Combining Transit and Radial Velocity Data

Leslie A. Rogers[1,2,3]

[1]Hubble Fellow, Cahill Center for Astronomy & Astrophysics, Caltech
[2]Sagan Fellow, Department of Earth and Planetary Sciences, University of California Berkeley
501 Campbell Hall #3411, Berkeley, CA 94720, USA
[3]Department of Astronomy & Astrophysics, University of Chicago
5640 S Ellis Ave, Chicago, IL 60637, USA
email: `larogers@uchicago.edu`

Abstract. The *Kepler* Mission, combined with ground based radial velocity (RV) follow-up, has revolutionized the observational constraints on sub-Neptune-size planet compositions. *Kepler*'s unprecedentedly large and homogeneous samples of planets with both mass and radius constraints open the possibility of statistical studies of the underlying planet composition distribution. This presentation describes the application of hierarchical Bayesian models to constrain the underlying planet composition distribution from a sample of noisy mass-radius measurements. This approach represents a promising avenue toward a quantitative measurement of the amount of physical scatter in small planet compositions, the identification of planet sub-populations that may be tied to distinct formation pathways, and empirical constraints on the dominant compositional trends in the planet sample. Both the transit and radial velocity techniques are subject to selection effects, and approaches to mitigate the resulting biases will be addressed. In addition to distilling composition-distribution insights from the current sample of Kepler planets with RV masses, this framework may be used to optimize the target selection for future transiting planet RV follow-up surveys.

Keywords. methods: data analysis, methods: statistical, planetary systems, planets and satellites: composition, techniques: photometric, techniques: radial velocities

1. Introduction

In terms of constraining planet bulk compositions, the subset of planets detected both with transits and with radial velocity (RV) measurements or transit timing variations (TTVs) are most valuable. The planet radius measured from the transit depth and the planet mass measured from the planet's gravitational influence (on their host star or other planets in the system) together combine to give the planet's mean density and some handle on the planet's bulk composition. Planet bulk compositions in turn provide insights into the a planet's formation and evolution. The presence of a H/He envelope surrounding a planet constrains the timing of the planet's formation relative to the protoplanetary disk lifetime, $t_{\rm disk} \sim 3$ Myr. The water-to-rock mass ratio of a planet is a tracer of the formation location relative to the snow line in the protoplanetary disk.

In this paper, we focus on how to combine insights from transit and RV and/or TTV measurements to constrain the range of possible compositions for a single planet (Section 2), and to derive insights into the composition distribution of planets from a statistical sample of noisy mass-radius measurements (Section 3).

2. Single Planets

The main planet properties that can be observationally constrained by transit and RV data are mass (M_p), radius (R_p), and incident flux (F_p). Inferring planet compositions

from mass-radius data can be cast in a Bayesian framework as first formulated by Rogers & Seager (2010a),

$$p\left(M_p, \{x_i\} \mid \mathcal{D}\right) \propto \left\{ \int p\left(\mathcal{D} \mid M_p, R_p\right) p\left(R_p \mid M_p, \{x_i\}\right) dR_p \right\} p\left(M_p, \{x_i\}\right). \quad (2.1)$$

We denote by $\{x_i\}$ the composition of the planet (with x_i representing the mass fraction of the planet in the ith component). The likelihood function, $p\left(\mathcal{D} \mid M_p, R_p\right)$ represents the likelihood of the transit and RV data (denoted by $\mathcal{D}$) given true planet M_p and R_p. The next factor, $p\left(R_p \mid M_p, \{x_i\}\right)$, encapsulates insights from planet interior structure and evolution models, which provide a mapping $f : [M_p, \{x_i\}] \mapsto R_p$. Most analyses to date have effectively placed absolute trust in the models, treating this term as a Dirac-delta function $p\left(R_p \mid M_p, \{x_i\}\right) = \delta\left(R_p - f\left(M_p, \{x_i\}\right)\right)$. Finally, $p\left(M_p, \{x_i\}\right)$ denotes the priors on the mass-composition distribution given the assumptions and information available before any data is collected. Equation 2.1 above is distilled for clarity; it could readily be generalized *i)* to be cast in terms of the radial velocity amplitude, transit depth and host star properties, and/or *ii)* to explicitly include additional dependencies (e.g., on age, F_p, host star metalicity, host star chemical abundances etc.).

Marginalizing over the planet mass gives the posterior pdf of the planet compositions,

$$p\left(\{x_i\} \mid \mathcal{D}\right) = \int p\left(M_p, \{x_i\} \mid \mathcal{D}\right) dM_p. \quad (2.2)$$

Model degeneracies complicate exoplanet compositional inferences (e.g, Valencia *et al.* 2006, Adams *et al.* 2008, Rogers & Seager 2010ab). The mapping $\tilde{f} : [M_p, \{x_i\}] \mapsto [M_p, R_p]$ is not one-to-one; planets with very different compositions can have identical masses and radii, especially if > 2 distinct chemical components are considered (e.g., H/He, ices, rock). Consequently, single planet compositional inferences are intrinsically in an under-constrained regime where the prior assumptions strongly affect the results (posteriors) obtained. It is thus crucial to ask the questions that can be most robustly answered. For instance, the mass-radius relations of rocky planets are most sensitive to the total iron mass fraction of the planets and are less sensitive to the degree of differentiation or oxidation within the planets (e.g., Elkins-Tanton & Seager 2008, Rogers *et al.* 2011, Zeng & Sasselov 2013). Similarly the mass-radius relations for planets with gas envelopes are most sensitive to the H_2/He gas mass fraction, but less sensitive to the ice-to-rock mass ratio (e.g., Rogers *et al.* 2011, Valencia *et al.* 2013). As a result, the total iron mass fraction of rocky planets and the H/He mass fraction of gas-laden planets are more readily constrained than other aspects of the planets' compositions. Examples of Bayesian compositional inferences for individual planets include Kepler-36 b, c (Carter *et al.* 2012) and PH3 b, c, d (Schmitt *et al.* 2014).

3. Planet Populations

The accumulating census of exoplanets with observational constraints on both their mass and radius offers the opportunity to move beyond case studies of individual objects to studies of the overall population. Hierarchical models provide a natural framework to constrain the underlying composition distribution of planets from a noisy sample of mass-radius measurements. Hierarchical models open the priors on the planet compositions to modeling. Given observations of N planets and a mod el for the joint planet mass-composition distribution that depends on some hyperparameters ϕ, we may generalize Equation 2.1,

$$p(\phi, \{M_{pj}\}, \{x_{ij}\}|\mathcal{D}) \propto \prod_{j=1}^{N} \left[\left\{ \int p(\mathcal{D}|M_{pj}, R_{pj}) p(R_{pj}|M_{pj}, \{x_{ij}\}) \mathrm{d}R_{pj} \right\} \right.$$
$$\left. \times p(M_{pj}, \{x_{ij}\}|\phi) \right] p(\phi) \tag{3.1}$$

Equation 3.1 is simplified to treat each jth planet as independent; it can, however, be readily generalized to account for correlations between multiple planets within the same system (from corellated measurements and/or from physical correlations between the true planet properties). So far, hierarchical models have been applied to the sample of $\mathcal{O}(50)$ *Kepler* planets having Keck HIRES RV follow-up to constrain, as a function of planet size, the fraction of planets that are sufficiently dense to be rocky (Rogers 2015), and to fit a power-law mass-radius relationship allowing for intrinsic scatter in the planet masses at a given radius (Wolfgang *et al.* 2015). Once the sample of small transiting planets with RV mass constraints grows to $\mathcal{O}(200)$, analyses focused on the joint mass-composition distribution of planets (as in Equation 3.1) could plausibly constrain the typical mass scale of rocky planets, the core mass-envelope mass relationship of gas-laden planets, and extent of physical scatter in exoplanet compositions (Rogers, in prep.).

A future with $\mathcal{O}(200)$ small planet mass-radius measurements may be only a couple years away. There is a suite of space-based transit surveys on the horizon (K2, TESS, CHEOPS and PLATO) that will find planets transiting bright stars that are more amenable to ground-based RV follow-up than typical *Kepler* planets. This combined with ongoing investment in the technological development of ground-based spectrographs (e.g., MAROON-X, Carmenes, HPF, SHREK, SPIRou, ESPRESSO, MINERVA) means the pace of exoplanet discovery and characterization is poised to continue accelerating.

We conclude with a note of caution. To ensure that future transit-search missions and RV follow-up instruments are leveraged to maximum advantage in characterizing the composition distribution of small exoplanets, care must be taken to characterize and to mitigate selection effects and biases. *Kepler* mass-radius populations studies (e.g., Rogers 2015, Wolfgang *et al.* 2015) have so far focused on mass distributions conditioned on planet radius to factor out the dominant selection effects. The next major advances in the study of the planet composition distribution will require a more sophisticated treatment of selection effects. Transiting planet radial velocity follow-up surveys directed at population level inferences (as opposed to case studies of individual systems) should *i)* aim to follow an algorithmic approach to selecting their targets and *ii)* report all measurements in the survey (including non-detections and RV upper-limits).

References

Adams, E. R., Seager, S., & Elkins-Tanton, L. 2008, *ApJ*, 673, 1160
Carter, J. A., Agol, E., Chaplin, W. J., *et al.* 2012, *Science*, 337, 556
Elkins-Tanton, L. T. & Seager, S. 2008, *ApJ*, 688, 628
Rogers, L. A. 2015, *ApJ*, 801, 41
Rogers, L. A., Bodenheimer, P., Lissauer, J. J., & Seager, S. 2011, *ApJ*, 738, 59
Rogers, L. A. & Seager, S. 2010a, *ApJ*, 712, 974
—. 2010b, *ApJ*, 716, 1208
Schmitt, J. R., Agol, E., Deck, K. M., *et al.* 2014, *ApJ*, 795, 167
Valencia, D., O'Connell, R. J., & Sasselov, D. 2006, *Icarus*, 181, 545
Valencia, D., Guillot, T., Parmentier, V., & Freedman, R. S. 2013, *ApJ*, 775, 10
Wolfgang, A., Rogers, L. A., & Ford, E. B. 2015, ArXiv e-prints
Zeng, L. & Sasselov, D. 2013, *PASP*, 125, 227