EXOPLANETARY DETECTION BY MULTIFRACTAL SPECTRAL ANALYSIS

SAHIL AGARWAL, FABIO DEL SORDO, and JOHN S. WETTLAUER

1 Program in Applied Mathematics, Yale University, New Haven, CT, USA
2 Department of Astronomy, Yale University, New Haven, CT, USA
3 Department of Geology & Geophysics, Yale University, New Haven, CT, USA
4 Nordita, Royal Institute of Technology and Stockholm University, SE-10691 Stockholm, Sweden
5 Department of Mathematics, Yale University, New Haven, CT, USA
6 Department of Physics, Yale University, New Haven, CT, USA
7 Mathematical Institute, University of Oxford, Oxford, UK

Received 2016 September 19; revised 2016 November 2; accepted 2016 November 4; published 2016 December 20

ABSTRACT

Owing to technological advances, the number of exoplanets discovered has risen dramatically in the last few years. However, when trying to observe Earth analogs, it is often difficult to test the veracity of detection. We have developed a new approach to the analysis of exoplanetary spectral observations based on temporal multifractality, which identifies timescales that characterize planetary orbital motion around the host star and those that arise from stellar features such as spots. Without fitting stellar models to spectral data, we show how the planetary signal can be robustly detected from noisy data using noise amplitude as a source of information. For observation of transiting planets, combining this method with simple geometry allows us to relate the timescales obtained to primary and secondary eclipse of the exoplanets. Making use of data obtained with ground-based and space-based observations we have tested our approach on HD 189733b. Moreover, we have investigated the use of this technique in measuring planetary orbital motion via Doppler shift detection. Finally, we have analyzed synthetic spectra obtained using the SOAP 2.0 tool, which simulates a stellar spectrum and the influence of the presence of a planet or a spot on that spectrum over one orbital period. We have demonstrated that, so long as the signal-to-noise-ratio \( \geq 75 \), our approach reconstructs the planetary orbital period, as well as the rotation period of a spot on the stellar surface.

**Key words:** methods: analytical – methods: data analysis – methods: observational – methods: statistical – planets and satellites: detection – techniques: spectroscopic

1. INTRODUCTION

The last three decades have seen the birth of exo-planetary science. With the advent of various techniques, which include, but are not limited to, pulsar timing (Wolszczan & Frail 1992), Doppler measurements (Mayor & Queloz 1995), transit photometry (Charbonneau et al. 2000), micro-lensing (Beaulieu et al. 2006) and direct imaging (Chauvin et al. 2004), thousands of planets have been detected orbiting distant stars. A central focus is the detection of so-called Exo-Earths; Earth-like planets in terms of mass and radius, orbiting around a star at a distance (habitable zone) that, given sufficient atmospheric pressure, would allow for the existence of liquid water on its surface (e.g. Kopparapu et al. 2013, and references therein). The techniques that are most commonly used in discovering other planets are transit photometry (e.g., Lissauer et al. 2014, and references therein) or Doppler measurements (e.g., Mayor et al. 2014, and references therein). Recently, using the transit method, detection of nine candidates for habitable planets was announced that may fall within the habitable zones of their host stars (Anglada-Escudé et al. 2016; Morton et al. 2016).

While the combination of these approaches have provided an impressive range of observations, the detection of Earth analogs is a challenging problem. Indeed, the presence of instrumental and astrophysical noise are sources of uncertainty for such discoveries (e.g., Fischer et al. 2016). The fingerprints of an exo-Earth could easily be hidden in stellar noise, or stellar signals might mimic the presence of an exoplanet. Moreover, when such noise is modeled, there is a risk of introducing spurious signals in the analysis of data (e.g., Dumusque et al. 2015; Rajpaul et al. 2016).

Contemporary studies aim to understand and correctly evaluate the imprint of stellar activity on exoplanet detections (e.g., Lanza et al. 2007, 2011; Aigrain et al. 2012; Korhonen et al. 2015). Nonetheless, radial velocity and transit photometry methods can still produce false detections, especially when dealing with Earth-like planets, or planets characterized by a signal \( \leq 1 \text{ m s}^{-1} \) (Dumusque et al. 2016; Morton et al. 2016), which is, in part, due to the models upon which these methods are based. Such exoplanet evolution and stellar models over-fit parameters to the data, which turns out to be crucial when one aims to detect terrestrial-like planets (Dumusque et al. 2012; Rajpaul et al. 2016). These parameters include, but are not limited to, planet and stellar radius, their masses, eccentricity, impact factor, transit duration, transit ingress/egress duration, transit depth, orbital inclination, distance of the planet from the star, limb-darkening parameters (which themselves may vary with the law used), and shape of the transit curve. In some cases, instrumental signals such as the wobble of the instrument cluster aboard the Spitzer satellite or the latent charge build-up in the pixels (ramp) are also modeled (Grillmair et al. 2007; Todorov et al. 2014). Additionally, noise sources such as granulation over the stellar surface employ model fitting to estimate the effect of that noise on the data (Dumusque et al. 2012), stellar activity concurrent with the stellar rotation are modeled by fitting sine waves through the radial velocity data, and long term stellar activity, light contamination from near by stars, among others are parametrically modeled. This fitting of models has led to the introduction of spurious signals in the observed data and hence to false detections.
In order to fit these stellar models to data, one must begin by considering a particular system, for example an unblended eclipsing binary or a transiting planet. Hence, the list of models that the data can represent needs to be complete, which constitutes a weakness for these types of fitting schemes (Morton et al. 2016). Moreover, these methods cannot always distinguish whether the signal is instrumental or from stellar activity (Morton et al. 2016). Finally, because it affects the details of the fitting and hence the results, the observations must have a high signal-to-noise ratio (S/N).

A key aspect of the analysis of exo-planetary systems is the identification of periodic timescales. This is also central to the study of multi-planetary systems (e.g., Millholland et al. 2016), and often requires the combined use of the two most successful approaches in exoplanet searches; the transit and radial velocity methods. The most widely used method for identifying periodic timescales is the Lomb–Scargle periodogram, which is based on the assumption that the data can be interpreted as a sum of periodic signals. However, this technique is used on data that have been “filtered” using models for stellar noise and terrestrial atmospheric contamination. Hence, it is again possible to accumulate artifacts in the data through the use of such models. Many approaches have been developed by the exoplanetary science community to address these matters. For example, Dumusque et al. (2016) have compared state-of-the-art methods on simulated radial velocity data and concluded that the detection of planets below the 1 m s⁻¹ threshold is still controversial. Moreover, the extraction of radial velocities from spectra relies on cross-correlation techniques that are based on the use of spectral templates based on stellar models. Therefore, the determination of the Doppler shift remains a method-dependent challenge. For example, in the recent radial velocity study of Proxima Centauri (Anglada-Escudé et al. 2016), the use of the TERRA algorithm (Anglada-Escudé & Butler 2012) rather than the HARPS pipeline, has been crucial for the quantification of the signal.

Here, to extract the timescales that characterize a planetary system and stellar features without a priori assumptions about the data itself, we introduce a new approach to spectral analysis. Namely, in the spirit of the Langevin theory of Brownian motion, we quantify a signal coming from a star as, rather than using temporal-weighting, e.g., Agarwal et al. 2012, and references herein, which does not have a priori assumptions about the temporal structure of the data. The approach has four stages, which we describe briefly below.

1. We construct a non-stationary profile \( Y(i) \) of the original time series \( X_i \) as,

\[
Y(i) = \sum_{k=1}^{i} (X_k - \bar{X}_k), \quad \text{where } i = 1, \ldots, N. \tag{1}
\]

The profile is the cumulative sum of the time series and \( \bar{X}_k \) is the average of the time series \( X_1 \ldots X_k \).

2. This non-stationary profile is divided into \( N_s = \text{int}(N/s) \) non-overlapping segments of equal length \( s \), where \( s \) is an integer and varies in the interval \( 1 < s < N/2 \). Each value of \( s \) represents a timescale \( s \times \Delta t \), where \( \Delta t \) is the temporal resolution of the time series. The time series has a length that is rarely an exact multiple of \( s \), which is handled by repeating the procedure from the end of the profile and returning to the beginning, thereby creating \( 2N_s \) segments.

3. A point by point approximation \( \tilde{Y}_s(i) \) of the profile is made using a moving window, smaller than \( s \) and weighted by separation between the points \( j \) to the point \( i \) in the time series such that \( |i - j| \leq s \). A larger (or smaller) weight \( w_j \) is given to \( \tilde{Y}_s(i) \) according to whether \( |i - j| \) is small (large) (Agarwal et al. 2012). This approximated profile is then used to compute the variance spanning up \( (\nu = 1, \ldots, N_s) \) and down \( (\nu = N_s + 1, \ldots, 2N_s) \) the profile

\[\text{In regular MF-DFA, rather than using temporal-weighting, } n\text{th order polynomials } \tilde{Y}_s(j) \text{ are used to approximate } Y(i) \text{ within a fixed window, without reference to points in the profile outside that window.}\]
as
\[
\text{Var}(\nu, s) \equiv \frac{1}{s} \sum_{i=1}^{s} [Y([\nu - 1]s + i) - \bar{Y}([\nu - 1]s + i)]^2
\]
for \(\nu = 1, \ldots, N_s\) and
\[
\text{Var}(\nu, s) \equiv \frac{1}{s} \sum_{i=1}^{s} [Y(N - [\nu - N_s]s + i) - \bar{Y}(N - [\nu - N_s]s + i)]^2
\]
for \(\nu = N_s + 1, \ldots, 2N_s\). (2)

4. Finally, a generalized fluctuation function is obtained and written as
\[
F_q(s) = \left[\frac{1}{2N_s} \sum_{i=1}^{2N_s} \left(\text{Var}(\nu, s)\right)^{q/2}\right]^{1/q}.
\] (3)

The behavior of \(F_q(s)\) depends on the choice of time segment \(s\) for a given order \(q\) of the moment taken. The principal focus is to study the scaling of \(F_q(s)\) as characterized by a generalized Hurst exponent \(h(q)\) viz.,
\[
F_q(s) \propto s^{h(q)}.
\] (4)

When \(h(q)\) is independent of \(q\) the time series is said to be monofractal, in which case \(h(q)\) is equivalent to the classical Hurst exponent \(H\). For \(q = 2\), regular MF-DFA and DFA are equivalent (Kantelhardt et al. 2002), and \(h(2)\) can also be related to the decay of the power spectrum \(S\). If \(S(f) \propto f^{-\beta}\), with frequency \(f\) then \(h(2) = (1 + \beta)/2\) (e.g., Rangarajan & Ding 2000). For white noise \(\beta = 0\) and hence \(h(2) = 1/2\), whereas for Brownian or red noise \(\beta = 2\) and hence \(h(2) = 3/2\). The dominant timescales in the data set are the points where the fluctuation function \(\log_{10}F_2(s)\) changes slope with respect to \(\log_{10}s\). At each wavelength a crossover in the slope of a fluctuation function is calculated if the change in slope of the curve exceeds a set threshold, \(C_{th}\). Because the window length is constrained as \(1 < s \leq N/2\) (Zhou & Leung 2010), this analysis is limited to timescales of \(1 \leq t_{up} = N\Delta t/2\). Dobson et al. (1990) studied stellar features by focusing on time-series observations of H and K lines of Ca II. Their method constructs diagrams based on the concept of pooled variance, and has been subsequently used in other analyses of stellar activity (Donahue et al. 1997; Scholz & Eislöffel 2004; Lanza et al. 2006; Messina et al. 2016). Pooled variance is defined as the mean variance at a particular timescale, \(\tau_p\), by stepping through a time series in consecutive bins of size \(\tau_p\) and calculating the variance within each bin. A pooled variance diagram (PVD) plots this mean variance versus \(\tau_p\), thereby examining the timescale dependence of the mean variance.

The key differences between a PVD and our method are as follows. First, while a PVD computes a mean variance of the data itself, MF-TW DFA examines the variability relative to the temporally weighted fit to the profile of the data, thereby exploiting the intuition that points closer in time are more likely to be related than more distant points. Second, because the procedure of MF-TW DFA produces a smooth profile of the data, which is the core time series analyzed, it does not suffer from the intrinsic noise present in the data (see Figure 4 of Agarwal et al. 2012). Third, in MF-TW DFA the fluctuation functions are calculated for an arbitrary range of moments (up to ten in geophysical data; Agarwal et al. 2012), both positive and negative, thereby providing a rich tapestry of the temporal dynamics underlying variability, as well as explicit information about the processes producing them (e.g., pink or Brown noise processes). Finally, to the best of our knowledge PVD’s have not been used to quantify exoplanetary timescales.

3. DEMONSTRATING THE METHOD WITH THREE TYPES OF DATA

To test the method described in Section 2 we apply it to three different kinds of data. The first two are spectral observations of HD 189733b, a well-known transiting exoplanet discovered in 2005 (Bouchy et al. 2005). The third is a set of simulated stellar spectra affected by the Doppler shift induced by an orbiting planet. These simulated data are obtained with the Spot Oscillation And Planet (SOAP) 2.0 tool (Dumusque et al. 2014).

3.1. HD 189733b

First detected in 2005 (Bouchy et al. 2005), HD 189733b is a hot Jupiter orbiting the star HD 189733A in the constellation

---

**Table 1**

| Observation Date | Program ID |
|------------------|------------|
| Night 1          | 2006 Sep 07| 072.C-04888(E) |
| Night 2          | 2007 Jul 19| 079.C-0828(A)  |
| Night 3          | 2007 Aug 28| 079.C-0127(A)  |

**Table 2**

| AOR Key | Observation Date | Wavelength (µm) |
|---------|------------------|-----------------|
| 18245632 | 2006 Oct 21 | 7.4–14.0 |
| 20645376 | 2006 Nov 21 | 7.4–14.0 |
| 23437824 | 2008 May 24 | 7.4–14.0 |
| 23438080 | 2008 May 26 | 7.4–14.0 |
| 23438336 | 2008 Jun 02 | 7.4–14.0 |
| 23438592 | 2008 May 31 | 7.4–14.0 |
| 23438488 | 2007 Oct 31 | 7.4–14.0 |
| 23439104 | 2007 Nov 02 | 7.4–14.0 |
| 23439360 | 2007 Jun 26 | 7.4–14.0 |
| 23439616 | 2007 Jun 22 | 7.4–14.0 |
| 23440384 | 2008 Jun 09 | 5.0–7.5  |
| 23440640 | 2008 Jun 04 | 5.0–7.5  |
| 23440896 | 2007 Dec 07 | 5.0–7.5  |
| 23441152 | 2007 Nov 06 | 5.0–7.5  |
| 23441408 | 2007 Nov 11 | 5.0–7.5  |
| 23441664 | 2007 Nov 09 | 5.0–7.5  |
| 23441920 | 2007 Nov 24 | 5.0–7.5  |
| 23442176 | 2007 Nov 15 | 5.0–7.5  |
| 23439872 | 2007 Nov 04 | 13.9–21.3|
| 23440128 | 2007 Jun 17 | 13.9–21.3|
| 23442432 | 2007 Dec 10 | 19.9–39.9|
| 23442688 | 2007 Jun 20 | 19.9–39.9|

Note. A Astronomical Observation Request.
by severe weather (Triaud et al. 2009; Wytenbach et al. 2015). Each night is treated as an independent data set. On night 1, these spectra are taken approximately every 10.5 minutes, and on nights 2 and 3 every 5.5 minutes.

3.1.2. Spitzer Data

The Spitzer space mission observed secondary eclipses of HD 189733b. We use Basic Calibrated Data (.bcd) files from the Spitzer pipeline version 18.18.0, obtained with the optimal extraction tool in the SPICE software that employs the Optimal Extraction Algorithm of Horne (1986) to obtain the reduced 1D spectra from the observed images (see Table 2).

3.2. Simulations

The use of simulated data provides a good test and application of our method to radial velocity measurements, because we can control the input and thereby facilitate a clear interpretation of the results. We use data produced with the SOAP 2.0 tool (Dumusque et al. 2014). These spectra are generated by filling a grid of cells with a telluric cleaned solar spectrum from the National Solar Observatory, simulating a rotating star. Depending on the grid cell position, the spectrum is Doppler shifted to account for stellar rotation, and the intensity of each cell is weighted by a limb-darkening law. To simulate the presence of a spot, a typical spectrum of a solar spot is enclosed in a cell. Additionally, the intensity is reduced in the presence of a dark spot or increased in the presence of a bright plage. Such features follow the rotation of the star. The final integrated spectrum is the sum of the spectra from all of the cells.

First, we use stellar spectra in the absence of stellar spots and thus influenced solely by an orbiting planet. This allows us to focus on, for example, the “bare” effect of the Doppler shift on the spectra. Because we would like to examine orbital periods, we stack the 25 spectra corresponding to the orbital period of the planet to provide 8 orbital periods worth of data. The time units are arbitrary, so the analysis can be compared to a wide range of observations, provided they are evenly spaced in time. Moreover, since we are interested in the effect of noise on our data analysis, Gaussian white noise with a specific S/N per

Figure 1. Second moment of the fluctuation functions for (a) Night 1 (time resolution 10.602988 ± 0.072236 minutes), (b) Night 2 (time resolution 5.517202 ± 0.005683 minutes), and (c) Night 3 (time resolution 5.523740 ± 0.004415 minutes), with original wavelength resolution (plotted at alternate wavelengths). The abscissa (log_{10}c) is measured using the number of data points in the time series, 1 < s < N/2, where s represents a timescale s × ∆t, with ∆t being the temporal resolution of the time series. Different colors represent different wavelengths. In all of the fluctuation functions, wavelength generally increases in the positive y-direction. The straight blue line has a slope of 0.5, the slope of the fluctuation function denoting white noise. Shorter wavelengths show higher amounts of noise, which may be associated with atmospheric turbulence and/or telluric effects on Earth or the exoplanet.

Figure 2. Crossover times plotted for all wavelengths, for all three nights for HD 189733b. Only the 85 minute timescale is present for all three nights.
resolution element is then added at each time to obtain data sets with different S/Ns. This is important for the comparison of results obtained with this simulated data set to real observations from spectrographs characterized by different S/Ns.

Second, we use stellar spectra in the absence of a planet and thus influenced solely by a stellar spot. This allows us to focus on the “bare” effect of a spot, which can in principle mimic the Doppler shift in radial velocity measurements induced by a spot. Here again, as we have 25 spectra corresponding to one full rotation period of the star, we stack these to give 8 rotation periods worth of data, and then add Gaussian white noise as in the first (planet only) case discussed above. The spot covers 5% of the star, has a contrast of 663 K and it is set to rotate on the stellar equator (i.e., a latitude of 0°). The star rotates “equator- on,” meaning the rotational axis is at 90° relative to the line of sight. We have selected such values in order to deal with the simplest possible realistic observational case. For each of these spectra the continuum has been subtracted.

4. DISCUSSION

4.1. HARPS Analysis of HD 189733b

The 1-D spectra from the HARPS instrument provide the time series for flux at each of the wavelengths for each night separately. These are analyzed using MF-TWDF as described in Section 2 above. The second moment of the associated fluctuation functions are shown in Figures 1(a)–(c) for Nights 1, 2, and 3 respectively. In all of the fluctuation functions, wavelength generally increases in the positive y-direction.

The key aspects of the fluctuation functions are as follows. First, nearly all of these curves are parallel to each other, demonstrating that the spectra evolve in time with a similar
noisy behavior at all wavelengths. Second, those fluctuation functions that show deviations, do so principally at smaller wavelengths and can thus be ascribed to atmospheric interference from the Earth or the exoplanet, such as variability of air masses, telluric contamination and/or turbulent effects.

Third, for all nights and all the wavelengths we see a timescale of 85 minutes. Fourth, for times longer than 85 minutes, the dynamics for all wavelengths exhibit a white noise structure. In Figure 2 we plot all of the times at which the slope of the fluctuation functions changes—the crossover times—for the three nights for all of the wavelengths. The robustness of the 85 minute timescale is reflected by its presence on all nights, whereas other timescales are present for only one or two nights. Lines at shorter times only appear to be continuous due to the large number of points, but are in fact quite noisy, as shown by the fluctuation functions. Possible origins of the other timescales include turbulent and convective processes in Earth’s and the exoplanet’s atmosphere, stellar activity or instrumental noise. It is important that we have identified these scales here and this provides a foundation for systematic examination of them in future studies as a possible means of systematically filtering them out.

To understand this 85 minute timescale, we study the following parameters of the planet HD 189733b, as found from previous studies: (1) the ratio of the planetary ($R_p$) to the stellar ($R_s$) radii, \((R_p/R_s)^2 = 0.02391 \pm 0.00007\) (Torres et al. 2008), and (2) the duration of the transit of the planet in front of the star \(\tau_{14} = 0.07527 \pm 0.00037\) days \(\approx 108.4\) minutes, or the time between first and last “contact” of the planet and the star (Triaud et al. 2009) (see Figure 3).

Due to the time-resolution of the data for each night, as well as the total length of these time series, we are not able to observe the shortest and the longest timescales in the system. The 85 minute timescale corresponds to the interval between the second and third contact of the planet, \(t_{23}\), i.e., the period during which the planet is completely between the observer and the star (Figure 3). If we assume that the planet is transiting in the equatorial plane of the star, and hence the impact parameter is \(b = 0\), we calculate this time by tracking the center of the planetary disk across the stellar disk as

\[
\tau_{14} = \frac{2(R_s + R_p)}{2(R_s - R_p)},
\]

from which we obtain \(\tau_{23} \approx 0.0551\) days \(= 79.659\) minutes, using \(R_p/R_s\) from Torres et al. (2008). Within the resolution of the time series, this value of \(\tau_{23}\) is consistent with the 85 minute timescale we have found. The timescale of the full transit \(\tau_{14}\) would generally be the dominant timescale of the system. Because the Rossiter–McLaughlin effect for this system operates on the same timescale (Winn et al. 2006), it might underlie the spectral modifications detected here. Our method only works up to \(N/2\), namely half the duration of the observational time, \(t_{\text{up}} = N \Delta t/2\), and thus due to the length of the time series from the HARPS data, the method does not allow \(\tau_{14}\) to be detected. To test how wavelength-resolution affects our results, we degrade the resolution of the spectra in wavelength and repeat our analysis. We decrease the resolution.
Figure 10. Crossover times plotted for all wavelengths, with $C_{\text{in}} = 0.08$, for the simulated SOAP 2.0 spectra in presence of an orbiting planet. We show results for different S/N. (a) No Noise, (b) S/N = 1000, (c) S/N = 100, (d) S/N = 500, (e) S/N = 200, (f) S/N = 150, (g) S/N = 250, (h) S/N = 100, (i) S/N = 80, (j) S/N = 75, (k) S/N = 50, (l) S/N = 10, (l) S/N = 1.
Figure 11. Crossover times plotted for all wavelengths, with $C_0 = 0.01$, for the simulated SOAP 2.0 spectra in presence of an orbiting planet. We show results for different S/N. (a) No Noise, (b) S/N = 1000, (c) S/N = 500, (d) S/N = 250, (e) S/N = 200, (f) S/N = 150, (g) S/N = 100, (h) S/N = 50, (i) S/N = 10, (j) S/N = 1.
by a factor 100 by using the average of each 100 points, and plot the fluctuation functions for the degraded spectra from Night 1 in Figure 4. Clearly, the structure at longer timescales is unchanged, whereas the noisy behavior at short wavelengths is less evident than in the analysis of high resolution spectra, showing that the averaging acts as a crude high-pass filter. Importantly, the 85 minute scale persists as does the general behavior of the fluctuation functions for this reduction in resolution.

4.2. Spitzer Analysis of HD 189733b

Because stars can only be observed at night, all Earth-based instruments provide a strong constraint on the detection of timescales. Moreover, telluric contamination is a well known problem that is also evident in our method at the shorter timescales in the HARPS data. To bypass these limitations, we examine data from the Spitzer mission, which observed HD189733b for 22 nights, looking at the secondary eclipses of the planet, i.e., when the planet is behind the star.

Although the data from Spitzer is of low resolution in wavelength space, this does not affect the identification of robust timescales, which we have demonstrated in the case of the HARPS data by comparing the results using full resolution data with those from degraded data. Regardless of the wavelength resolution, the time resolution as well as the total duration make the Spitzer data sets compelling.

As was done in Figure 2, we plot these crossovers as a function of wavelength in Figure 5. It is clear that relative to the HARPS data the Spitzer data is substantially more noisy. We note that typically the raw data are filtered to remove the noisy characteristics, but given experience with other systems (Agarwal et al. 2012; Agarwal & Wetlaufer 2016) we take the perspective that noise can be an essential source of information.

The four prominent timescales in this data are: (1) $\tau_{12} = 15.4804 \pm 3.7660$ minutes, (2) $55.0966 \pm 5.8851$ minutes, (3) $\tau_{23} = 87.0947 \pm 2.7888$ minutes, and (4) $\tau_{14} = 118.9671 \pm 5.5764$ minutes, where the uncertainty is one standard deviation about the mean. First, the $55.0966 \pm 5.8851$ minute timescale is the pointing wobble in the Infrared Array Camera of the Spitzer telescope, which is due to the battery heater in the telescope (Grillmair et al. 2007; Spitzer 2010). Second, the 85 minute timescale obtained from the HARPS data is within one standard deviation of $\tau_{23}$ and hence is robust. The other two timescales are related to the transit of the planet behind the star, namely the secondary eclipse.

The necessary and sufficient condition for these timescales to represent the transit (Figure 3) is

$$\tau_4 = \tau_{23} + 2 \times \tau_{12}. \tag{6}$$

It is evident that, within one standard deviation, Equation (6) is satisfied for the above timescales. Because $\tau_4 < \tau_{up}$, the Spitzer data sets span a sufficient time to allow the detection of the secondary eclipse timescale $\tau_4$. Using the values for $\tau_4$ and $\tau_{14}$, we can then estimate the ratio of the radius of the planet to that of the star by rearranging Equation (5) to obtain

$$\frac{R_p}{R_*} = \frac{\tau_{14} - \tau_{23}}{\tau_{14} + \tau_{23}} = 0.1543 \pm 0.0283. \tag{7}$$

Therefore, we have shown here that without the use of any fitting of model parameters such as the epoch of mid-transit, orbital period, fractional flux deficit, total duration of transit, the impact parameter of the planet’s path across the stellar disc, the transit depth, the shape parameter for transit, the transit ingress/egress times, and many more (Collier Cameron et al. 2007; Morton et al. 2016), we can now calculate most of these from the timescales alone (Seager & Mallén-Ornelas 2003). We note that, to within the precision of $\pm 0.0283$, the result in Equation (7) is the same as that found by Bouchy et al. (2005) and the average of that found by Pont et al. (2007).

In this analysis, we did not take into account the possibility that the planet may be transiting with a non-zero impact parameter ($b \neq 0$, where $b = 0$ implies the planet is transiting the equator of the star from the observer’s perspective). This would add another unknown in Equation (7). By using the simple geometry shown in Figure 6, another equation can be derived that also involves $\tau_{12}$. Each of these two second order equations can be solved to determine the relationship between the ratio of radii and the impact parameter, as follows.

First, we can write $D_{12} = D_{34}$, the distance between the first-second or third-fourth contact from the Pythagorean theorem as

$$D_{12} = D_{34} = \sqrt{(R_* + R_p)^2 - b^2 R_*^2} - \sqrt{(R_* - R_p)^2 - b^2 R_*^2}. \tag{8}$$

Second, $D_{23}$, the distance between the second and third contact is

$$D_{23} = 2\sqrt{(R_* - R_p)^2 - b^2 R_*^2}, \tag{9}$$

and finally we have

$$D_{14} = 2\sqrt{(R_* + R_p)^2 - b^2 R_*^2}. \tag{10}$$

Therefore,

$$\frac{\tau_4}{D_{14}} = \frac{\tau_{23}}{D_{23}}. \tag{11}$$
which, upon substitution of Equations (9) and (10), gives
\[ \tau_{23} = \frac{\tau_{14}}{\rho} \times \sqrt{\frac{1 + \rho^2 - 2\rho + b^2}{1 + \rho^2 + 2\rho - b^2}}, \]
(12)
where \( \rho = \frac{R_p}{R_*} \). Similarly, Equations (8) and (10) give
\[ \tau_{12} = \frac{\tau_{14}}{2} \left( 1 - \sqrt{\frac{1 + \rho^2 - 2\rho + b^2}{1 + \rho^2 + 2\rho - b^2}} \right) \]
(13)
and thus Equations (12) and (13) can be used to calculate \( \rho \) and \( b \), simultaneously. As a check, Equations (12) and (13) combine to give Equation (5), and (12) becomes Equation (7) in the case \( b = 0 \).

In Figures 7 and 8 we show the crossovers from all the available data sets, as shown in Table 2. Additionally, Figure 8 includes those data sets that are not used in other studies due to low S/N, or other issues (Todorov et al. 2014). Moreover, we see the clear emergence of the significant timescales discussed above from all of these data sets. Importantly, we note the amount of noise in these data sets, which as we will see below can be a source of information for the robust estimation of these timescales.

4.3. SOAP Simulated Data

Thus far, we have focused our analysis on data observed during a primary or secondary eclipse of an exoplanet. However, these measurements are not always available since only a small percentage of exoplanet transits occur between the observer and the host star. Instead, all the exoplanets produce an effect on the radial velocity of the host star.
Current technology only allows detection of the red and blue shifts of the spectrum when the motion of the star is sufficiently large, and although modern high-resolution spectrographs can attain a resolution of the order of 1 m s\(^{-1}\), detections of small planets often remain controversial. Moreover, these data are contaminated by various sources of noise, such as instrumental noise, atmospheric noise, stellar noise, and telluric absorption lines. Presently, the observations are fit to a radial velocity curve, which may or may not be robust and have a non-negligible false detection rate (see Section 4.3.2. of Fischer et al. 2016).

To address this problem, we test the detectability of timescales related to the radial velocity effect with our method. We start from the analysis of a well defined case using data from the SOAP 2.0 simulations. This allows us to study the effect of noise as well as the robust estimation of crossover times, which we know to be present in the data set.

The original (“noise-free”) data describe the shifts in the spectrum of the star with a radial velocity of amplitude 40 ms\(^{-1}\), due to a planet orbiting around it. We add Gaussian white noise to the spectra at each time to get a specific S/N, and thereby obtain 12 such data sets for S/N of 1, 10, 50, 75, 80, 100, 150, 200, 250, 500, and 1000. Figure 9 shows the second moment of the fluctuation functions for all the wavelengths in the reduced spectra without noise, and subfigures in Figure 10 show the crossover timescales extracted for the different noise cases. As explained in Section 2.1, we evaluate these crossovers using a threshold value \(C_{\text{th}}\) for the slope change. The plots in Figure 10 use a threshold value \(C_{\text{th}} = 0.08\). Whence, we are able to extract the exact timescale for the orbital period of the exoplanet. Even for a S/N = 150, the methodology is efficient and robust against noise, and we extract the correct timescale. As the signal quality is degraded further, we see the noise starting to affect the calculated crossovers with scatter above the actual value. Finally, as we further degrade the signal quality the crossover times disappear altogether. As the noise begins to dominate, the signal becomes white. We know this quantitatively from our analysis, as fluctuation functions with a constant slope of 0.5 are, by definition, white noise processes (e.g., Figure 1). To examine how to capture timescales when the signal quality is poor, we decrease the threshold to \(C_{\text{th}} = 0.01\) (Figure 11). Even this threshold is able to approximately capture (31 time units) the orbital period, but it also detects other multiple timescales, which may be spurious. The important point here is the ability to study how noise affects these timescales; as the S/N is decreased timescales remain robust, but become harmonics of the robust scales for a further increase in the S/N. These may be due to a form of stochastic resonance between the threshold, radial velocity measurements and noise. The value of threshold is thus crucial.

In actual data, one cannot calculate this threshold for all the wavelengths separately and for each night, and hence one value is chosen in accordance with the observed noise characteristics. Presently, use of periodograms and fitting to sine curves is performed to model the radial velocity curves from the observations. The above analysis shows how noise can lead to spurious estimation of orbital periods, and thus potentially result in spurious detection of exoplanets.
In the same manner as above, we have analyzed the case in which no planet is orbiting around the star, but a spot covers 5% of its surface. In a similar fashion we have analyzed 8 rotation periods of the star and reconstructed its rotational period (Figures 12–14). The use of this method to distinguish stellar features from planets will be subject of future work.

5. CONCLUSION

We have presented and tested a new multi-fractal approach for the analysis of exo-planetary spectral observations. The goal is to use a fit-free procedure to identify robust timescales associated with the exo-planetary orbital motion around a host star, as well as to detect timescales associated with stellar features. With these timescales in hand, one can compute key system parameters such as the ratio of the size of the planet to that of the star and the latitude of transit, without use of stellar evolution models, data fitting, noise filtering, and the additional wide variety of other assumptions about the system that are typically made. The concept of the approach is to take an agnostic (or model free) view of the observed spectral structure. The method makes no a priori assumptions about the temporal structure in any observed spectra and makes use of only one number, the generalized Hurst exponent, the value of which underlies the identification of the key timescales. We have reconstructed the primary and secondary transit times of the exoplanet HD 189733b, using data from both ground-based (HARPS spectrograph) and space-based (Spitzer) observations.

Using the SOAP 2.0 tool, which simulates a stellar spectrum and how it can be affected by the presence of a planet across one orbital period, our method is further tested in the context of measuring planetary orbital motion via Doppler shift detection. Because the SOAP tool can also simulate the presence of a spot on the stellar surface across one rotational period of the star, we have demonstrated that our approach reconstructs the planetary orbital period, as well as the rotation period of a spot covering 5% of the stellar surface. Moreover, we have tested the analysis with a wide range of S/Ns. We can reconstruct: (a) the period of a planet producing a 40 m s⁻¹ Doppler shift of the stellar spectrum, and (b) the rotational period of the star, based on the presence of a spot on its surface, provided that the S/Ns are ≥75 or ≥30 respectively. Importantly, to avoid introduction of errors arising from intrinsic irregularities in the analyzed time series, this method has the highest fidelity when observations are carried out at sufficiently high frequency that the time-difference between each measurement is less than the shortest relevant timescale that may be present in the system.

In conclusion, this method based on Multi-fractal Temporally Weighted Detrended Fluctuation Analysis of time series is a robust way to measure planetary orbital motion. It provides a fertile framework to examine other data sets and to explore trying to systematically distinguish stellar noise from planetary motion.

The authors thank the referee for comments that have helped to improve our paper, Xavier Dumusque for providing the SOAP 2.0 data used for this work and Debra Fischer and Allen Davis for discussions on this and other detection schemes. S.A. and J.S.W. acknowledge NASA Grant NNH13ZDA001N-CRYO for support. F.D.S. acknowledges the Swedish Research Council International Postdoc fellowship for support. J.S.W. acknowledges Swedish Research Council grant no. 638-2013-9243 and a Royal Society Wolfson Research Merit Award for support.

REFERENCES

Agarwal, S., Moon, W., & Wentura, J. S. 2012, RSPSA, 468, 2416
Agarwal, S., & Wentura, J. S. 2016, Ph.D., 380, 142
Aigrain, S., Pont, F., & Zucker, S. 2012, MNRAS, 419, 3147
Anglada-Escudé, G., Amado, P. J., Barnes, J., et al. 2016, Nat, 536, 437
Anglada-Escudé, G., & Butler, R. P. 2012, ApJS, 200, 15
Beaulieu, J.-P., Bennett, D. P., Fouqué, P., et al. 2006, Natur, 439, 437
Bouchy, F., Udry, S., Mayor, M., et al. 2005, A&A, 444, L15
Charbonneau, D., Brown, T. M., Latham, D. W., & Mayor, M. 2000, ApJL, 529, L45
Chauvin, G., Lagrange, A.-M., Dumas, C., et al. 2004, A&A, 425, L29
Collier Cameron, A., Wilson, D. M., West, R. G., et al. 2007, MNRAS, 380, 1230
Dobson, A. K., Donahue, R. A., Radick, R. R., & Kadlec, K. L. 1990, in ASP Conf. Ser. 9, Cool Stars, Stellar Systems and the Sun, ed. G. Wallerstein (San Francisco, CA: ASP), 132
Donahue, R. A., Dobson, A. K., & Baliunas, S. L. 1997, SoPh, 171, 211
Dumusque, X., Boisse, I., & Santos, N. C. 2014, ApJ, 796, 132
Dumusque, X., Borsa, F., Damasso, M., et al. 2016, arXiv:1609.03674
Dumusque, X., Pepe, F.,Lovis, C., et al. 2012, Natur, 491, 207
Dumusque, X., Pepe, F.,Lovis, C., & Latham, D. W. 2015, ApJ, 808, 171
Fischer, D. A., Anglada-Escude, G., Arriagada, P., et al. 2016, PASP, 128, 066001
Grillmair, C. J., Charbonneau, D., Burrows, A., et al. 2007, ApJL, 658, L115
Horne, K. 1986, PASP, 98, 609
Kantelhardt, J. W., Zschiegner, S. A., Koscielny-Bunde, E., et al. 2002, PhyA, 316, 87
Kopparapu, R. K., Ramirez, R., Kasting, J. F., et al. 2013, ApJ, 765, 131
Korhonen, H., Andersen, J. M., Piskunov, N., et al. 2015, MNRAS, 448, 3038
Lanza, A. F., Boisse, I., Bouchy, F., Bonomo, A. S., & Moutou, C. 2011, A&A, 533, A44
Lanza, A. F., Bonomo, A. S., & Rodonó, M. 2007, A&A, 464, 741
Lanza, A. F., Plisso, N., Rodonó, M., Messina, S., & Cutispoto, G. 2006, A&A, 455, 595
Lissauer, J. J., Dawson, R. I., & Tremaine, S. 2014, Nat, 513, 336
Mayor, M., Lovis, C., & Santos, N. C. 2014, Nat, 513, 328
Mayor, M., Pepe, F., & Queloz, D. 2003, Msngr, 114, 20
Mayor, M., & Queloz, D. 1995, Natur, 378, 355
Messina, S., Parirah, P., Biazzo, K., et al. 2016, MNRAS, 457, 3372
Millholland, S., Wang, S., & Laughlin, G. 2016, ApJL, 823, L7
Morton, T. D., Bryson, S. T., Coughlin, J. L., et al. 2016, ApJ, 822, 86
Pont, F., Gilliland, R. L., Moutou, C., et al. 2007, A&A, 476, 1347
Rajpaul, V., Aigrain, S., & Roberts, S. 2016, MNRAS, 456, L6
Rangarajan, G., & Ding, M. 2000, PrRV, 61, 4991
Scholz, A., & Eislöffel, J. 2004, A&A, 419, 249
Seager, S., & Mullén-Oroelos, G. 2003, ApJ, 585, 1038
Spitzer, 2010, Modification to Observed Pointing Wobble During Staring Observations, Tech. Rep., Caltech
Todorov, K. O., Deming, D., Burrows, A., & Grillmair, C. J. 2014, ApJ, 796, 100
Torres, G., Winn, J. N., & Holman, M. J. 2008, ApJ, 677, 1324
Triaud, A. H. M. J., Queloz, D., Bouchy, F., et al. 2009, A&A, 506, 377
Winn, J. N., Johnson, J. A.,Marcy, G. W., et al. 2006, ApJL, 653, L69
Wolszczan, A., & Frail, D. A. 1992, Natur, 355, 145
Wytenbach, A., Ehrenreich, D., Lovis, C., Udry, S., & Pepe, F. 2015, A&A, 577, A62
Zhou, Y., & Leung, Y. 2010, JSMTE, 6, 06021