j)$ and model \mathcal{M}_i and calculated the BIC (Schwarz 1978)

$$\text{BIC}(\mathcal{M}_i|F) = m_i \ln N - 2 \ln p(F|\theta_{\max}, \mathcal{M}_i). \quad (2)$$

The advantage of the BIC in comparison to χ^2 minimization, for example, is in its relation to the number of model parameters (m_i) and data points. The more free parameters in the model, the stronger the weight of the first penalty term in Eq. (2), thereby mitigating the effects of overfitting. Details of the actual computer code implementation or transit simulations aside, this Bayesian framework is essentially what the Hunt for Exomoons

with *Kepler* survey used to identify and rank exomoon candidates (Kipping et al. 2012), which ultimately led to the detection of the exomoon candidate around Kepler-1625 b after its first detection via the orbital sampling effect (Heller 2014; Heller et al. 2016a).

2.1. Data preparation

In a first step, we used *Kepler*’s Pre-search Data Conditioning Simple Aperture Photometry (PDCSAP) and the *Hubble* Wide Field Camera 3 (WFC3) light curve as published by Teachey & Kipping (2018) based on their quadratic detrending. Then we executed our PTMCMC fitting and derived the ΔBIC values and the posterior parameter distributions.

In a second step, we did our own extraction of the *Hubble* light curve including an exponential ramp correction for each *Hubble* orbit. Then we performed the systematic trend correction together with the transit fit of a planet-moon model. Our own detrending of the light curves is not a separate step, but it is integral to the fitting procedure. For each calculation of the likelihood, we find the best fitting detrending curve by dividing the observed light curve by the transit model and by fitting a third-order polynomial to the resulting light curve. Then we remove the trend from the original light curve by dividing it through the best-fit detrending polynomial and evaluate the likelihood. We also performed a test in which the detrending parameters were free PTMCMC model parameters and found similar results for the parameter distributions but at a much higher computational cost. We note that the resulting maximum likelihood is (and must be) the same by definition if the PTMCMC sampling converges.

Kepler-1625 was observed by *Hubble* under the GO program 15149 (PI Teachey). The observations were secured from October 28 to 29, 2017, to cover the ~ 20 h transit plus several hours of out-of-transit stellar flux (Teachey & Kipping 2018). The F130N filter of WFC3 was used to obtain a single direct image of the target, while 232 spectra were acquired with the G141 grism spanning a wavelength range from 1.1 to 1.7 μm . Due to the faintness of the target, it was observed in staring mode (e.g. Berta et al. 2012; Wilkins et al. 2014) unlike the most recent observations of brighter exoplanet host stars, which were monitored in spatial scanning mode (McCullough & MacKenty 2012). Hence, instead of using the IMA files as an intermediate product, we analyzed the FLT files, which are the final output of the calwfc3 pipeline of *Hubble* and allow a finer manipulation of the exposures during consecutive non-destructive reads. Each FLT file contains measurements between about 100 and 300 electrons per second, with exposure times of about 291 s.

We used the centroid of the stellar image to calculate the wavelength calibration, adopting the relations of Pirzkal et al. (2016). For each spectroscopic frame, we first rejected the pixels flagged by calwfc3 as “bad detector pixels”, pixels with unstable response, and those with uncertain flux value (Data Quality condition 4, 32, or 512). Then we corrected each frame with the flat field file available on the Space Telescope Science Institute (STScI) website³ by following the prescription of the WFC3 online manual. We performed the background subtraction on a column-by-column basis. Due to a number of contaminant stars in the observation field (Fig. 1, top panel), we carefully selected a region on the detector that was as close as possible to the

¹ Available at www.astro.ucla.edu/~ianc/files

² Available at <http://dfm.io/emcee/current/user/pt>

³ www.stsci.edu/hst/wfc3/analysis/grism_obs/calibrations/wfc3_g141.html

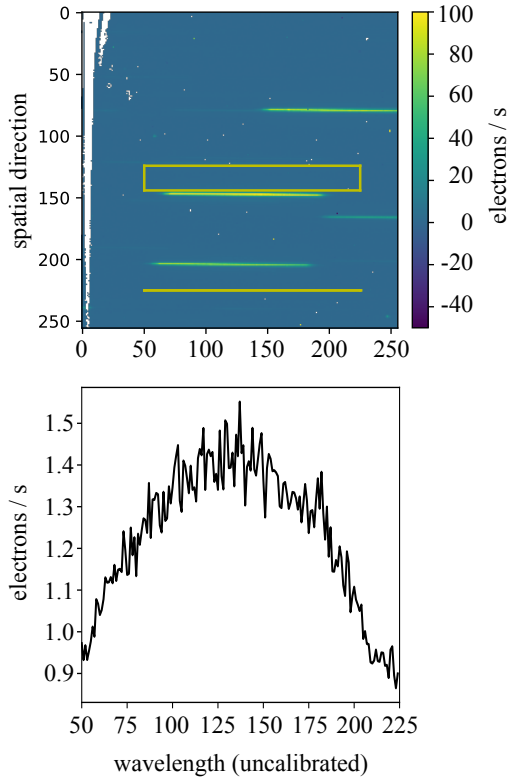


Fig. 1. *Top:* example of a WFC3 exposure of Kepler-1625. The abscissa shows the column pixel prior to wavelength calibration. The yellow box indicates the region used for background estimation. The spectrum of Kepler-1625 is at the center of the frame, around row 150 in the spatial direction, while several contaminant sources are evident in other regions of the detector. The color bar illustrates the measured charge values. *Bottom:* background value measured across the rows of the same frame.

spectrum of Kepler-1625, close to row 150 in spatial dimension, and far from any contaminant. For each column on the detector, we applied a 5σ clipping to reject the outliers and then calculated the median background flux value in that column. Following STScI prescriptions, we also removed pixels with an electron-per-second count larger than 5. An example for the background behavior is shown in the bottom panel of Fig. 1.

We inspected each frame with the image registration package (Baker & Matthews 2001) to search drifts in both axes of the detector with respect to the very last frame, and then extracted the spectrum of Kepler-1625 by performing optimal extraction (Horne 1986) on the detector rows containing the stellar flux. This procedure automatically removes bad pixels and cosmic rays from the frames by correcting them with a smoothing function. We started the extraction with an aperture of a few pixels centered on the peak of the stellar trace and gradually increased its extension by one pixel per side on the spatial direction until the flux dispersion reached a minimum.

We performed another outlier rejection by stacking all the one-dimensional spectra along the time axis. We computed a median-filtered version of the stellar flux at each wavelength bin and the median filter. Finally, we summed the stellar flux across all wavelength bins from 1.115 to 1.645 μm to obtain the band-integrated stellar flux corresponding to each exposure.

Before performing the PTMCMC optimization, we removed the first *Hubble* orbit from the data set and the first data point of each *Hubble* orbit, as they are affected by stronger instrumental effects than the other observations (Deming et al. 2013) and cannot be corrected with the same systematics model. We also removed the last point of the 12th, 13th, and 14th *Hubble* orbit since they were affected by the passage of the South Atlantic Anomaly (as highlighted in the proposal file, available on the STScI website).

2.2. Proposed unseen planet

2.2.1. Mass-orbit constraints for a close-in planet

According to Teachey et al. (2018), the 2017 *Hubble* transit of Kepler-1625 b occurred about 77.8 min earlier than predicted, an effect that could be astrophysical in nature and is referred to as a transit timing variation (TTV). As proposed by Teachey et al. (2018), this TTV could either be interpreted as evidence for an exomoon or it could indicate the presence of a hitherto unseen additional planet. Various planetary configurations can cause the observed TTV effect such as an inner planet or an outer planet. At this point, no stellar radial velocity measurements of Kepler-1625 exist that could be used to search for additional nontransiting planets in this system.

In the following, we focus on the possibility of an inner planet with a much smaller orbital period than Kepler-1625 b simply because it would have interesting observational consequences. We use the approximation of Agol et al. (2005) for the TTV amplitude (δt) due to a close inner planet, which would impose a periodic variation on the position of the star, and solve their expression for the mass of the inner planet ($M_{p,\text{in}}$) as a function of its orbital semimajor axis ($a_{p,\text{in}}$),

$$M_{p,\text{in}} = \delta t M_{\star} \frac{a_{p,\text{out}}}{a_{p,\text{in}} P_{p,\text{out}}}, \quad (3)$$

where $a_{\text{out}} = 0.87$ AU is the semimajor axis of Kepler-1625 b. The validity of this expression is restricted to coplanar systems without significant planet–planet interaction and with $a_{\text{out}} \gg a_{\text{in}}$, so that TTVs are only caused by the reflex motion of the star around its barycenter with the inner planet.

As we show in Sect. 3.2, the proposed inner planet could be a hot Jupiter. The transits of a Jupiter-sized planet, however, would be visible in the *Kepler* data. As a consequence, we can estimate the minimum orbital inclination (i) between Kepler-1625 b and the suspected planet to prevent the latter from showing transits. This angle is given as per $i = \arctan(R_{\star}/a_{p,\text{in}})$ and we use $R_{\star} = 1.793^{+0.263}_{-0.488} R_{\odot}$ (Mathur et al. 2017).

2.2.2. Orbital stability

We can exclude certain masses and orbital semimajor axes for an unseen inner planet based on the criterion of mutual Hill stability. This instability region depends to some extent on the unknown mass of Kepler-1625 b. Mass estimates can be derived from a star–planet–moon model, but these estimates are irrelevant if the observed TTVs are due to an unseen planetary perturber. Hence, we assume a nominal Jupiter mass (M_{Jup}) for Kepler-1625 b.

The Hill sphere of a planet with an orbital semimajor axis a_p around a star with mass M_{\star} can be estimated as $R_{\text{H}} = a_p (M_p / [3M_{\star}])^{1/3}$, which suggests $R_{\text{H}} = 125 R_{\text{Jup}}$ for Kepler-1625 b. We calculate the Hill radius of the proposed

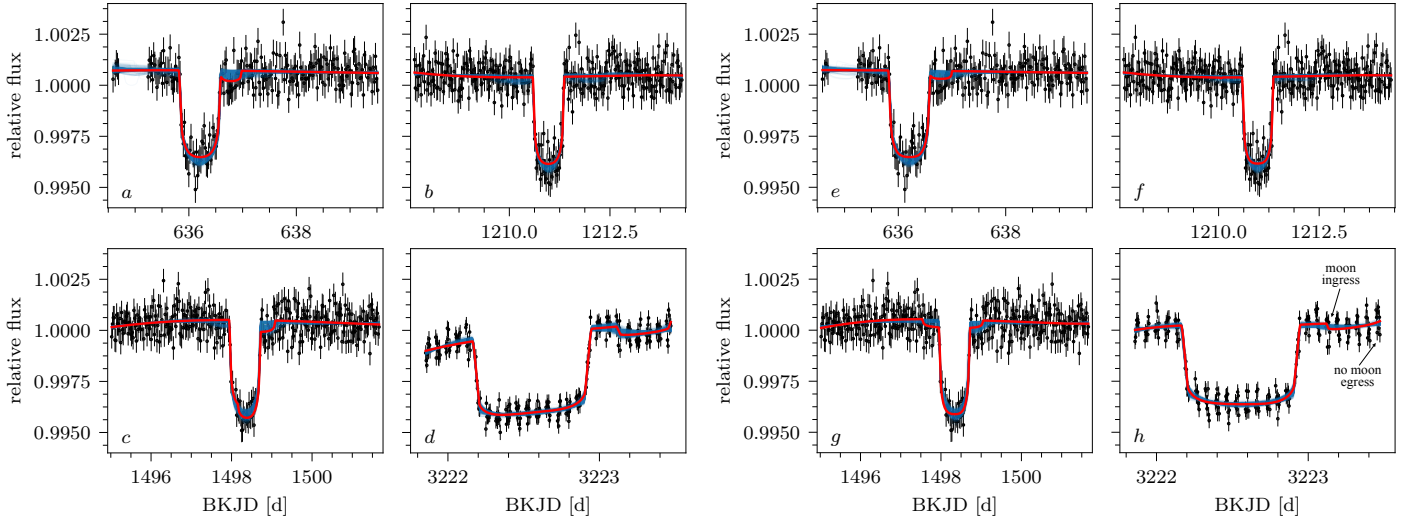


Fig. 2. Orbital solutions for Kepler-1625 b and its suspected exomoon based on the combined *Kepler* and *Hubble* data. *Panels a–c:* *Kepler* PDCSAP flux. *Panel d:* quadratic detrending of the *Hubble* data from [Teachey & Kipping \(2018\)](#). The blue curves show 1000 realizations of our PTMCMC fitting of a planet-moon model. Our most likely solution (red line) is very similar to the one found by [Teachey & Kipping \(2018\)](#), but differs significantly from the one initially found by [Teachey et al. \(2018\)](#). *Panels e–g:* *Kepler* PDCSAP flux. *Panel h:* our own detrending of the *Hubble* light curve (in parallel to the fitting). The ingress and egress of the model moon are denoted with arrows and labels in panel has an example.

inner planet accordingly, and identify the region in the mass-semimajor axis diagram of the inner planet that would lead to an overlap of the Hill spheres and therefore to orbital instability.

3. Results

3.1. PTMCMC sampling and ΔBIC

Regarding the combined data set of the *Kepler* and *Hubble* data as detrended by [Teachey & Kipping \(2018\)](#), we find a ΔBIC of -44.5 between the most likely planet-only and the most likely planet-moon solution. A combination of the *Kepler* and *Hubble* light curves based on our own extraction of the WFC3 data yields a ΔBIC of -31.0 . Formally speaking, both of these two values can be interpreted as strong statistical evidence for an exomoon interpretation. The two values are very different, however, which suggests that the detrending of the *Hubble* data has a significant effect on the exomoon interpretation. In other words, this illustrates that the systematics are not well-modeled and poorly understood.

In Figs. 2a–d, we show our results for the PTMCMC fitting of our planet-moon model to the four transits of Kepler-1625 b including the *Hubble* data as extracted and detrended by [Teachey & Kipping \(2018\)](#) using a quadratic fit. Although our most likely solution shows some resemblance to the one proposed by [Teachey & Kipping \(2018\)](#), we find that several aspects are different. As an example, the second *Kepler* transit (Fig. 2b) is fitted best without a significant photometric moon signature, that is to say, the moon does not pass in front of the stellar disk⁴, whereas the corresponding best-fit model of [Teachey & Kipping \(2018\)](#) shows a clear dip prior to the planetary transit (see their Fig. 4). What is more, most of our orbital solutions (blue lines) differ substantially from the most likely solution (red line). In other words, the orbital solutions do not converge and various planet-moon orbital configuration are compatible with the data, though with lower likelihood.

⁴ [Martin et al. \(2019\)](#) estimate that failed exomoon transits should actually be quite common for misaligned planet-moon systems, such as the one proposed by [Teachey & Kipping \(2018\)](#).

In Figs. 2e–h, we illustrate our results for the PTMCMC fitting of our planet-moon model to the four transits of Kepler-1625 b including our own extraction and detrending of the *Hubble* transit. Again, the orbital solutions (blue lines) do not converge. A comparison of panels d and h shows that the different extraction and detrending methods do have a significant effect on the individual flux measurements, in line with the findings of [Rodenbeck et al. \(2018\)](#). Although the time of the proposed exomoon transit is roughly the same in both panels, we find that the best-fit solution for the data detrended with our own reduction procedure does not contain the moon egress (panel h), whereas the best-fit solution of the data detrended by [Teachey & Kipping \(2018\)](#) does contain the moon egress (panel d). A similar fragility of this particular moon egress has been noted by [Teachey & Kipping \(2018\)](#) as they explored different detrending functions (see their Fig. 3).

Our Fig. 3 illustrates the distribution of the differential likelihood for the planet-moon model between the most likely model parameter set (θ_{max}) and the parameter sets (θ') found after five million steps of our PTMCMC fitting procedure, $p(\theta'|F, \mathcal{M}_1) - p(\theta_{\text{max}}|F, \mathcal{M}_1)$. For the combined *Kepler* and *Hubble* data detrended by [Teachey & Kipping \(2018\)](#) (left panel) and for our own *Hubble* data extraction and detrending (right panel), we find that most model solutions cluster around a differential likelihood that is very different from the most likely solution, suggesting that the most likely model is, in some sense, a statistical outlier. We initially detected this feature after approximately the first one hundred thousand PTMCMC fits. Hence, we increased the number of PTMCMC samplings to half a million and finally to five million to make sure that we sample any potentially narrow peaks of the likelihood function near the best-fit model at $p(\theta'|F, \mathcal{M}_1) - p(\theta_{\text{max}}|F, \mathcal{M}_1) = 0$ with sufficient accuracy. We find, however, that this behavior of the differential likelihood distribution clustering far from the best-fit solution persists, irrespective of the available computing power devoted to the sampling.

Figure 4 shows the posterior distributions of the moon parameters of our planet-moon model. The top panel refers to

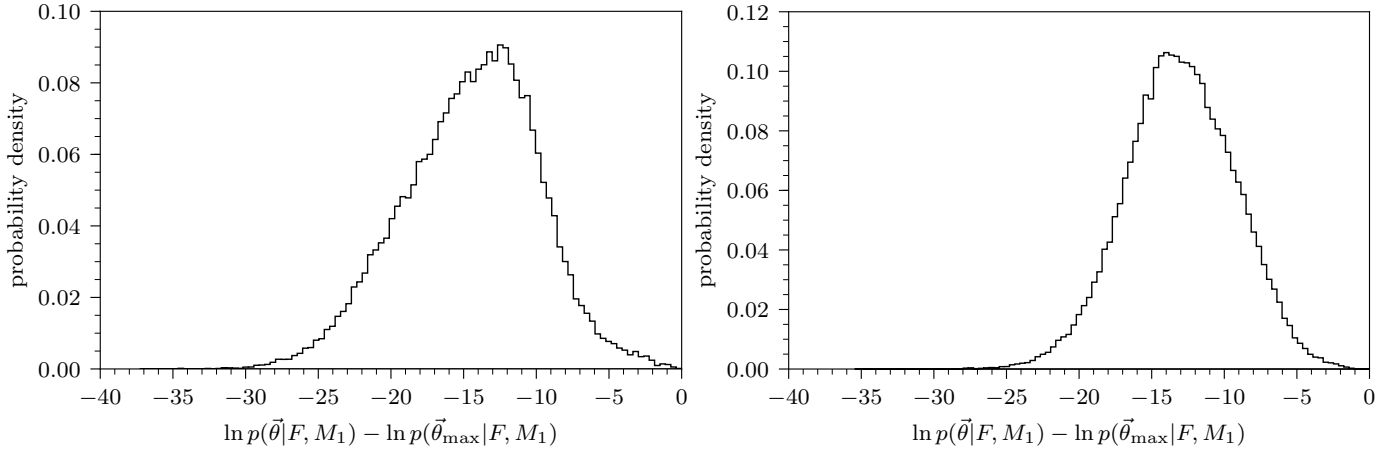


Fig. 3. Differential likelihood distribution between the most likely planet-moon model and the other solutions using 10^6 steps of our PTMCMC fitting procedure. *Left:* results from fitting our planet-moon transit model to the original data from [Teachey & Kipping \(2018\)](#). *Right:* results from fitting our planet-moon transit model to our own detrending of the *Kepler* and WFC3 data. In both panels the most likely model is located at 0 along the abscissa by definition. In both cases the models do not converge to the best-fit solution, suggesting that the best-fit solution could in fact be an outlier.

our PTMCMC fitting of the combined *Kepler* and *Hubble* data (*Hubble* data as detrended and published by [Teachey & Kipping 2018](#)), and the bottom panel shows our PTMCMC fitting of the *Kepler* data combined with our own extraction and detrending of the *Hubble* light curve. The respective median values and standard deviations are noted in the upper right corners of each subpanel and summarized in Table 1.

A comparison between the upper and lower corner plots in Fig. 4 reveals that the different detrending and fitting techniques have a significant effect on the resulting posterior distributions, in particular for i_s and Ω_s , the two angles that parameterize the orientation of the moon orbit. At the same time, however, the most likely values (red dots above the plot diagonal) and median values (blue crosses below the plot diagonal) of the seven parameters shown are well within the 1σ tolerance.

The following features can be observed in both panels of Fig. 4. The moon-to-star radius ratio (Col. 1, leftmost) shows an approximately normal distribution, whereas the scaled planet-moon orbital semimajor axis (Col. 2) shows a more complicated, skewed distribution. The solutions for the orbital period of the exomoon candidate (Col. 3) show a comb-like structure owing to the discrete number of completed moon orbits that would fit a given value of the moon’s initial orbital phase (Col. 4), which is essentially unconstrained. The moon-to-planet mass ratio (Col. 5) then shows a skewed normal distribution with a tail of large moon masses. Our results for the inclination i_s between the satellite orbit (around the planet) and the line of sight, and for the longitude of the ascending node of the moon orbit are shown in Cols. 6 and 7. The preference of i_s being either near 0 or near $\pm\pi$ (the latter is equivalent to a near-coplanar retrograde moon orbit) illustrates the well-known degeneracy of the prograde/retrograde solutions available from light curve analyses ([Lewis & Fujii 2014](#); [Heller & Albrecht 2014](#)).

3.2. Transit timing variations

Next we consider the possibility of the transits being caused by a planet only. Neglecting the *Hubble* transit, our PTMCMC sampling of the three *Kepler* transits with our planet-only transit model gives an orbital period of $P = 287.3776 \pm 0.0024$ d and an initial transit midpoint at $t_0 = 61.4528 \pm 0.0092$ d in units of

the Barycentric Kepler Julian Day (BKJD), which is equal to $\text{BJD} - 2,454,833.0$ d. The resulting transit time of the 2017 *Hubble* transit is 3222.6059 ± 0.0182 d.

Our planet-only model for the 2017 *Hubble* transit gives a transit midpoint at 3222.5547 ± 0.0014 d, which is $73.728(\pm 2.016)$ min earlier than the predicted transit midpoint. This is in agreement with the measurements of [Teachey & Kipping \(2018\)](#), who found that the *Hubble* transit occurred 77.8 min earlier than predicted. This observed early transit of Kepler-1625 b has a formal $\sim 3\sigma$ significance. We note, however, that this 3σ deviation is mostly dictated by the first transit observed with *Kepler* (see Fig. S12 in [Teachey et al. 2018](#)). We also note that this transit was preceded by a ~ 1 d observational gap in the light curve, about 0.5 d prior to the transit, which might affect the local detrending of the data and the determination of the transit mid-point of a planet-only model. Moreover, with most of the TTV effect being due to the large deviation from the linear ephemeris of the first transit, stellar (or any other systematic) variability could have a large (but unknown) effect on the error bars that go into the calculations.

In Fig. 5 we show the mass of an unseen inner planet that is required to cause the observed 73.7 min TTV amplitude from our PTMCMC fit as a function of its unknown orbital semimajor axis. The mass drops from $5.8 M_{\text{Jup}}$ at 0.03 AU to $1.8 M_{\text{Jup}}$ at 0.1 AU. Values beyond 0.1 AU cannot be assumed to fulfill the approximations made for Eq. (3) and are therefore shown with a dashed line. The actual TTV amplitude of Kepler-1625 b could even be higher than the ~ 73 min that we determined for the *Hubble* transit, and thus the mass estimates shown for a possible unknown inner planet serve as lower boundaries.

The resulting radial velocity amplitudes of the star of 923 m s^{-1} (at 0.03 AU) and 151 m s^{-1} (at 0.1 AU), respectively, are indicated along the curve. Even if the approximations for a coplanar, close-in planet were not entirely fulfilled, our results suggest that RV observations of Kepler-1625 with a high-resolution spectrograph attached to a very large (8 m class) ground-based telescope could potentially reveal an unseen planet causing the observed TTV of Kepler-1625 b. Also shown along the curve in Fig. 5 are the respective minimum orbital inclinations (rounded mean values shown) between Kepler-1625 b and the suspected close-in planet required to

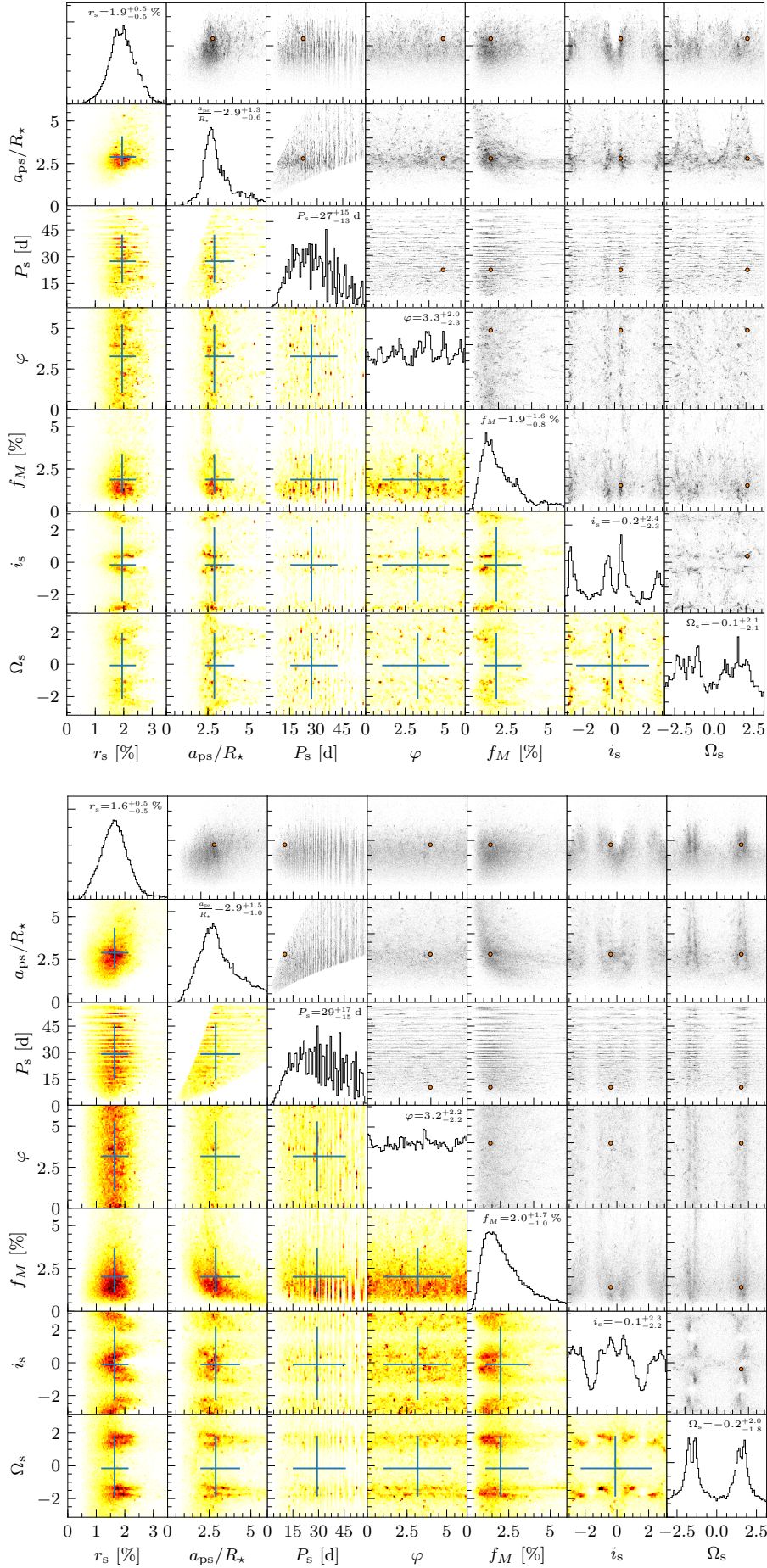


Fig. 4. Posterior distributions of a parallel tempering ensemble MCMC sampling of the combined *Kepler* and WFC3 data with our planet-moon model. *Top*: results for the original data from Teachey & Kipping (2018). *Bottom*: results for our own detrending of the *Kepler* and WFC3 data. In both figures, scatter plots are shown with black dots above the diagonal, and projected histograms are shown as colored pixels below the diagonal. The most likely parameters are denoted with an orange point in the scatter plots. Histograms of the moon-to-star radius ratio r_s , scaled semimajor axis of the planet-moon system (a_{ps}/R_*), satellite orbital period (P_s), satellite orbital phase (φ), moon-to-planet mass ratio (f_M), orbital inclination of the satellite with respect to our line of sight (i_s), and the orientation of the ascending node of the satellite orbit (Ω_s) are shown on the diagonal. Median values and standard deviations are indicated with error bars in the histograms.

Table 1. Results of our PTMCMC fitting procedure to the combined *Kepler* and *Hubble* data.

	TK18b HST photometry	Our HST photometry
r_s (%)	$1.9^{+0.5}_{-0.5}$	$1.6^{+0.5}_{-0.5}$
a_{ps} (R_\star)	$2.9^{+1.3}_{-0.6}$	$2.9^{+1.5}_{-1.0}$
P_s (d)	27^{+15}_{-13}	29^{+17}_{-15}
φ (rad)	$3.3^{+2.0}_{-2.3}$	$3.2^{+2.2}_{-2.2}$
f_M (%)	$1.9^{+1.6}_{-0.8}$	$2.0^{+1.7}_{-1.0}$
i_s (rad)	$-0.2^{+2.4}_{-2.3}$	$-0.1^{+2.3}_{-2.2}$
Ω_s (rad)	$-0.1^{+2.1}_{-2.1}$	$-0.2^{+2.0}_{-1.8}$

Notes. The *Hubble* data was either based on the photometry extracted by Teachey & Kipping (2018, TK18b; central column) or based on our own extraction (right column). Figure 4 illustrates quite clearly that the posterior distributions are not normally distributed and often not even representative of skewed normal distributions. The confidence intervals stated in this table have thus to be taken with care.

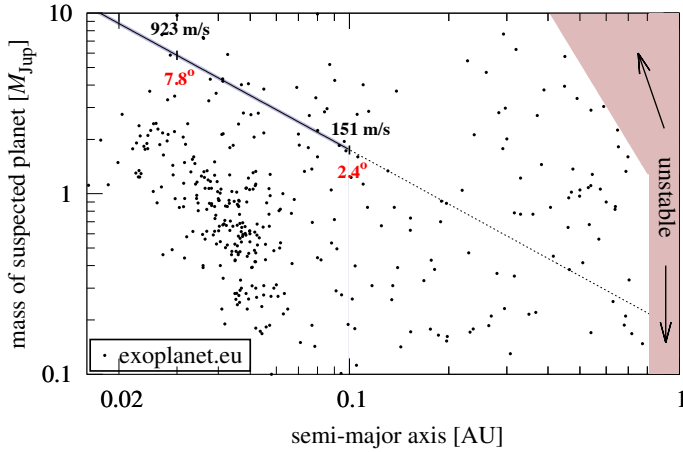


Fig. 5. Mass estimate for the potential inner planet around Kepler-1625 based on the observed TTV of 73.728 min. The thin pale blue fan around the solid curve shows the 1σ tolerance fan of ± 2.016 min. Values for semimajor axes >0.1 AU are poor approximations and thus shown with a dashed line. Black points show masses and semimajor axes of all planets from exoplanet.eu (as of 26 October 2018) around stars with masses between $0.75 M_\odot$ and $1.25 M_\odot$. A conservative estimate of a dynamically unstable region for the suspected inner planet, where its Hill sphere would touch the Hill sphere of Kepler-1625 b with an assumed mass of $1 M_{Jup}$, is shaded in pale red. RV amplitudes and minimum orbital inclination with respect to Kepler-1625 b are noted along the curve for the planetary mass estimate.

prevent Kepler-1625 b from transiting the star. The exact values are $i = 7.8^{+1.1}_{-2.0}$ degrees at 0.03 AU and $i = 2.4^{+0.3}_{-0.6}$ degrees at 0.1 AU.

The pale red shaded region is excluded from a dynamical point of view since this is where the planetary Hill spheres would overlap. The extent of this region is a conservative estimate because it assumes a mass of $1 M_{Jup}$ for Kepler-1625 b and neglects any chaotic effects induced by additional planets in the system or planet–planet cross tides etc. The true range of unstable orbits is probably larger. The black dots show all available exoplanet masses and semimajor axes from the Exoplanet Encyclopaedia, which illustrates that the suspected planet could be more massive than most of the known hot Jupiters.

4. Conclusions

With a ΔBIC of -44.5 (using published *Hubble* data of Teachey & Kipping 2018) or -31.0 (using our own *Hubble* extraction and detrending) between the most likely planet-only model and the most likely planet-moon model, we find strong statistical evidence for a roughly Neptune-sized exomoon. In both cases of the data detrending, the most likely orbital solution of the planet-moon system, however, is very different from most of the other orbital realizations of our PTMCMC modeling and the most likely solutions do not seem to converge. In other words, the most likely solution appears to be an outlier in the distribution of possible solutions and small changes to the data can have great effects on the most likely orbital solution found for the planet-moon system. As an example, we find that the two different detrending methods that we explored produce different interpretations of the transit observed with *Hubble*: in one case our PTMCMC sampling finds the egress of the moon in the light curve, in the other case it does not (Fig. 2).

Moreover, the likelihood of this best-fit orbital solution is very different from the likelihoods of most other solutions from our PTMCMC modeling. We tested both a standard MCMC sampling and a parallel-tempering MCMC (Foreman-Mackey et al. 2013); the latter is supposed to explore both the parameter space at large and the tight peaks of the likelihood function in detail. Our finding of the nonconvergence could imply that the likelihood function that best describes the data is non-Gaussian. Alternatively, with the BIC being an asymptotic criterion that requires a large sample size by definition (Stevenson et al. 2012), our findings suggest that the available data volume is simply too small for the BIC to be formally applicable. We conclude that the ΔBIC is an unreliable metric for an exomoon detection for this data set of only four transits and possibly for other data sets of *Kepler* as well.

One solution to evaluating whether the BIC or an alternative information criterion such as the Akaike information criterion (AIC; Akaike 1974) or the deviance information criterion (DIC; Spiegelhalter et al. 2002) is more suitable for assessing the likelihoods of a planet-only model and of a planet-moon model could be injection-retrieval experiments of synthetic transits (Heller et al. 2016b; Rodenbeck et al. 2018). Such an analysis, however, goes beyond the scope of this paper.

We also observe the TTV effect discovered by Teachey & Kipping (2018). If the early arrival of Kepler-1625 b for its late-2017 transit was caused by an inner planet rather than by an exomoon, then the planet would be a super-Jovian mass hot Jupiter, the exact mass limit depending on the assumed orbital semimajor axis. For example, the resulting stellar radial velocity amplitude would be about 900 m s^{-1} for a $5.8 M_{Jup}$ planet at 0.03 AU and about 150 m s^{-1} for a $1.8 M_{Jup}$ planet at 0.1 AU. From the absence of a transit signature of this hypothetical planet in the four years of *Kepler* data, we conclude that it would need to have an orbital inclination of at least $i = 7.8^{+1.1}_{-2.0}$ (if it were at 0.03 AU) or $i = 2.4^{+0.3}_{-0.6}$ degrees (if it were at 0.1 AU). If its inclination is not close to 90° , at which point its effect on the stellar RV amplitude would vanish, then the hypothesis of an unseen inner planet causing the Kepler-1625 TTV could be observationally testable.

Ground-based photometric observations are hardly practicable to answer the question of this exomoon candidate because continuous in- and near-transit monitoring of the target is required over at least two days. Current and near-future space-based exoplanet missions, on the other hand, will likely

not be able to deliver the signal-to-noise ratios (S/N) required to validate or reject the exomoon hypothesis. With a *Gaia* *G*-band magnitude of $m_G = 15.76$ (Gaia Collaboration 2016, 2018) the star is rather faint in the visible regime of the electromagnetic spectrum and the possible moon transits are therefore beyond the sensitivity limits of the TESS, CHEOPS, and PLATO missions. 2MASS observations suggest that Kepler-1625 is somewhat brighter in the near-infrared (Cutri et al. 2003), such that the *James Webb* Space Telescope (launch currently scheduled for early 2021) should be able to detect the transit of the proposed Neptune-sized moon, for example via photometric time series obtained with the NIRC*am* imaging instrument.

All things combined, the fragility of the proposed photometric exomoon signature with respect to the detrending methods, the unknown systematics in both the *Kepler* and the *Hubble* data, the absence of a proper assessment of the stellar variability of Kepler-1625, the faintness of the star (and the resulting photometric noise floor), the previously stated coincidence of the proposed moon's properties with those of false positives (Rodenbeck et al. 2018), the existence of at least one plausible alternative explanation for the observed TTV effect of Kepler-1625 b, and the serious doubts that we have about the ΔBIC as a reliable metric at least for this particular data set lead us to conclude that the proposed moon around Kepler-1625 b might not be real. We find that the exomoon hypothesis heavily relies on a chain of delicate assumptions, all of which need to be further investigated.

A similar point was raised by Teachey & Kipping (2018), and our analysis is an independent attempt to shed some light on the “unknown unknowns” referred to by the authors. For the time being, we take the position that the first exomoon has yet to be detected as the likelihood of an exomoon around Kepler-1625 b cannot be assessed with the methods used and data currently available.

Acknowledgements. The authors thank Kevin Stevenson, Hannah Wakeford and Megan Sosey for their help with the data analysis, and Nikole Lewis for the feedback on the manuscript. The authors would also like to thank the referee for a challenging and constructive report. This work was supported in part by the German space agency (Deutsches Zentrum für Luft- und Raumfahrt) under

PLATO Data Center grant 50001501. This work made use of NASA's ADS Bibliographic Services and of the Exoplanet Encyclopaedia (<http://exoplanet.eu>). RH wrote the manuscript, proposed Figs. 2–4, generated Fig. 5, and guided the work. K.R. derived the star-planet-moon orbital simulations and the respective statistics and generated Figs. 2–4. G.B. performed the light curve extraction from the WFC3 *Hubble* data and generated Fig. 1. All authors contributed equally to the interpretation of the data.

References

- Agol, E., Steffen, J., Sari, R., & Clarkson, W. 2005, *MNRAS*, **359**, 567
- Akaike, H. 1974, *IEEE Trans. Automat. Contr.*, **19**, 716
- Baker, S., & Matthews, I. 2001, *Proceedings of the 2001 IEEE Conference on Computer Vision and Pattern Recognition*, **1**, 1090
- Berta, Z. K., Charbonneau, D., Désert, J.-M., et al. 2012, *ApJ*, **747**, 35
- Cutri, R. M., Skrutskie, M. F., van Dyk, S., et al. 2003, *VizieR Online Data Catalog*: **II/246**
- Deming, D., Wilkins, A., McCullough, P., et al. 2013, *ApJ*, **774**, 95
- Foreman-Mackey, D., Hogg, D., Lang, D., & Goodman, J. 2013, *PASP*, **125**, 306
- Gaia Collaboration (Prusti, T., et al.) 2016, *A&A*, **595**, A1
- Gaia Collaboration (Brown, A. G. A., et al.) 2018, *A&A*, **616**, A1
- Heller, R. 2014, *ApJ*, **787**, 14
- Heller, R. 2018, *A&A*, **610**, A39
- Heller, R., & Albrecht, S. 2014, *ApJ*, **796**, L1
- Heller, R., Hippke, M., & Jackson, B. 2016a, *ApJ*, **820**, 88
- Heller, R., Hippke, M., Placek, B., Angerhausen, D., & Agol, E. 2016b, *A&A*, **591**, A67
- Horne, K. 1986, *PASP*, **98**, 609
- Kipping, D. M., Bakos, G. Á., Buchhave, L., Nesvorný, D., & Schmitt, A. 2012, *ApJ*, **750**, 115
- Lewis, K. M., & Fujii, Y. 2014, *ApJ*, **791**, L26
- Mandel, K., & Agol, E. 2002, *ApJ*, **580**, L171
- Martin, D. V., Fabrycky, D. C., & Montet, B. T. 2019, *ApJL*, submitted [arXiv:1901.06366]
- Mathur, S., Huber, D., Batalha, N. M., et al. 2017, *ApJS*, **229**, 30
- McCullough, P., & MacKenty, J. 2012, Considerations for using Spatial Scans with WFC3, Tech. rep
- Pirzkal, N., Ryan, R., & Brammer, G. 2016, Trace and Wavelength Calibrations of the WFC3 G102 and G141 IR Grisms, Tech. rep
- Rodenbeck, K., Heller, R., Hippke, M., & Gizon, L. 2018, *A&A*, **617**, A49
- Schwarz, G. 1978, *Ann. Stat.*, **6**, 461
- Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & Van Der Linde, A. 2002, *J. R. Stat. Soc. Ser B Stat. Methodol.*, **64**, 583
- Stevenson, K. B., Harrington, J., Fortney, J. J., et al. 2012, *ApJ*, **754**, 136
- Teachey, A., & Kipping, D. M. 2018, *Sci. Adv.*, **4**, eaav1784
- Teachey, A., Kipping, D. M., & Schmitt, A. R. 2018, *AJ*, **155**, 36
- Wilkins, A. N., Deming, D., Madhusudhan, N., et al. 2014, *ApJ*, **783**, 113