

# MITIGATING EXOPLANET RADIUS BIASES FROM WAVELENGTH DEPENDENCE OF STARSPOT CONTRAST

## 1 Introduction

Recent work has shown that ignoring starspots systematically biases measurements of exoplanet radii (Rackham et al. 2018). Here we propose a strategy for mitigating starspot-induced exoplanet radius biases, based on a large sample of spotted stars observed in *TESS* Cycle 2. The approach will deliver constraints on both the typical starspot coverage fraction and the typical starspot temperature contrast, as a function of spectral type and rotation rate. This mapping can then serve as a “lookup table” for practitioners to gauge the extent to which a given source is likely to suffer from starspot-induced exoplanet radius biases. No such reliable lookup table presently exists for main sequence stars, owing to the inability of differential lightcurves to account for *total* starspot coverage fractions (Basri 2018). This program will leverage the *TESS* Cycle 2 overlap with the *Kepler* prime field and portions of K2 fields.

## 2 Scientific Justification: The Transit Light Source Effect

The diminution of flux arising from occultations of the stellar disk contains information the relative sizes of the occulting bodies. A commonsense exoplanet heuristic dictates that the relative transit depth  $D$  equals the ratio of the bodies’ projected areas ( $R_p^2/R_\star^2$ ). Additional correction factors account for limb darkening (Mandel and Agol 2002). Starspots on a stellar surface can exhibit spatial segregations that yield *unocculted starspots*: spots that are present on the surface but are not traversed during transits. These unocculted starspots confound the mapping of transit depth to planet radius: a planet “looks bigger” than it really is if it blocks the brighter-on-average flux of the stellar disk (Morris et al. 2018). **You cannot measure a planet radius without making some assumptions about the extent, contrast, and distribution of starspots.**

The prevailing assumption within the exoplanet community has been that starspot covering fractions are small enough—as in Sunspots—to be ignored. New theoretical models (Rackham et al. 2018) and observational evidence (Fang et al. 2016) indicate that starspot covering fractions are large enough to bias transit depths beyond the level of precision of *Kepler* and *TESS*. This so-called “Transit Light Source Effect” (TLSE) can bias derived exoplanet radii up to 17% in the *TESS* bandpass, and can bias derived exoplanet *densities* up to 50%, since the radius error propagates into volume error with a factor of 3. Therefore, TLSE threatens placement and interpretation of the exoplanet mass-radius relation science requirements of the *TESS* mission. The interpretation of *JWST* exoplanet transmission spectroscopy is also at risk (Wakeford et al. 2019).

## 3 Analysis Plan

**Kepler and TESS perceive different starspot contrasts.**

Starspots emit light, which is often parameterized as *spot contrast*,  $c$ , where 0 means the spots are non-emitting, and  $c = 1$  means the spot matches the specific intensity of the ambient photosphere. Spot contrast is wavelength dependent ( $c_\lambda$ ) and can be thought of as possessing a characteristic spot temperature,  $T_{spot} < T_{phot}$  that parameterizes the wavelength dependence of the contrasts. The analysis technique proposed here leverages the different wavelength dependence of the *Kepler* and *TESS* system throughputs to directly measure the ratio of the spot contrasts in their respective bandpasses. The spot contrast ratio—evaluated at these two pairs of bandpasses—provides a unique

solution<sup>1</sup> to  $T_{spot}(\text{SpT}, P_{\text{rot}})$ .

**Spot-induced modulation amplitudes depend on contrast.** Rotationally-modulated lightcurves exhibit semi-sinusoidal undulations as starspots (or spot groups) enter and exit the projected stellar disk. The most-spotted projected hemisphere is directed towards the observer at the moment of the lightcurve minimum, and the least-spotted projected hemisphere at lightcurve maximum. The amplitude of the lightcurve, therefore, equals the difference in starspot coverage fractions  $\Delta f_{spot} \equiv f_{max} - f_{min}$  between the most-and-least spotted projected hemispheres weighted by the spot contrast:  $A_{\lambda} = \Delta f_{spot} \cdot (1 - c_{\lambda})$ . Therefore the ratio of *TESS* to *Kepler* amplitudes yields a constraint independent of the actual spot coverage fraction:

$$\frac{A_{TESS}}{A_{Kepler}} = \frac{\Delta f_{spot}}{\Delta f_{spot}} \cdot \frac{1 - c_{TESS}}{1 - c_{Kepler}}$$

Recent stellar surface inhomogeneity modeling (Rackham et al. 2018; Guo et al. 2018; Basri 2018) has yielded coarse scaling factors  $\zeta$  for estimating the total starspot surface coverage fraction from the observed lightcurve amplitude  $f_{spot}$ , yielding  $p(f_{spot}) = \zeta \cdot A$ , where  $\zeta$  represents a probability distribution with values typically between 1 and 50, depending on the degree of assumed starspot longitudinal symmetry. Equipped solely with *TESS*-measured lightcurve amplitudes, we can provide what practitioners seek: a per-star estimate of  $f_{spot}$  and  $T_{spot}$ .

#### Distinguishing Kepler/K2 and TESS amplitudes with statistically similar samples

Ideally, we would have simultaneously observed the same targets with *Kepler* and *TESS* to measure the contrast for individual stars. Since this observational strategy was not possible for a large number of stars, we instead construct statistical subsamples of stars that are identical in every way, except for their observation band.

The main two factors that affect spot-modulation lightcurve amplitudes are effective temperature (*i.e.* mass), and rotation period (*i.e.* stellar activity/magnetic fields) (McQuillan et al. 2014). Other important-albeit-unknown factors like stellar inclination angle and multiplicity are statistically identical between the *TESS* and *Kepler* samples. We, therefore, propose to follow McQuillan et al. (2014) by constructing plots of amplitude versus period, subdivided into bins of effective temperature. The two samples should look identical in these plots, but the locus of *TESS* sources is expected to fall below the *Kepler* sample, as in panel D of Figure 1.

*TESS* Cycle 2 will have significant overlap with both the *Kepler* prime field (up to 3 sectors of overlap) and 16438 source in the northern edges of K2 campaign fields (K. Colón, priv. comm.). This overlap sample offers an independent check on any statistical biases since the exact same stars can be compared. The fact that observations are not contemporaneous adds some *cosmic variance*, but on average stars will enter and exit stellar activity cycles in equal measure, injecting per-star variance, but keeping the subsample mean amplitudes the same.

## 4 Technical Feasibility

Measuring periods and amplitudes of *Kepler* and *TESS* lightcurves is a routine operation, with relatively low risk. The PI has significant experience determining amplitudes with a Fourier-based decomposition technique in young stars observed in K2. The derived periods and amplitudes need not be *perfect*, so long as the biases are about the same in both the *Kepler* and *TESS* subsamples.

The yield of spotted stars in *TESS* will number in the hundreds of thousands. The *TESS* mission was designed to target cool main sequence stars with an abundance of K and M dwarfs.

<sup>1</sup>Operationally, integrating the specific intensity  $I_{\lambda}(T_{spot})$  over the system throughputs produces the expected contrast ratios. For now we assume  $I_{\lambda}(T_{spot})$  can be approximated as a cooler stellar photosphere, such as the PHOENIX model grid (Husser et al. 2013), an inexact but adequate assumption for the large bandpasses.

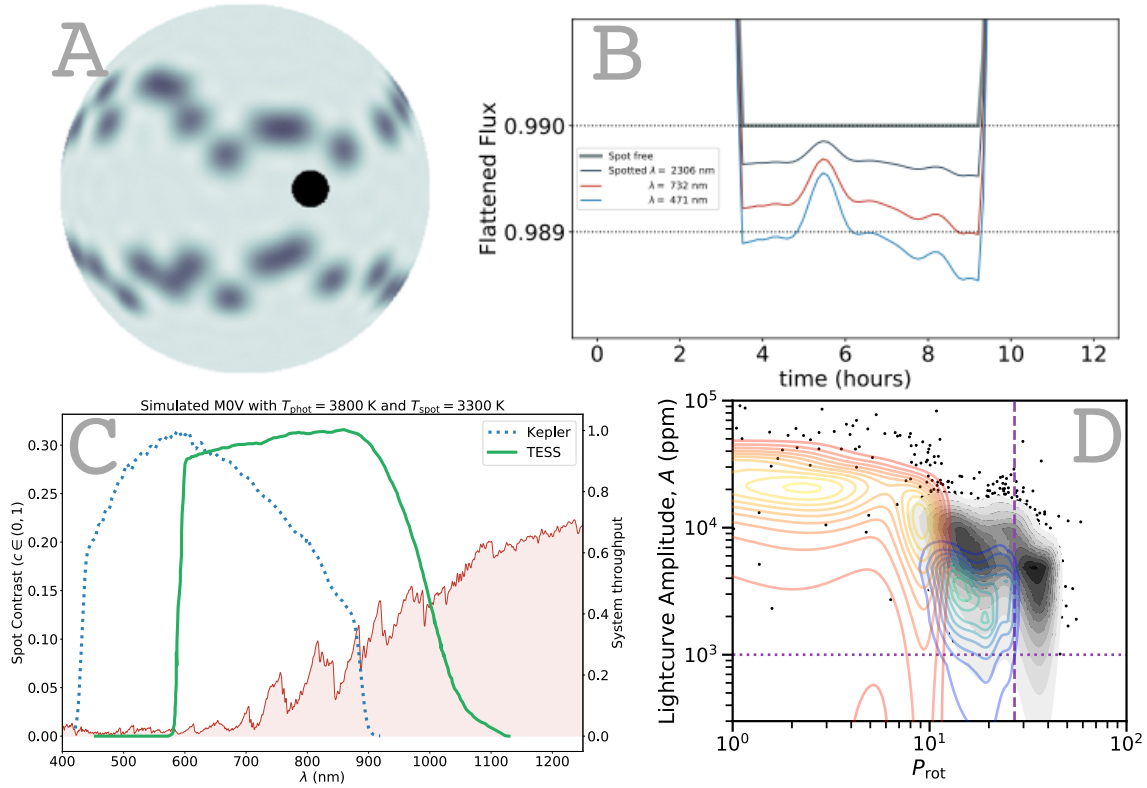


Figure 1: A) Schematic of a planet ( $R_p = 0.1R_\star$ ) transiting a star with unocculted starspots as the star moves from left to right. B) The resulting transit depth shows a systematic bias and a wavelength dependence due to starspots. C) Illustration of TLSE, D) Prediction for *TESS* amplitudes (colorful contours) relative to the *Kepler* sample (gray shaded background contours and black points) for  $4000 < T_{\text{phot}}(K) < 4500$  (McQuillan et al. 2014).

This program will focus on the rapidly rotating stars ( $P_{\text{rot}} < 20$  days), which exhibit relatively large amplitudes ( $A > 2000$  ppm), illustrated in panel D of Figure 1. This program will attempt—on a best effort basis—to measure the slowest rotating stars, which have periods much longer than the typical observation window of 27–54 days, and amplitudes comparable to the precision limits of *TESS*; these slow rotating sources must reside near the continuous viewing zone in the North Ecliptic Pole to have any hope of detecting spot modulation.

One potential risk to this program is the that the *Kepler* and *TESS* populations will be statistically different from each other in their unobserved properties, confounding the assumption that the samples’ only difference is the observation bandpass. For example, metallicity, age, extinction, and stellar background contamination may all be different between the *Kepler* samples and *TESS* samples. This risk is alleviated by the subsample of *Kepler* sources reobserved in *TESS* Cycle 2.

## 5 Expected Impact

The main deliverable of this work will be a probabilistic assessment of the starspot temperature and starspot coverage fraction for every star with a measured starspot modulation period and amplitude in *TESS*. These probability distributions will empower exoplanet researchers to correct for biases in exoplanet transit radii. The probability distributions will reflect all the available information

garnered from *TESS*, and reasonable assumptions about starspot geometrical distributions.

Stellar surface structures serve as a stepping stone for brown dwarf cloud modeling, a poorly understood phenomenon that will inform our understanding of exoplanet spectroscopy. The proposed work will demonstrate how to gain insights about probable surface structures in the presence of significant geometrical degeneracies.

The starspot properties are also of intrinsic astrophysical interest. Trends in  $T_{spot}$  with rotation rate and mass can reveal important clues to stellar activity, magnetic fields, and stellar atmospheres. This work will directly address starspot-induced scatter and bias in isochronal-derived stellar ages (Somers and Pinsonneault 2015; Gully-Santiago et al. 2017).

## 6 Work Plan

PI Gully-Santiago will carry out the data analysis using custom scripts that build upon the heritage of the *Kepler* data analysis ecosystem, including `lightkurve`, `astropy`, `celerite`, `BombScargle`, `PyTorch`, and `bokeh`. The PI will first apply these scripts to Kepler/K2 lightcurves while in the role as a support scientist at the Kepler/K2 Guest Observer Office. The PI will relocate to the University of Texas at Austin in February 2020 as a Research Fellow. UT Austin undergraduate and graduate students will lead the application of TESS-adapted scripts to *TESS* FFI lightcurves with `TESS-cut` and `elenor`. Half of the funding will go to 1 undergraduate and 1 graduate student Summer 2020 salary, and 0.3 FTE to support the PI. All reproducible scripts—including Jupyter Notebook tutorials—will be publicly available and developed openly on GitHub. MGS will provide students with an interactive `bokeh` tool to visualize periodic lightcurve structures. This tool will leverage the PI’s familiarity with `bokeh` to aid rapid visual spot-checking of the multidimensional analysis output. Finally, a High-Level Science Product (HLSP) will be delivered to MAST that contains the following information for each star found to possess detectable starspot-induced lightcurve modulation: 1) the period or periods, 2) an iteratively determined outlier mask of stellar flares, occultations, or spurious data artifacts, 3) a background-star amplitude-dilution estimate, 4) a probability distribution for the spot temperature and coverage fraction, 5) addition model fitting metadata. MGS will co-advise—with Co-I C. Morley—students in the writeup of the paper(s), authored and/or coauthored by students.

## References

- Gibor Basri. *ApJ*, 865:142, Oct 2018. • Xiang-Song Fang, Gang Zhao, Jing-Kun Zhao, Yu-Qin Chen, and Yerra Bharat Kumar. *MNRAS*, 463:2494–2512, Dec 2016. • M. A. Gully-Santiago, G. J. Herczeg, I. Czekala, G. Somers, K. Grankin, K. R. Covey, J. F. Donati, S. H. P. Alencar, G. A. J. Hussain, B. J. Shappee, G. N. Mace, J.-J. Lee, T. W.-S. Holoién, J. Jose, and C.-F. Liu. *ApJ*, 836:200, February 2017. • Z. Guo, M. Gully-Santiago, and G. J. Herczeg. *ApJ*, 868:143, December 2018. • T.-O. Husser, S. Wende-von Berg, S. Dreizler, D. Homeier, A. Reiners, T. Barman, and P. H. Hauschildt. *A&A*, 553:A6, May 2013. • Kaisey Mandel and Eric Agol. *ApJ*, 580:L171–L175, Dec 2002. • A. McQuillan, T. Mazeh, and S. Aigrain. *The Astrophysical Journal Supplement Series*, 211:24, Apr 2014. • Brett M. Morris, Eric Agol, Leslie Hebb, and Suzanne L. Hawley. *AJ*, 156:91, Sep 2018. • Benjamin V. Rackham, Dániel Apai, and Mark S. Giampapa. *ApJ*, 853:122, Feb 2018. • G. Somers and M. H. Pinsonneault. *ApJ*, 807:174, July 2015. • H. R. Wakeford, N. K. Lewis, J. Fowler, G. Bruno, T. J. Wilson, S. E. Moran, J. Valenti, N. E. Batalha, J. Filippazzo, V. Bourrier, S. M. Hörst, S. M. Lederer, and J. de Wit. *AJ*, 157:11, Jan 2019.