This project summary will go in NSPIRES. Put here for proofing

Exploring the Critical Radius Between mini-Neptunes and super-Earths

We propose to combine a Bayesian reanalysis of short-period Kepler exoplanet transits with standard models of tidal theory in order to identify the planetary radius that separates rocky and gaseous exoplanets. We exploit the conventional assumption that gaseous planets dissipate orders of magnitude less tidal energy than rocky planets, leading to the expectation that the latter will be on circular orbits out to larger orbital periods. Preliminary dynamical simulations show that short period (2-10 days) gaseous bodies tend to be found with eccentricities near their primordial value, but rocky bodies are preferentially found at low eccentricity due to tidal circularization. Thus, a study of the eccentricities of short-period planets can constrain the planetary radius of this transition. The identification of the boundary between rocky and gaseous bodies, independently from mass measurements, will be vital for Kepler's long-term goal of discovering a habitable planet around a solar analogue.

A lower limit to the orbital eccentricity can be calculated by comparing the difference between the modeled transit duration and the transit duration that would be seen if the orbit were circular. To assess this difference, we analyze Kepler lightcurves using a purely geometric model that includes no assumptions about the orbital dynamics. We cast our measurement of minimum eccentricity in terms of two model parameters, whose posterior distributions we explore using Markov Chain Monte Carlo methods, and one physical parameter that must be estimated from other means. We have run a suite of simulations using a grid in transit depth, stellar brightness (lightcurve signal-to-noise), and the number of transits included in the model to gauge our sensitivity to these model parameters. We validated this method on the confirmed exoplanet system Kepler 62-b, and successfully recovered the published results and expected parameter uncertainties. We propose here to extend this analysis to an ensemble of 890+ KOIs that have been selected based upon their Kepler-reported periods and planetary radius. This reanalysis will enable the first measurements of the boundary between gaseous and rocky exoplanets, as well as of tidal dissipation as a function of planetary radius.

This project spans the fields of high-performance computation, statistical modeling of experimental data, celestial mechanics, and tidal dynamics, which will make it a valuable contribution to the field of exoplanet studies. We will release code and data using open-source collaboration tools, and help to guide the adoption of reproducible research standards by releasing interactive analysis packages as part of our publication process.

I. Introduction

The discovery of an inhabited planet is a primary goal of exoplanetary science. The *Kepler* spacecraft has now found several small candidates in potentially habitable orbits, (e.g. Kepler–62 f; Borucki et al., 2013), and radial velocity surveys are also detecting apparently low–mass planets in habitable zones (HZs), e.g. Gl 667C c (Anglada-Escudé et al., 2012). While the location of an orbit with respect to the HZ is an important first cut on habitability, the composition of the planet is at least as important. Life as we understand it cannot survive on gaseous planets, yet we do not know the mass and/or radius that separates gaseous from terrestrial planets. In particular, the identification of the critical radius between the two, R_{crit}, would provide crucial information for *Kepler* 's primary mission to discover a terrestrial planet in the HZ of a G dwarf.

The primary goal of this proposal is to determine this critical planetary radius between rocky and gaseous bodies using *Kepler* data. We will exploit the expected discrepancy in tidal dissipation of gaseous and rocky bodies to determine the largest orbital periods at which the two classes of bodies circularize. We will use the so–called transit duration deviation, the difference between the observed transit duration and that of a circular orbit, to estimate the minimum eccentricity permitted from the transit data. We will also include the effects of additional planetary companions and atmospheric mass loss, as they also influence the evolution of eccentricity. In order to successfully constrain these theoretical results from *Kepler* data, we require robust characterization of each candidate, and hence we will perform state–of–the–art modeling of all relevant *Kepler* transits. To optimize our sensitivity to these effects, we will examine only those systems with short periods (less than 10 days) and having planetary radii near the anticipated boundary (less than 4 $R_\oplus$). Our proposed research offers the best route to determine which *Kepler* candidates are rocky, independent of mass measurements, which is vital information in *Kepler* 's quest to discover potentially habitable planets orbiting G dwarfs. In summary:

> **We propose to use the observed minimum eccentricity and theory of tidal dynamics to determine the critical radius between rocky and gaseous exoplanets R_{crit}, constrain their tidal quality factors (Q_r and Q_g, respectively), and understand the efficiency of hydrogen loss for close–in, small, gaseous exoplanets, using *Kepler* lightcurves.**

II. Objectives and Significance

Tidal Theory

Tidal dissipation is celestial bodies in extremely challenging to measure (Goldreich & Soter, 1966; Hut, 1981; Aksnes & Franklin, 2001; **?**; Jackson et al., 2009; Lainey et al., 2012) due to 1) the dearth of known worlds in highly dissipative configurations, 2) the long timescales involved (Gyrs), and 3) the intractability of derivations based on first principles. However, the *Kepler* space telescope has now discovered thousands of exoplanet candidates (Batalha et al., 2013), of which about 1000 orbit FGK stars with orbital periods less than 10 days, and that may experience significant tidal evolution (Rasio et al., 1996; **?**; Matsumura et al., 2010).

In the equilibrium tide model (Darwin, 1880; MacDonald, 1964; Goldreich & Soter, 1966;

Hut, 1981; Ferraz-Mello et al., 2008; Leconte et al., 2010), the figure of a tidally deformed body is a superposition of surface waves with different frequencies. The sum of these waves corresponds to the tidally–deformed figure, and allows for the relatively simple derivation of the time rates of change of orbital and spin properties. While two qualitatively different models have emerged, the constant–phase–lag (CPL) and constant–time–lag (CTL) models (Greenberg, 2009), both rely on this assumption of superposition, and neither has been rejected observationally. Both models make a critical prediction that we will exploit in this proposal: The tidal dissipation in rocky planets is orders of magnitude larger than in gaseous bodies. This disparity implies that rocky planets will evolve much more rapidly than gaseous bodies and we expect rocky exoplanets to be tidally circularized on larger orbits than for gaseous bodies. As we show below, the canonical values for Q_g of 10^6 and Q_r of 100 should be measurable in the *Kepler* field, if the transit data and stellar properties can be known to sufficient accuracy. The key is to recognize that gaseous bodies will tidally circularize more slowly than rocky planets, and may still retain non–zero eccentricities after Gyrs. While transit data cannot measure the eccentricity (Barnes, 2007), they *can* provide a lower limit, which we outline below.

The Transit Duration Deviation (TDD)

The transit duration is the time required for a planet to traverse the disk of its parent star, and to first order is:

$$T = \frac{2\sqrt{R_*^2 - b^2}}{v}, \tag{1}$$

where R_* is the radius of the star, b is the impact parameter, and v is the instantaneous velocity of the planet. On a circular orbit, v is constant and we expect

$$T_c = \frac{\sqrt{R_*^2 - b^2}}{\pi a} P, \tag{2}$$

where P is the orbital period. However, for an eccentric orbit the orbital velocity is a function of longitude (Kepler's 3rd Law), and is given by

$$v(\theta) = \frac{2\pi a}{P} \sqrt{\frac{1 + 2e\cos(\theta) + e^2}{1 - e^2}}, \tag{3}$$

where e is the eccentricity and θ is the true anomaly, the angle between the longitude of pericenter and the actual position of the planet in its orbit. From transit data alone, the value of θ is unknown, and hence so is e.

However, we can exploit the difference between T and T_c to obtain a minimum value of the eccentricity, e_{min} (Barnes, 2007). The situation is somewhat complicated because T can be larger or smaller than T_c depending on θ. If the planet is close to apoapse, $T > T_c$, while at periapse $T < T_c$. To derive e_{min} we must assume that $\theta = 0$ or π. While the velocity could be larger at some other position in the orbit, we know that the maximum deviation from the circular velocity is at least as large as the measured velocity, and hence e must be

at least a certain value. If we define the transit duration deviation, Δ, as

$$\Delta = \begin{cases} \dfrac{T - T_c}{T_c} & T_c < T \\ \dfrac{T_c - T}{T_c} & T_c > T, \end{cases} \tag{4}$$

and then for ease of notation,

$$\Delta' \equiv \begin{cases} \Delta + 1 & T_c < T \\ \Delta - 1 & T_c > T, \end{cases} \tag{5}$$

then we find that

$$e_{min} = \begin{cases} \dfrac{\Delta'^2 - 1}{\Delta'^2 + 1} & T_c < T \\ \dfrac{1 - \Delta'^2}{1 + \Delta'^2} & T_c > T \end{cases} \tag{6}$$

is the minimum eccentricity permitted by transit data.

The Boundary Between Rocky and Gaseous Planets

That transit data provide a minimum eccentricity, while tidal theory damps eccentricity to zero, is crucial for our proposed research. If $e_{min} = 0$, *then the orbit is circular.* Of course circular orbits could be primordial, but Jackson et al. (2008) and Matsumura et al. (2010) showed that the observed radial–velocity–detected planets in tight orbits could have formed with eccentricities consistent with the more distant planets and were subsequently tidally damped in both the CPL and CTL framework. If $e_{min} > 0$ then tides have not tidally damped the eccentricity, and, if we know the age of the system, then we can estimate the tidal Q (or in the CTL model, the time lag factor).

As an example consider the two curves in Fig. 1. The line shows 10 Gyr of tidal evolution of a 2 $R_\oplus$planet with a density of 1 g/cm^3 and tidal Q of 10^6 (i.e. a 3.8 $M_\oplus$"mini-Neptune"), while the filled circles represent the orbit of a 2 $R_\oplus$planet with a mass of 10 $M_\oplus$and a tidal Q of 100 (i.e. a "super-Earth") every 100 Myr. Both objects start with the same orbital initial conditions. The super–Earth circularizes in about 1 Gyr; the mini–Neptune does not evolve significantly, even after 10 Gyr. This discrepancy is evident despite the fact that equilibrium tidal models predict that evolution scales as mass to the 3/2 power and radius to the fifth power – instead, the large difference between the Qs dominates. We therefore hypothesize that the TDD may be able to identify the radius that separates gaseous planets from rocky planets.

To test this possibility, we performed the following test. We created 25,000 synthetic star–planet configurations with initial semi–major axes uniformly in the range [0.01,0.1] AU, radii in the range [0.5,10] $R_\oplus$, stellar masses in the range [0.8,1.2] $M_\odot$, and ages in the range [2,8] Gyr. If the radius is less than 2 $R_\oplus$, we scale the mass as $(R/R_\oplus)^{3.68}M_\oplus$(Sotin et al., 2007) and we assign a tidal Q in the range [30,300]. If larger, then we assume the density is 1 g/cm^3, and a tidal Q in the range [10^6,10^7]. The initial eccentricity is drawn from the currently observed distribution of distant planets ($a > 0.2$ AU). We then integrate the

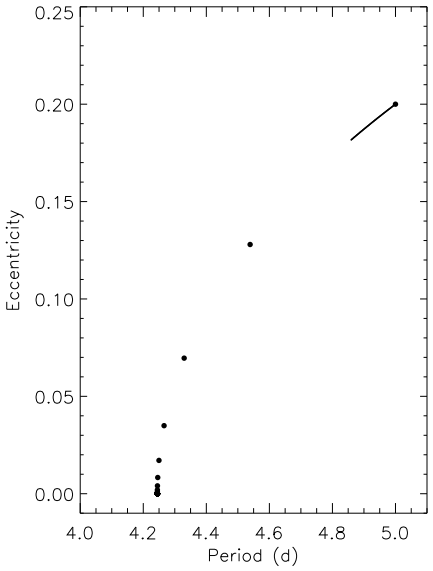

Figure 1: Comparison of the tidal evolution of a 2 $R_\oplus$mini–Neptune (solid line; for 10 Gyr) and a 2 $R_\oplus$super–Earth (circles; in 100 Myr intervals). The mini–Neptune experiences little orbital evolution, but the super–Earth circularizes in about 1 Gyr. This discrepancy is due to the 4 orders of magnitude difference in tidal dissipation between gaseous and rocky planets.

system forward for the randomly chosen age and assume we observe the system in that final configuration. In Fig. 2, we show the resulting average eccentricities of these planes as a function of planetary radius, R_p, and orbital period, P. The small values of e at low R_p and P shows this effect. Furthermore, we can see the features that correspond directly to three parameters that are currently very poorly constrained: R_{crit} via the rapid rise in $\langle e \rangle$ at 2 $R_\oplus$; Q_g via the rapid rise in $\langle e \rangle$ at 1 day above 2 $R_\oplus$; and Q_r via the rapid rise at 4 days and below 2 $R_\oplus$. Thus in this simple model, we see three constraints for three unknowns, yielding the possibility that the right observations may be able to provide values for these elusive quantities.

However, *Kepler* data do not provide eccentricity, but rather the minimum eccentricity. We must therefore transform these simulations into a form that is directly comparable to an observable, e.g. e_{min}. In order to calculate e_{min}, we choose a random value for θ and calculate the velocity according to Eq. 3. We calculate the average minimum eccentricity, $\langle e_{min} \rangle$ in 1 $R_\oplus$ planetary radius bins and 1 day orbital period bins and plot $\langle e_{min} \rangle$ as a function of orbital period for different radii as solid lines in Fig. 3. For $R < 2$ $R_\oplus$, $\langle e_{min} \rangle \sim 0$ up to about a 4 day period. However, for larger radii, circular orbits are only guaranteed for periods less than about 2 days. *Despite the order of magnitude ranges for each physical planetary property, the disparity in tidal Q's produces a strong signal in* $\langle e_{min} \rangle$ *that distinguishes the rocky and gaseous planets.*

This pilot study is encouraging, but its feasibility rests on the precision of the models of the *Kepler* data. Specifically, impact parameters, stellar and planetary radii, orbital period, and stellar mass must be known and their uncertainties well–modeled. The first four properties are measurable from transit data alone, while the fifth must be estimated by other means. The *Kepler* team has provided these data in various publications and websites. The solid squares show the values of $\langle e_{min} \rangle$ with quantities from the *Kepler* Planet Candidate Data Explorer[1]. Nearly all the observed data are above the predictions. There are two

[1]http://planetquest.jpl.nasa.gov/kepler

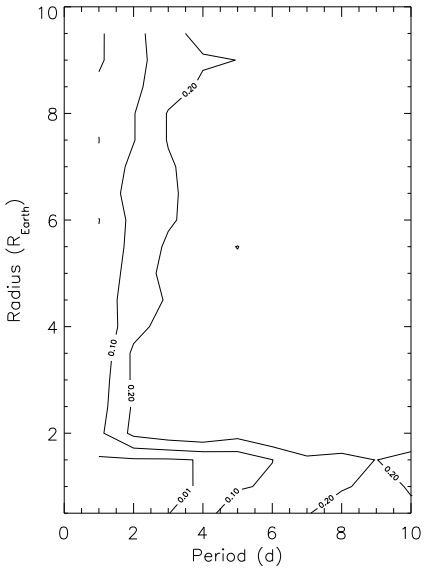

Figure 2: Average final eccentricity of a suite of 25,000 systems of one star and one planet. Planets larger than 2 $R_\oplus$ are gaseous, and those smaller are rocky. The bin-sizes are 0.5 $R_\oplus$in radius and 0.5 d in period. Gaseous planets retain a residual eccentricity if $P > 1.5$ d, while rocky planets require $P > 4$ d. The transition occurs at 2 $R_\oplus$, which in this case is R_{crit}.

possible explanations for this discrepancy: 1) The theory is wrong, or 2) the reported model parameters are of poor quality. We outline limits of the theory below.

Non–Tidal Effects

First, we note that additional companions can pump eccentricity through mutual gravitational interactions, even if tidal damping is ongoing (Mardling & Lin, 2002; Bolmont et al., 2013). Therefore we must be cautious when interpreting Fig. 3, as additional companions, both seen and unseen, can maintain non–zero eccentricities. However, there are limits: Bolmont et al. (2013) showed that planet–planet interactions cannot maintain the eccentricity of 55 Cnc e above 0.1. That system is particularly relevant as there are many close–in planets orbiting G dwarfs. Therefore, we can conclude that discrepancy between the observed and simulated systems is not due to planet–planet interactions.

Another possibility is that stellar winds and activity can strip an atmosphere, reducing the mass and radius, and potentially changing the planet from a mini–Neptune to a super–Earth (Jackson et al., 2010; Valencia et al., 2010; Leitzinger et al., 2011; Poppenhaeger et al., 2012). Recently, Owen & Wu (2013) argued that the *Kepler* sample is consistent with hydrodyanmic mass loss, and that some low–mass planets could have formed with substantially more mass. Mass loss should decrease the time to circularize the orbit, assuming the radius doesn't become very large, which is unlikely after about 100 Myr (Lopez et al., 2012). Therefore, mass loss could stall circularization for mini–Neptunes, but not for super–Earths. Although few radial velocity measurements exist, planets with radii less than ~ 1.5 $R_\oplus$ have densities consistent with silicate compositions (Batalha et al., 2011). Thus, mass loss seems unlikely to explain the discrepancy seen for the smallest candidates in the *Kepler* field.

Other effects, such as stellar mass loss or the galactic tide will be negligible, but to properly treat the problem, planetary mass loss and planet–planet perturbations must be considered. On the theoretical side, the path forward to determine R_{crit}, Q_r and Q_g is clear: We must first model the full range of plausible values for these three parameters to

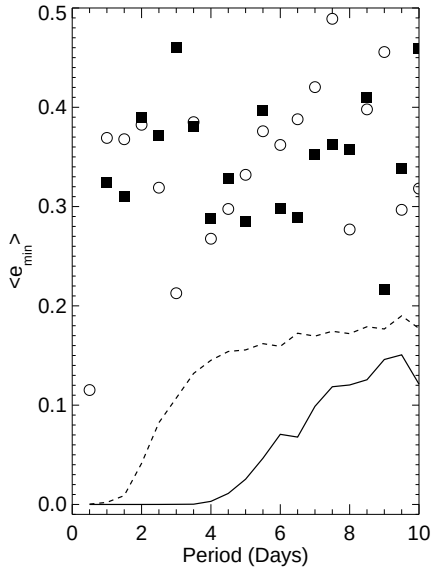

Figure 3: Average minimum eccentricities for transiting exoplanets as a function of period and radius. The solid curve and filled squares represent planets with radii below the selected critical radius of 2 $R_\oplus$, while the dashed curve and open circles are larger planets. The lines are the relationships for the simulated data set, symbols for KOIs. The latter all have values near 0.3–0.4 days regardless of period, whereas model gaseous planets have non–zero eccentricities if the period is larger than 1.5 days, and rocky planets at larger than 4. If tidal dissipation is a function of radius, it should be detectable. However, tidal circularization appears not to be present in the *Kepler* data, in sharp contradiction with the radial velocity exoplanet sample. Publicly available fits to *Kepler* data are therefore inadequate to identify R_{crit} and tidal Qs, and a state–of–the–art statistical re–analysis of *Kepler* photometry is required to determine these fundamental parameters.

calculate $\langle e_{min} \rangle (R_p, P)$. We must then include the planet–planet interactions of multiple planet systems over the lifetimes of the systems. Finally, we must incorporate mass loss, including the possibility that a mini–Neptune can become a super–Earth with its associated change in tidal Q.

III. Technical Approach and Methodology

We describe below our technical approach to the re–analysis of Kepler data, including simulations to understand how we will be able to constrain system parameters and validation of this technique using a known exoplanet system, and how we will fold these results into tidal theory to understand R_{crit}, Q_r and Q_g.

Transit Model

To avoid a dependency between the fitted model and a physical model that includes orbital dynamics, we parameterize the lightcurve model in purely geometric terms. To do so, we adopt the quadratic limb–darkened model of Mandel & Agol (2002), which describes transit lightcurves in terms of two (nuisance) limb–darkening coefficients and two (important) system parameters. The first of the system parameters is the planetary radius divided by the stellar radius ($\zeta \equiv Rp/R_*$), which determines the fractional area of the stellar disk that may be occulted by the planet. The second is the impact parameter of the planet ($\beta \equiv b/R_*$). This variable is a function of time, due to the objects' relative motion. This function is dependent on the chord that the planet takes across the stellar disk, itself typically estimated using the orbital parameters semi–major axis ($\alpha \equiv a/R_*$) and inclination.

Instead we use here a purely geometric parameterization, represented in Figure 4. We describe the impact parameter as a function of time using the minimum impact parameter β_0

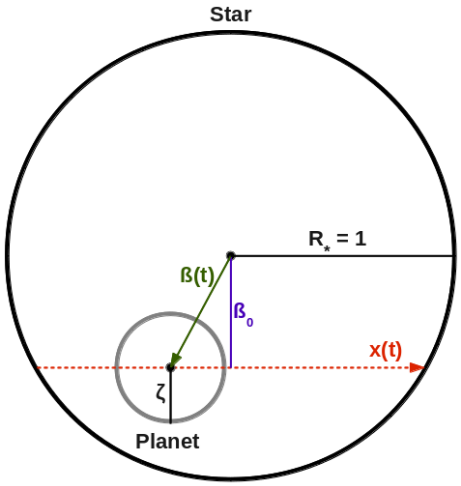

Figure 4: Schematic detailing our geometric model for the impact parameter as a function of time, $\beta(t)$. Relevant variables include the minimum impact parameter scaled by the stellar radius β_0, the radius of the planet scaled by the stellar radius ζ, and the time-dependent position of the planet $x(t)$. The fitted parameters are t_0 (defined where $\beta(t_0) = \beta_0$), the minimum impact parameter β_0, scaled planetary radius ζ, and τ (which represents the time for the planet to travel the angular distance subtended by the stellar radius).

– when the centers of the sources are aligned along the x–axis at center–of–transit time t_0 – and the location of the planet on the transit chord across the stellar disk. The x–coordinate of the planet as a function of time is represented as $x(t)/R_* = (t - t_0) * v_\perp/R_* = (t - t_0)/\tau$, where $v_\perp$ is the (unknown) perpendicular velocity, and τ is the (fitted) amount of time it takes the planet to traverse a distance equal to the stellar radius assuming no acceleration. This allows us to express geometrically the impact parameter as a function of time:

$$\beta(t) \quad = \quad \sqrt{\beta_0^2 + ((t - t_0)/\tau)^2}, \tag{7}$$

which is then used along with ζ to generate a model transit lightcurve. This model yields a 4–parameter fit to each transit: $t_0, \beta_0^2, \tau, \zeta$. The system period P may be determined using the multiple (N) transits and the ensemble of $t_{0;i=1...N}$. The transit duration T may be found from the 2 solutions to $\beta(t) = 1$, and represents the time between the center of the planet crossing each limb of the star:

$$T \quad = \quad 2 * \tau\sqrt{1 - \beta_0^2}. \tag{8}$$

Combining Equations 4–6 with Equation 8, we express e_{min} in terms of our model parameters:

$$e_{min} \quad = \quad \left|\frac{P^2 - 4\pi^2\alpha^2\tau^2}{P^2 + 4\pi^2\alpha^2\tau^2}\right|. \tag{9}$$

All factors of the minimum impact parameter β_0 cancel out, which is fortuitous as this is typically the least constrained of all system parameters, especially for the long–cadence *Kepler* data. While the uncertainty in β_0 will affect our knowledge of the other system parameters, this uncertainty may be marginalized over by examining the posterior distributions, which drives us to use Markov–Chain Monte Carlo (MCMC) modeling (described below).

Equation 9 indicates that e_{min} is purely a function of the fitted parameter τ, the derived period P, and an externally estimated semi–major axis for the planet (in units of the stellar radius) α. The uncertainty in e_{min} scales as below for all 3 parameters:

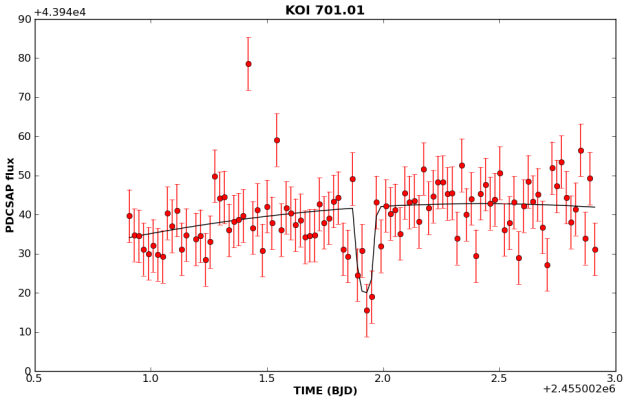

Figure 5: An example of our local detrending algorithm, applied to the first transit of KOI 701.01. The data and model (solid line) are in PDCSAP flux units. The solid curve is generated by first dividing the data by the transit model, applying a low–order spline to the normalized data, and then taking the product of the spline and the model.

$$\begin{aligned} \Xi &\equiv \frac{16\pi^2\alpha^2\tau^2P^2}{16\pi^4\alpha^4\tau^4 - P^4} \\ \frac{\delta e_{min}}{\delta\xi} &= \frac{\Xi\, e_{min}}{\xi} \qquad [\xi = \tau; P; \alpha]\,. \end{aligned} \tag{10}$$

In the limit of a circular orbit ($P \to 2\pi\alpha\tau$), $\Xi\, e_{min} = 1$.

KOI 701.01

We validate our proposed methodology by analyzing *Kepler* data from KOI 701.01 (Kepler 62–b; Borucki et al., 2013). This planet has a period of 5.715 days, $\zeta = 0.018$ ($R_p \sim 1.3$ R_E), and a transit depth of 4×10^{-4} %. We use the limb darkening parameters for the host star from Sing (2010). The *Kepler* data have correlated (red) noise which we must account for before model fitting. To do so, we perform a local detrending by first dividing the data by the proposed model, and fitting a low order spline to the result. The goodness of fit is determined by comparing the product of the spline and the model to the data. Figure 5 provides an example model fit determined in this manner. The data used are the PDCSAP fluxes and uncertainties. We were able to so model the first 32 transits for this proposal.

To examine how our knowledge of system parameters evolves as a function of number of transits, we have fit *all* the data up to the time of each transit, for each of $N = 32$ transits. This means that for transit $n \leq N$, we have common model parameters β_0^2, τ, ζ and per–transit parameters $t_{0;i=1..n}$, for a total of $n + 3$ model parameters. This yields an ensemble of N system models, each incorporating one more transit than the previous one.

We used the affine–invariant MCMC sampler `emcee` (Foreman-Mackey et al., 2013) to sample the posterior distribution of the model parameters. This program uses the method of Goodman & Weare (2010) to achieve high sampling performance independent of the aspect ratio of the posterior distribution, meaning covariances between parameters are less important to the efficacy of the MCMC sampling. This provided a set of MCMC chains that we examine to determine our constraints on the fitted parameters. We used the Gelman–Rubin $\hat{R}$–static (Gelman & Rubin, 1992) to assure that each chain sufficiently samples model space, and required effective chain lengths larger than 10^4 to ensure sufficient mixing in the MCMC sample (e.g. Tegmark et al., 2004). Our trial runs using KOI 701.01 indicated that

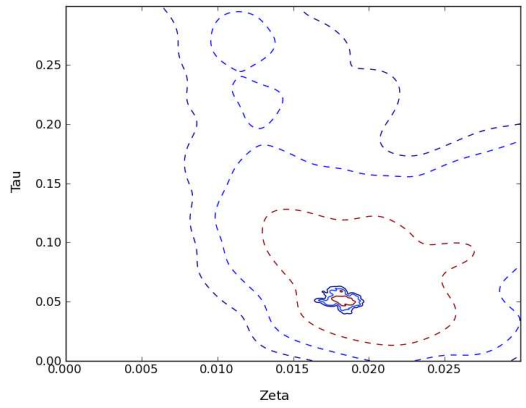

Figure 6: The joint distribution of τ vs. ζ for KOI 701.01. The *dashed* lines show the distribution when fitting 1 transit, with contours at the 68.3%, 95.5%, and 99.7% confidence limits. The *solid* lines show the distribution when fitting 32 transits simultaneously.

our chains typically have autocorrelation lengths of ~ 100, requiring a total number of steps per chain of 10^6. We used burn–in times having 10% the requested number of steps, which are then discarded before the final chain commences.

For each transit, we evaluated the joint and marginalized distributions of fitted parameters ζ and τ. Figure 6 demonstrates how the joint distribution evolves from fitting 1 transit (dashed contours, which represent 68.3%, 95.5%, and 99.7% confidence limits), to fitting 32 transits (solid contours, at the same confidence limits), for KOI 701.01. We then marginalized over all other parameters, to examine the per–parameter confidence limits. Figure 7 demonstrates how our marginalized constraints on ζ (left panel) and τ (right panel) evolved as a function of the number of transits used in the fit for 701.01. The solid line provides the maximum of the posterior distribution, and the dashed line indicates its median. The shaded area encloses 68.3% of the distribution. In this manner, we find a maximum likelihood value of $\tau = 0.051^{0.003}_{0.002}$. This may be contrasted to $\tau = 0.049 \pm 0.003$ derived from reported Borucki et al. (2013) parameters, where they used 171 transits. For completeness, we note our confidence limits on $\zeta = 0.0182^{0.0006}_{0.0003}$ vs the Borucki et al. (2013) result $\zeta = 0.0188 \pm 0.0003$.

To examine our constraints on the system period, we used the $t_{0;1..n}$ posterior distributions from the fits described above. We used `emcee` to sample the posterior space of (nuisance parameter) $t_{0;1}$ and period P. For a given trial $(t_{0;1}, P)$ pair, the likelihood was determined through:

$$\mathcal{L}(t_{0;1}, P) \quad = \quad \prod_{i=1}^{i=n} \kappa_i(t_{0;1} + P * (i-1)) \tag{11}$$

where κ_i is a kernel density estimate of each posterior distribution $t_{0;i}$, which is evaluated at the predicted time of transit $t_0 + P * (i-1)$. By modeling the times of transit separately in the original MCMC analysis, we open the possibility of using a more complex ephemeris model at this stage of the analysis, such as may be expected from transit timing variations (Agol et al., 2005; Holman & Murray, 2005). The results of this analysis for KOI 701.01 are presented in the right panel of Figure 7.

For the derived period after 32 transits, we find $P = 5.71484^{0.00009}_{0.00015}$. The uncertainty in the period roughly scales as a power law with an e–folding timescale of approximately

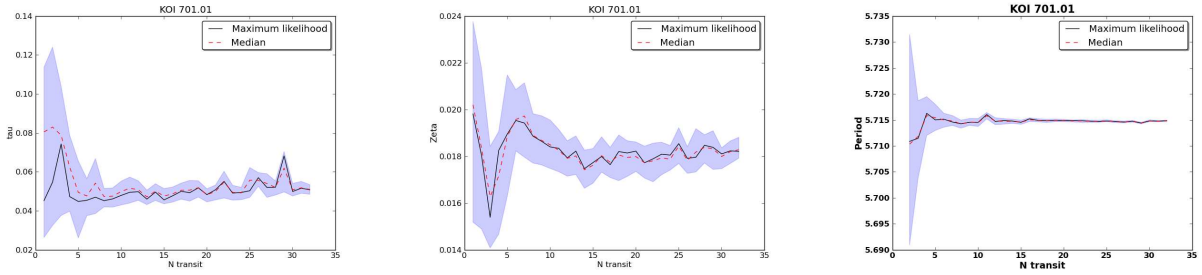

Figure 7: The marginalized distributions of τ (left panel), ζ (center panel), and period P (right panel) as a function of the number of transits being used, for KOI 701.01. For each transit $n \leq 32$ we use *all* data up to and including transit n to constrain ζ and τ (along with β_0^2 and times of transit $t_{0;i=1...n}$). The *solid* line represents the maximum of the posterior distribution, while the *dashed* line indicates its median. The shaded area contains 68.3% of the posterior samples.

---

15 transits. This may be contrasted to $P = 5.714932 \pm 0.000009$ days reported in Borucki et al. (2013). The differences in e_{min} are more substantial. Using their reported value of $\alpha = 18.7 \pm 0.5$, we derive $e_{min} = 0.043 \pm 0.009$, compared to $e_{min} = 0.021 \pm 0.005$ using the Borucki et al. (2013) results. This difference is almost entirely driven by the (1–sigma) differences in τ, which makes it of utmost importance to model this parameter directly.

Lightcurve Simulations

To examine how our abilities to constrain system parameters (in particular, τ) as a function of transit depth and signal–to–noise (S/N), we generated simulated lightcurves at the *Kepler* cadence. We used a subset of 6 of the synthetic systems described in the first section of this proposal. These were chosen to have transit depths of [5e-5/1e-4/5e-4/1e-3/5e-3/1e-2] percent. We use the system inclination, semi–major axis, planet–to–star radius ratio, and period (along with two limb darkening parameters) to generate fake system lightcurves using the method of Mandel & Agol (2002). The system lightcurve is evaluated once each minute and integrated over 30 evaluations to approximate a single *Kepler* long–cadence observation. This is done for a window of 1 day on either side of the given transit midpoint to ensure significant out–of–transit data to include in the fit.

A white–noise component is added to each lightcurve for each of 4 magnitude bins, using the precisions in parts–per–million (ppm) given on the *Kepler* calibration webpage [2]. We evaluate each lightcurve separately for magnitude 8/10/12/14 objects, adding a random contribution of amplitude 11.3/29/80/296 ppm (respectively) to each datapoint as generated above. We do not include red noise, or other transient gaps and features known to exist in the *Kepler* data. We then model these data using the geometric parameterization. Figure 8 outlines how our signal–to–noise on τ scales with the system brightness and transit depth *after fitting 10 transits only* (due to computational constraints). We are able to recover τ

[2] http://keplergo.arc.nasa.gov/CalibrationSN.shtml

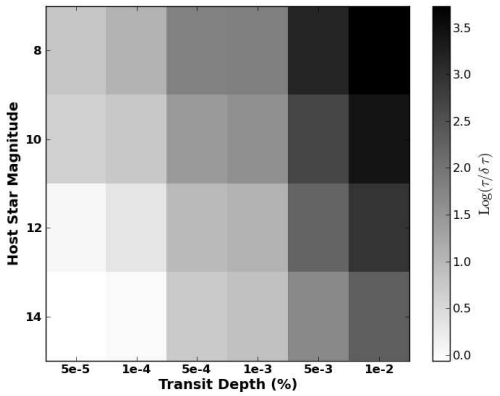

Figure 8: Grid of signal–to–noise in τ as a function of host star brightness and transit depth, derived from fitting our geometric model to simulated noisy lightcurves. The host star brightness sets the S/N of the photometry. These values were obtained after modeling 10 transits jointly. These simulations accurately predict the S/N of τ measurements in KOI 701.01 after 10 transits.

to better than ten percent for all transits deeper than 0.5% down to 14^{th} magnitude; for all transits deeper than 0.1% down to 12^{th}; for all transits deeper than 0.05% down to 10^{th}; and for all transits deeper than 0.01% down to 8^{th}. This analysis is consistent with our results from KOI 701.01, where the host star's magnitude is 13.6, the transit depth is 0.04%, and measured S/N in τ is ~ 7.8 after modeling 10 transits: $0.0479^{0.0073}_{0.0049}$. Figure 8 predicts a S/N of 5–6. Note the measured precision in τ increases to a S/N of approximately 20 after 32 transits (Figure 7).

Kepler Objects of Interest

★ **rory:** Enumerate what is out there to analyze ★ . We need α for R_{crit}, and ages for Q.

890+ KOIs. Which have supporting information on the host star density and age.

Computational Requirements

Using KOI 701.01 as a benchmark, we find a roughly quadratic relationship for the total `emcee` computation time vs the number of transits: $time(n) = -12.5 + 32.2 \times n + 11.9 \times n^2$ seconds (on a single 2.13 GHz core). This relationship was established using a fixed number of 100 burn–in and 1000 "final" steps for each of $2 \times n + 6$ walkers (i.e. 2 walkers for each parameter in the model). If we scale these chains up to 10^6 total steps, which should yield effective chain lengths of $\sim 10^4$, we find a linear relationship in the computation time required to reach 10^6 steps: $time_{10^6}(n) = -1105 + 5943 \times n$ seconds. For a 5–day period system having ~ 300 transits over the assumed 17–quarter operational lifetime of *Kepler* , this comes out to approximately 21 CPU–days of analysis. The `emcee` code is natively able to use multiprocessing capabilities, making this trivial to implement on a multi–core system. However, with ~ 1000 systems on our analysis path, this will require ~ 60 CPU–years of computation, requiring the use of NASA's High-End Computing (HEC) facilities. We will also make use of the local `Hyak` compute cluster when it is available.

Data Interpretation

In the following sections, we describe the theoretical component of our research plan in more detail. In Task A, we consider systems consisting of one star and one planet. In Task B, we include multiple planet systems. In Task C, we incorporate mass loss.

Task A: Tides in Star–Planet Systems

We begin with the simplest treatment, a single planet orbiting a single star in an orbit that can be modified by tides. We will mostly follow the procedure described for our pilot study, but will expand the analysis to a broader range of initial conditions as well as include alternative tidal models. The pilot study (25,000 simulations of $\sim$ 5 Gyr each) requires about 4 hours on a modern workstation, and therefore many such trials are easily tractable. Co–I Barnes is an expert in tidal effects on exoplanets (Barnes et al., 2008; Jackson et al., 2008; Barnes et al., 2009a, 2013), and will use his existing code `eqtide` that follows the prescriptions in Barnes et al. (2013). The free parameters are the mass–radius relationship for both rocky and gaseous bodies, the value of R_{crit}, the Qs of rocky and gaseous bodies, the initial eccentricity distribution, and the age distribution of *Kepler* stars.

The choices in the pilot study were necessarily limited, and we will explore many more options during this proposal. While we assumed that rocky bodies were Earth–like in their composition, other scaling laws are possible (e.g. Seager et al., 2007; Fortney et al., 2007; Lissauer et al., 2011). We will use these other scalings for the masses of the rocky planets, as well as mixing the models to allow for a range of compositions. For the gaseous planets, we will assume different densities in the range 0.5 – 3 g/cm^3, and also mixed cases where the density is chosen at random. The value of R_{crit} is the parameter we are most interested in, and we will grid it finely, varying it from 1 to 2.5 R$_\oplus$in 0.1 R$_\oplus$intervals. We will consider two different models for $Q(R)$, the tidal Q as a function of planetary radius. First we will use the same differences as in the pilot study, but we will also consider a three–tiered model, in which intermediate mass planets have intermediate Qs. Neptune, and possible even Saturn, have a Q value of 10^4 (**?**Lainey et al., 2012), and hence we must consider this possibility, too. This also introduces a new radius cut–off, R_{mid}, which we will allow to move from 2 to 5 R$_\oplus$. The Q range for these systems will have values between 3000 and 30,000. We will randomly choose a stellar mass in the range 0.7–1.4 M$_\odot$as we are interested in FGK stars. ★ **rory:** A sentence on the age distribution ★ . Finally, we will keep the initial eccentricity distribution consistent with that of more distant exoplanets. Ultimately we expect to run several hundred of these suites of systems, which is easily tractable on a modern multi–core workstation.

The examples in Figs. 1–3 used one tidal model, in which the lag angle between the tidal bulge and the perturber is constant regardless of frequency (e.g. Goldreich & Soter, 1966; Jackson et al., 2008). Another popular model assumes that lag angle is instead a function of frequency (e.g. Hut, 1981; Matsumura et al., 2010). We will also employ this model with the same ranges as described above, and relating Q to the time lag tau as $Q = 1/n\tau$, where n is the mean motion (e.g. Correia et al., 2012). In reality there is no general conversion between the two, but this relation is in common use. Thus, our work may also shed light on the frustrating ambiguity in determining the most appropriate equilibrium tidal model.

Task B: Multi–planet Systems

Many of the *Kepler* systems are multiple. Its likely that many singletons are as well, but with non–transiting companions. Mutual gravitational interactions between planets can maintain an eccentricity, even in the presence of strong tidal damping (Mardling & Lin, 2002; Greenberg & Van Laerhoven, 2011; Correia et al., 2012). While this pumping cannot

explain the discrepancy in Fig. 3, it can still prevent some planets from reaching $e_{min} = 0$, and lead to an incorrect determination of R_{crit}. To assess this possibility, we will perform simulations of multiplanet systems undergoing tidal damping.

Ideally these simulations would couple an N–body integrator to the tidal models, as the former can be very accurate. Unfortunately the timescales are so long that such an approach is intractable. We therefore must use classical secular theory in which the evolution for Gyr can be computed in seconds. The second order theory is insufficient for many of the cases we consider, which will have eccentricities up to 0.9. Therefore, we will use higher order theories, which have been previously developed (e.g. Ford et al., 2000; Veras & Armitage, 2004; Libert & Henrard, 2005). We may also take advantage of the high degree of coplanarity of close–in *Kepler* systems Fabrycky et al. (2012), and ignore all terms involving inclination. As each of these models is computationally cheap, we can safely use the 12th order theory of Libert & Henrard (2005) to evaluate the evolution.

More challenging is the choice of initial conditions. While many *Kepler* systems are multiple, we do not yet know the underlying distribution of orbital architectures. A full exploration of parameter space with arbitrary multiplicity and orbital elements would be intractable, and would be very challenging to interpret. We therefore will limit our study to suites of 25,000 systems with multiplicity that follows from the observations. Furthermore, we will limit the size of our planetary systems to < 0.5 AU. More distant companions could certainly affect the orbits of close–in companions, but the more distant companions may also be on inclined orbits, as demonstrated by υ Andromedae c and d (McArthur et al., 2010; Reffert & Quirrenbach, 2011). Therefore, to ease both the computational burden, and not to diverge too far into unconstrained parameter space, we will use these constraints. The physical properties of the planets will be in the same ranges as above. Co–I Barnes has experience with secular theory (Barnes & Greenberg, 2006b,a), and hence will lead this effort. We will perform numerical tests of stability of initial conditions, and will throw out systems that divergently cross strong mean motion resonances – such 2:1, 3:1, and 3:2 – that lead to system break–up (e.g. Gomes et al., 2005).

Additionally, it is likely that many of the close–in systems cannot have initial eccentricities comparable to the non–tidally–evolved planets because they would be unstable. In those cases, the orbits are probably close to their primordial morphologies and experience migration during the protoplanetary disk phase. Recently Dawson & Murray-Clay (2013) showed compelling evidence that high metallicity stars are more likely to host eccentric planets. Therefore, we will make two comparisons with the *Kepler* sample: the full set and the high metallicity set. Although this decreases our statistical robustness, it may yield the most appropriate comparison.

Task C: Atmospheric Mass Loss

In addition to tidal evolution, close–in exoplanets can also experience mass loss, as high energy radiation can liberate hydrogen (Watson et al., 1981; Vidal-Madjar et al., 2003). Jackson et al. (2010) showed that mass loss and tidal evolution are coupled for the case of CoRoT–7 b and that both positive and negative feedbacks are possible. Mass loss is especially important in our study as a planet could transition from a gaseous body to a rocky body during its lifetime, leading to a two–speed evolution.

We will employ the classic mass loss of model of Watson et al. (1981) in which XUV photons liberate hydrogen atoms in the upper atmosphere. Mass loss can be a very complicated process (Yelle, 2004; Lammer et al., 2007; Khodachenko et al., 2007; Leitzinger et al., 2011; Lammer et al., 2013), and with so many unknowns for any given planet, this simple model is most appropriate. The key parameter in this formulation is the efficiency of transforming incident XUV radiation into escaping atoms, ϵ. Most studies of hot Jupiter place ϵ in the range $0.1 - 0.4$. We will therefore explore a range of $0.05 - 0.5$ in increment of 0.05 and apply these to the configurations of Task A and Task B. For the XUV flux, we will use the empirical relationship derived in Ribas et al. (2005) for G dwarfs. While this formulation is not strictly valid for F and K dwarfs, analogous studies do not exist for those spectral types, and the Ribas model is probably a close approximation. The Watson and Ribas models are already in `eqtide` (Barnes et al., 2013), and hence minimal code improvements are required for this step. This final theoretical task requires about 2–3 times more computational resources than the other two combined, but is still dwarfed by the *Kepler* lightcurve analysis, and can easily be completed within a few months on a workstation, or on UW's local supercomputer.

Unfortunately the inclusion of mass loss leads to a degeneracy in our model. The three parameters R_{crit}, Q_r and Q_g, are all related to features in Figure 3. Mass loss complicates the picture by blurring these boundaries. On the other hand, it is entirely possible that we will fail to reproduce the observed $\langle e_{min} \rangle$ distribution without it. At the conclusion of Task C, we will have about 2000 suites of simulations with different physical parameters to compare to the *Kepler* planet candidates. If we succeed in identifying R_{crit}, then planets found in the HZ of *Kepler* targets can be characterized as gaseous or rocky (and potentially habitable), independent of knowledge of their masses. This is the ultimate goal of this proposal.

IV. Team Qualifications and Previous NASA Support

PI Becker is PI on NASA OSS grant NNX09AB32G, "3.5m Transit Timing Observations at 100% Duty Cycle", which observed multiple transiting exoplanet systems for evidence of transit timing variations (Kundurthy et al., 2011, 2013b; Becker et al., 2013; Kundurthy et al., 2013a). Much of the software developed for that project has been modified to operate with the *Kepler* data, and used in the analyses presented here. He has considerable expertise in using modern software packages implemented on distributed computing infrastructures to model multi–dimensional systems. Relevant examples include the work done under NASA ADP grant NNX09AC77G, "Time Domain Studies of the 2MASS Calibration Point Source Working Database", for which he is also PI (Davenport et al., 2012, 2013).

Co–I Barnes was a Co–I on NASA OSS grant 811073.02.07.01.15 "Simulating the Initial Planetesimal Disk" which produced the first N-body simulation of 1 km planetesimal accretion (Barnes et al., 2009b). As part of this effort, Barnes used several hundred thousand hours of CPU time at NASA HPC facilities, such as the Columbia and Plaiedes supercomputers. He has published $\sim$ 40 papers on tidal theory and orbital dynamics, focusing on exoplanets (e.g. Barnes & Quinn, 2001; Barnes & Raymond, 2004; Barnes & Greenberg, 2006b; Barnes et al., 2011, 2013).

Co–I Agol...

V. Relevance to NASA Programs

This project addresses directly multiple objectives that are in–scope for the Origins of Solar Systems call for proposals, including:

- **Observations and theoretical investigations related to the formation and evolution of planetary systems**: This proposal will explore the current minimum eccentricity distribution as a function of orbital period and planetary radius. These data will be combined with extensive and novel theoretical analyses to constrain the initial conditions and evolution of these systems. Critically, this proposal will measure for the first time tidal circularization model parameters R_{crit}, Q_r and Q_g.

- **Characterization of extra-solar planets to explain observations of extra-solar planets**: This proposal will help to draw the critical boundary between gaseous and rocky planets, helping to interpret the observations of exoplanet systems having longer orbital periods (and potentially in their host star habitable zones).

VI. Project Development Plan

PI Becker will be technical lead the project for the first 1.5 years, which will constitute the data analysis (MCMC) portion of the project. He will work at 50% FTE with the graduate student to implement the computations on the *Kepler* sample of data. Co–I Barnes will serve as technical lead for the project for the second 1.5 years, as the project transitions from analysis of the data to interpretation and constraints on tidal evolution theory. Co–I Agol has significant experience in dealing with the *Kepler* data, including implementing detrending algorithms to correct for correlated noise in the *Kepler* data. He will advise as–needed throughout the project. PI Becker will serve as the project lead throughout.

We regard the professional development of students as an important responsibility of *any* research project. In this regard, the graduate students funded by this proposal will have the opportunity to attend at least one relevant conference each year, and encouraged to give oral presentations on our work. This will become a requirement as the project progresses. We expect the graduate student to become an expert in both areas of this project, both the computational/modeling side and the theoretical side. This is a powerful combination and one not seen often enough in the field. We consider this dual–aspect training a strong component of this project.

Year 1 (2014): This first year of the project will start with PI Becker bringing the student up to speed in modeling transit lightcurves, in leaning Bayesian techniques, and in implementing a robust application of the `emcee` package (or other affine–invariant sampler, if needed). Making the software robust to detrending errors, missing data, and initial conditions for the samplers is a non–trivial exercise. Discovering and understanding the failure modes will be a main focus of this computationally–intensive first year. Becker will lead this effort, and transition the student into lead during the year. The generation of the MCMC chains is expected to take ~ 60 CPU–years, requiring the use of high–end computing facilities to allow hundreds of parallelized, multi–core jobs, which should reduce this computation time trivially to calendar–months. The validation of these chains (Gelman–Rubin $\hat{\mathrm{R}}$ and effective chain lengths) is expected to take a comparable amount of time, as some chains are expected to have to be re–run or extended. The goal of this first year is to have finalized

the MCMC chains on $t_0, \beta_0^2, \tau, \zeta$ for all transits of all KOIs. We will make these publicly available through a `github` site specifically designed for this project. Our modeling software will also be released via `github`.

Year 2 (2015): The second year is expected to begin with the MCMC analyses of the periods via times of transit, and transition to theoretical interpretation of the systems. We will start with linear ephemeris models for all systems to estimate periods; for those systems where this model is insufficient, we will examine transit timing models with more complicated ephemerides. We anticipate that this effort may result in the publication of 1 or more ancillary papers on the transit timings of the systems under study, with assistance from Agol on the interpretation of the results. During this year, Barnes will begin working with the graduate student on the theoretical aspects of the project. They will design and simulate the models described in Task A and publish a preliminary estimate of R_{crit}, Q_g, and Q_r. During this year we will also begin running simulations of multiplanet systems with tidal damping.

Year 3 (2016): During the final year, we will finish all theoretical modeling, including the incorporation of mass loss. We will publish a paper on the role of multiplicity in the e_{min} distribution. The graduate student will perform a final analysis of all available *Kepler* data and will compare this final data set to the synthetic data produced by the tides+multiplicity+evaporation model. A final paper will summarize the results of the investigation, including final values, with error estimates, for R_{crit}, Q_r, Q_g, and ϵ. We will characterize the degeneracies between these parameters using a Bayesian framework.

VII. Data Sharing Plan

All investigators are committed to the sharing of data and software. PI Becker has been behind real–time public alert systems for many time–domain astronomical surveys. This includes the MACHO survey, the Deep Lens Survey, the SuperMACHO and ESSENCE surveys, and the SDSS–II Supernova Survey, all of which have released their events to the public in near–real time through web pages, Astronomer's Telegrams, IAU circulars, and VOEvents. He has been diligent in releasing the software and data behind publications. This includes the image subtraction software `hotpants`, period finding software `Supersmoother`, and spatial clustering software `OPTICS`[3] which have been used in many subsequent publications. He is currently working part–time on the Large Synoptic Survey Telescope (LSST), which is both open–source and open–data.

We will version release all software developed for this project on the publicly available open–source collaboration website http://github.com (`github` hereafter). The website has become a leading collaboration platform; it enables distributed users to download code and contribute back to the project. All code we develop for this project will be made available under the terms of the open source BSD license[4] whenever possible. We will make a new `github` account for this project that we will use to stage code and data releases, as described in the project development plan. Analysis packages used in our publications will be released as `iPython`[5] notebooks to help establish reproducible research standards in the field. `iPython`

[3] http://www.astro.washington.edu/users/becker/c_software.html
[4] http://www.opensource.org/licenses/bsd-license.php
[5] http://ipython.org

allows the exchange of portable environments (notebooks) that enable the user to follow the analysis path leading to a given scientific result, and also to interact with it at the code level to understand (and verify) the methodology.

References

Agol, E., Steffen, J., Sari, R., & Clarkson, W. 2005, *MNRAS*, 359, 567

Aksnes, K., & Franklin, F. A. 2001, *AJ*, 122, 2734

Anglada-Escudé, G., et al. 2012, *ApJ*, 751, L16

Barnes, J. W. 2007, *PASP*, 119, 986

Barnes, R., & Greenberg, R. 2006a, *ApJ*, 652, L53

Barnes, R., & Greenberg, R. 2006b, *ApJ*, 638, 478

Barnes, R., Greenberg, R., Quinn, T. R., McArthur, B. E., & Benedict, G. F. 2011, *ApJ*, 726, 71

Barnes, R., Jackson, B., Greenberg, R., & Raymond, S. N. 2009a, *ApJ*, 700, L30

Barnes, R., Mullins, K., Goldblatt, C., Meadows, V. S., Kasting, J. F., & Heller, R. 2013, Astrobiology, 13, 225

Barnes, R., & Quinn, T. 2001, *ApJ*, 550, 884

Barnes, R., Quinn, T. R., Lissauer, J. J., & Richardson, D. C. 2009b, *Icarus*, 203, 626

Barnes, R., & Raymond, S. N. 2004, *ApJ*, 617, 569

Barnes, R., Raymond, S. N., Jackson, B., & Greenberg, R. 2008, Astrobiology, 8, 557

Batalha, N. M., et al. 2011, *ApJ*, 729, 27

Batalha, N. M., et al. 2013, *ApJS*, 204, 24

Becker, A. C., Kundurthy, P., Agol, E., Barnes, R., Williams, B. F., & Rose, A. E. 2013, *ApJ*, 764, L17

Bolmont, E., Selsis, F., Raymond, S. N., Leconte, J., Hersant, F., Maurin, A.-S., & Pericaud, J. 2013, ArXiv e-prints

Borucki, W. J., et al. 2013, ArXiv e-prints

Correia, A. C. M., Boué, G., & Laskar, J. 2012, *ApJ*, 744, L23

Darwin, G. H. 1880, Royal Society of London Philosophical Transactions Series I, 171, 713

Davenport, J. R. A., Becker, A. C., Kowalski, A. F., Hawley, S. L., Schmidt, S. J., Hilton, E. J., Sesar, B., & Cutri, R. 2012, *ApJ*, 748, 58

Davenport, J. R. A., et al. 2013, *ApJ*, 764, 62

Dawson, R. I., & Murray-Clay, R. A. 2013, *ApJ*, 767, L24

Fabrycky, D. C., et al. 2012, ArXiv e-prints

Ferraz-Mello, S., Rodríguez, A., & Hussmann, H. 2008, Celestial Mechanics and Dynamical Astronomy, 101, 171

Ford, E. B., Kozinsky, B., & Rasio, F. A. 2000, *ApJ*, 535, 385

Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, *PASP*, 125, 306

Fortney, J. J., Marley, M. S., & Barnes, J. W. 2007, *ApJ*, 659, 1661

Gelman, A., & Rubin, D. 1992, Statistical Science, 7, 457

Goldreich, P., & Soter, S. 1966, *Icarus*, 5, 375

Gomes, R., Levison, H. F., Tsiganis, K., & Morbidelli, A. 2005, *Nature*, 435, 466

Goodman, J., & Weare, J. 2010, Communications in Applied Mathematics and Computational Science, 5, 65

Greenberg, R. 2009, *ApJ*, 698, L42

Greenberg, R., & Van Laerhoven, C. 2011, *ApJ*, 733, 8

Holman, M. J., & Murray, N. W. 2005, Science, 307, 1288

Hut, P. 1981, *A&A*, 99, 126

Jackson, B., Barnes, R., & Greenberg, R. 2009, *ApJ*, 698, 1357

Jackson, B., Greenberg, R., & Barnes, R. 2008, *ApJ*, 678, 1396

Jackson, B., Miller, N., Barnes, R., Raymond, S. N., Fortney, J. J., & Greenberg, R. 2010, *MNRAS*, 407, 910

Khodachenko, M. L., et al. 2007, Astrobiology, 7, 167

Kundurthy, P., Agol, E., Becker, A. C., Barnes, R., Williams, B., & Mukadam, A. 2011, *ApJ*, 731, 123

Kundurthy, P., Barnes, R., Becker, A. C., Agol, E., Williams, B. F., Gorelick, N., & Rose, A. 2013a, ArXiv e-prints

Kundurthy, P., Becker, A. C., Agol, E., Barnes, R., & Williams, B. 2013b, *ApJ*, 764, 8

Lainey, V., et al. 2012, *ApJ*, 752, 14

Lammer, H., Erkaev, N. V., Odert, P., Kislyakova, K. G., Leitzinger, M., & Khodachenko, M. L. 2013, *MNRAS*, 430, 1247

Lammer, H., et al. 2007, Astrobiology, 7, 185

Leconte, J., Chabrier, G., Baraffe, I., & Levrard, B. 2010, *A&A*, 516, A64

Leitzinger, M., et al. 2011, *Planet. Space Sci.*, 59, 1472

Libert, A.-S., & Henrard, J. 2005, Celestial Mechanics and Dynamical Astronomy, 93, 187

Lissauer, J. J., et al. 2011, *Nature*, 470, 53

Lopez, E. D., Fortney, J. J., & Miller, N. 2012, *ApJ*, 761, 59

MacDonald, G. J. F. 1964, Reviews of Geophysics and Space Physics, 2, 467

Mandel, K., & Agol, E. 2002, *ApJ*, 580, L171

Mardling, R. A., & Lin, D. N. C. 2002, *ApJ*, 573, 829

Matsumura, S., Peale, S. J., & Rasio, F. A. 2010, *ApJ*, 725, 1995

McArthur, B. E., Benedict, G. F., Barnes, R., Martioli, E., Korzennik, S., Nelan, E., & Butler, R. P. 2010, *ApJ*, 715, 1203

Owen, J. E., & Wu, Y. 2013, ArXiv e-prints

Poppenhaeger, K., Czesla, S., Schröter, S., Lalitha, S., Kashyap, V., & Schmitt, J. H. M. M. 2012, *A&A*, 541, A26

Rasio, F. A., Tout, C. A., Lubow, S. H., & Livio, M. 1996, *ApJ*, 470, 1187

Reffert, S., & Quirrenbach, A. 2011, *A&A*, 527, A140

Ribas, I., Guinan, E. F., Güdel, M., & Audard, M. 2005, *ApJ*, 622, 680

Seager, S., Kuchner, M., Hier-Majumder, C. A., & Militzer, B. 2007, *ApJ*, 669, 1279

Sing, D. K. 2010, *A&A*, 510, A21

Sotin, C., Grasset, O., & Mocquet, A. 2007, *Icarus*, 191, 337

Tegmark, M., et al. 2004, *Phys. Rev. D*, 69, 103501

Valencia, D., Ikoma, M., Guillot, T., & Nettelmann, N. 2010, *A&A*, 516, A20

Veras, D., & Armitage, P. J. 2004, *Icarus*, 172, 349

Vidal-Madjar, A., Lecavelier des Etangs, A., Désert, J.-M., Ballester, G. E., Ferlet, R., Hébrard, G., & Mayor, M. 2003, *Nature*, 422, 143

Watson, A. J., Donahue, T. M., & Walker, J. C. G. 1981, *Icarus*, 48, 150

Yelle, R. V. 2004, *Icarus*, 170, 167

Budget Justification

PI Salaries:

We include 6 months salary in the first year for PI Becker, and 3 months in the second year. Becker is on the Research Faculty at UW, meaning his salary must be obtained through grants such as this one. Benefits are calculated at the rate of 26.9%. We budget for a 2% annual increase in these salaries.

We include 3 months salary in the second year for Co–I Barnes, and 4 months in the third year. Barnes is on the Research Staff at UW, meaning his salary must be obtained through grants such as this one. Benefits are calculated at the rate of 34.0%. We budget for a 2% annual increase in these salaries.

Other Salaries:

One graduate student will be funded for 3 academic quarters per year at 60% time, and 2 summer quarters at 100% time. Benefits are calculated at the rate of 14.2%. We budget for a 2% annual increase in these salaries.

Tuition Costs:

We include tuition fees for one graduate student at the rate of $\$4,689$ per quarter (first year), for 3 academic quarters per year. We budget for a projected 10% annnual increase in these fees after the first year, and 12% subsequently.

Equipment:

None

Travel:

We budget for 2 domestic trips per year for collaboration and conferences, at the rate of $1,500$ per trip (to cover travel to and 4 days lodging to the East Coast). This will be shared by the investigators and graduate student.

Publication Charges:

Publication costs are budgeted at the electronic publishing charge of \$110/page, for 20 pages/year.

Computing Fees:

Computer support fees are budgeted at the nominal rate of \$67 per person per month by the Physics and Astronomy Computing Services group (PACS) at UW.

Indirect Costs:

Indirect costs are based on the MTDC rate of 54.5% per the negotiated agreement with DHHS dated 3/5/2013.

Personnel and Work Effort:

PI Becker will work on this project at 50% effort for the first year of the project, and 25% for the second. His focus will be on implementing a robust MCMC analysis of all the Kepler data, in debugging failure modes, and in training the graduate student in the art of Bayesian analysis.

Co–I Barnes will work on this project at 25% effort in the second year of the project, and 33% in the third.

Co–I Agol will assist as–needed.

The graduate student will work on this project at 60% FTE for 3 academic quarters, and 100% FTE for 2 summer quarters, for all 3 years of the project. Their role will be to become an expert in the analysis of the data, and in the interpretation of the results and how it relates to tidal dissipation in the population of planets studied.

Facilities and Equipment:

As this is a compute–heavy proposal, we will be making use of the local University of Washington `Hyak` compute cluster, as well as applying for time on NASA's High-End Computing (HEC) facilities. The University provides office equipment for all investigators.