A cautionary tale: limitations of a brightness-based spectroscopic approach to chromatic exoplanet radii


Published in:
Astronomy & Astrophysics

Document Version:
Publisher's PDF, also known as Version of record

Queen's University Belfast - Research Portal:
Link to publication record in Queen's University Belfast Research Portal

Publisher rights
Reproduced with permission from Astronomy & Astrophysics, © ESO 2017

General rights
Copyright for the publications made accessible via the Queen's University Belfast Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The Research Portal is Queen's institutional repository that provides access to Queen's research output. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact openaccess@qub.ac.uk.
LETTER TO THE EDITOR

A cautionary tale: limitations of a brightness-based spectroscopic approach to chromatic exoplanet radii

H. M. Cegla1,2,3, C. Lovis2, V. Bourrier2, C. A. Watson1, and A. Wytenbach2

1 Astrophysics Research Centre, School of Mathematics & Physics, Queen’s University Belfast, University Road, Belfast BT7 1NN, UK
2 Observatoire de Genève, Université de Genève, 51 chemin des Maillettes, 1290 Versoix, Switzerland e-mail: h.cegla@unige.ch
3 Swiss National Science Foundation NCCR-PlanetS CHEOPS Fellow, Switzerland

Received 31 October 2016 / Accepted 23 December 2016

ABSTRACT

Determining wavelength-dependent exoplanet radii measurements is an excellent way to probe the composition of exoplanet atmospheres. In light of this, Borsa et al. (2016) sought to develop a technique to obtain such measurements by comparing ground-based transmission spectra to the expected brightness variations during an exoplanet transit. However, we demonstrate herein that this is not possible due to the transit light curve normalisation necessary to remove the effects of the Earth’s atmosphere on the ground-based observations. This is because the recoverable exoplanet radius is set by the planet-to-star radius ratio within the transit light curve; we demonstrate this both analytically and with simulated planet transits, as well as through a reanalysis of the HD 189733 b data.

Key words. methods: data analysis – planets and satellites: atmospheres – planets and satellites: fundamental parameters – planets and satellites: individual: HD 189733 b – techniques: radial velocities – techniques: spectroscopic

1. Introduction

Transmission spectroscopy is an essential tool for characterising the atmospheres of transiting exoplanets (see e.g. Charbonneau et al. 2002; Pont et al. 2013; Madhusudhan et al. 2014; Sing et al. 2016, and references therein). Snellen (2004) demonstrated that narrowband exoplanet features (e.g. sodium) could be probed by analysing the shape of the stellar absorption lines as a planet occults its host star, that is, by studying the planet occulted stellar absorption spectrum normalised; they can then be multiplied by a transit light curve to allow one to study the local, occulted stellar profiles directly by subtracting the in-transit profile from the out-of-transit profile (see, e.g. Cegla et al. 2016a). Since the transit light curve is dependent on the planet-to-star radius ratio, it sets the planet radius that is recoverable when examining the brightness ratio between the local and out-of-transit profiles (meaning one cannot recover new radius variations following B16).

In this study, we first break down the physical implications of the B16 technique, and then simulate the planet transit of HD 189733 b to illustrate the impact of the technique’s shortcomings on the measured planet radius; we also reapply the B16 technique to HARPS data, with a more rigorous error propagation. In Sect. 2, we demonstrate how the choice of the transit light curve normalisation sets the recoverable planet radius. We present the simulated planet transits, and our reanalysis of the HARPS data in Sect. 3, and show how an underestimation of the errors can lead to spurious claims of planetary radius variations. Finally, we summarise our conclusions in Sect. 4.

2. Limitations of a brightness-based approach to chromatic exoplanet radii

2.1. The Borsa et al. 2016 approach

In B16, the authors attempted to study passband-dependent planet radius variations by averaging together the cross-correlation functions (CCFs) for subsets of HARPS spectral orders. All of these CCFs, regardless of passband, were continuum normalised and then further scaled using a Mandel & Agol (2002) transit light curve based on the system parameters determined in the full HARPS passband, except for passband dependent limb darkening coefficients.

For each passband, they created master out-of-transit CCFs by averaging together all the individual out-of-transit CCFs in
a given night, and then subtracting the out- from the in-transit data to obtain CCFs for the starlight occulted by the planet. The authors then fitted Gaussian functions to the respective CCFs, presumably to act as a proxy for the integrated flux within, and argued that the ratio between the areas (determined from the Gaussian fits) of the local to the out-of-transit CCFs could serve as a measure of the brightness of the missing starlight, $\beta$, occulted during the in-transit observations.

The authors then compared this empirical $\beta$ to the expected brightness ratio based on an approximation to the solution for integrating the starlight behind the planet given the particular geometry of the system (and assuming a particular function for the limb darkening). Since this solution was dependent on the planet radius, the authors argued that they were able to disentangle a planet radius measurement for each passband.

We demonstrate in Sect. 2.2 that the technique outlined by B16 can only recover the planetary radius set by the transit light curve used in the initial normalisation. However, we emphasise that such a limitation is set due to a flux-based approach and would not be present in an RV-based approach, such as that in Di Gloria et al. (2015) or that in traditional line profile tomography (i.e. following the Collier Cameron et al. 2010, formulation).

2.2. Transit light curve normalisation

B16 started by determining the area of their out-of-transit master CCFs, $A_{\text{out}}$, for each passband. Since the out-of-transit master CCFs ($CCF_{\text{out}}$) were continuum normalised their area is equal to their equivalent width, $EW_{\text{out}}$, that is,

$$A_{\text{out}} = EW_{\text{out}}.$$  \hspace{1cm} (1)

The authors also measured the areas of the local CCFs behind the planet, $A_{\text{loc}}$. Since the in-transit CCFs ($CCF_{\text{in}}$) are normalised by the transit light curve, the area of the local CCF ($CCF_{\text{loc}}$) is

$$A_{\text{loc}} = (1 - f_{\text{lc}}) \cdot EW_{\text{loc}} = \beta_{\text{lc}} \cdot EW_{\text{loc}},$$ \hspace{1cm} (2)

where $f_{\text{lc}}$ is the flux from the light curve, $EW_{\text{loc}}$ is the equivalent width of the local CCF, and $\beta_{\text{lc}}$ is the fraction of starlight occulted by the planet under the assumptions of the transit light curve used in the normalisation.

In an ideal case, where the local stellar photospheric profiles can be represented by constant Gaussian functions (assumed both above and in B16), then the only difference between the local and disc-integrated out-of-transit CCFs is the broadening by stellar rotation present in the disc-integrated observations. Since the rotational broadening preserves the equivalent width, then

$$EW_{\text{out}} = EW_{\text{loc}} = EW_{\text{in}},$$  \hspace{1cm} (3)

where the $EW_{\text{out}}$ and the $EW_{\text{loc}}$ are not only equal to each other but are also equal to the in-transit equivalent width, $EW_{\text{in}}$ (since it is simply a summation of local profiles of equal EW), and the ratio of the areas becomes:

$$\frac{A_{\text{loc}}}{A_{\text{out}}} = \beta_{\text{lc}}.$$ \hspace{1cm} (4)

Hence, under the above assumptions one can only recover the planet radius injected into the model transit light curve, regardless of which passband is studied. This is because the transit light curve normalisation effectively sets the area of the local profile. We note, that the transit light curve normalisation for the above follows Cegla et al. (2016a, hereafter C16), where

$$CCF_{\text{loc}} = CCF_{\text{out}} - f_{\text{lc}} \cdot CCF_{\text{in}},$$ whereas, the transit light curve normalisation in B16 was

$$CCF_{\text{B16}}^\text{loc} = CCF_{\text{in}} + f_{\text{lc}} \cdot CCF_{\text{out}} + (1 - f_{\text{lc}})$$ \hspace{1cm} (5)

(Borsa, priv. comm.), such that

$$CCF_{\text{B16}}^\text{loc} = -CCF_{\text{loc}}^\text{C16} + \beta_{\text{lc}}.$$ \hspace{1cm} (6)

where $CCF_{\text{loc}}^\text{C16}$ is the local CCF obtained following C16. Since the continuum is a free parameter in the Gaussian fits, the ratio of the areas is still equal to that in Eq. (4).

As such, the technique implemented by B16 represents a circular argument. Moreover, if one had the broadband photometry necessary for the correct spectral normalisation, then the planet radii could be determined directly from the light curves alone.

We stress that a solely brightness-based approach to transmission spectroscopy uses only the equivalent width, and therefore precludes any retrieval of information on exoplanet radii. It is the inclusion of the spectral dimension that is necessary to determine $R_p$ (i.e. utilising the Doppler information, as is done in Collier Cameron et al. 2010 and Di Gloria et al. 2015).

3. Systematic effects on chromatic radius measurements

To try to understand how B16 obtained results mimicking Rayleigh scattering, we simulated the transit of HD 189733b and applied their technique. To demonstrate that we could recover our model inputs, we also present results wherein we applied the normalisation from C16 with the correct planet radii for each passband; in doing so we discovered the approximations for $\beta$ in B16 underestimated the planet radii, and thus we also explored a numerical approach for calculating this brightness ratio that was more accurate than the approximation used in B16.

3.1. Analytical brightness ratio approximation

Unfortunately, integrating the limb darkened brightness underneath a planet lying off stellar disc centre is not straightforward and even approximate analytical expressions are quite complex (especially if considering ingress and egress regions). For this reason, B16 used the $\beta$ formalism presented in Collier Cameron et al. (2010), which was based on the analytical approximations of Ohta et al. (2005) for the RM effect. Therein the brightness ratio for a fully in-transit planet (i.e. no ingress or egress regions were considered) was defined, under the standard linear limb darkening law, as:

$$\beta \approx \frac{1 + u_1}{1 - u_1/3} \mu,$$ \hspace{1cm} (7)

where $R_*$ is the stellar radius, $R_p$ is the planet radius, $u_1$ is the linear limb darkening coefficient, and $\mu$ is the centre-to-limb planet position. We note that $\mu = \cos(\theta) = \sqrt{1 - \frac{x_p^2 + y_p^2}{R_*^2}}$, where $\theta$ is the centre-to-limb angle, and $x_p, y_p$ is the centre of the planet (see Collier Cameron et al. 2010; Cegla et al. 2016a, for details). We also note that Ohta et al. (2005) state that the accuracy of this approximation diminishes with increasing $R_p/R_*$, arguing that the additional terms in the analytical solution contribute to ~1% if $R_p/R_* \lesssim 0.1$ and up to few percent if $R_p/R_* \sim 0.3$. Given that $R_p/R_*$ is only predicted to vary a few percent in wavelength for particular atmospheric characteristics, such as Rayleigh scattering, this approximation may inject systematic errors that could be misinterpreted as having a physical origin (even if the light curve normalisation is done correctly).
3.2. Numerical brightness ratio approximation

To investigate the impact of the accuracy of Eq. (7), we decided to calculate $\beta$ numerically. For this approach, the in-transit starlight blocked by the planet is still defined as $\beta = F_{\lambda}/F_\star$, with $F_\star$ and $F_{\lambda}$ defined as the fluxes of the total stellar disc and the stellar disc under the planet, respectively. Note the observed brightness of the un-occulted star can be analytically determined by integrating a given limb darkening law over the projected stellar disc; for a linear limb darkening this is

$$F_\star = R_\star^2 \int_0^{\pi/2} \int_0^1 I(\mu) \mu \mathrm{d}\mu \mathrm{d}\phi = \pi R_\star^2 \left[ 1 - \frac{\mu^2}{3} \right],$$

(8)

where $\phi$ is the azimuthal angle. As previously stated, calculating the flux behind the planet analytically is not trivial. Hence, we calculated $F_{\lambda}$ numerically by constructing a square stellar grid with a width of $2R_p$ centred about the planet position $(x_p, y_p)$, with $n$ equal steps in the vertical and horizontal direction. Contributions from steps that did not lie beneath the planet and/or on the stellar disc were excluded. Thus, we approximated the flux behind the planet as

$$F_{\lambda} \approx \sum_{x,y} \left( \frac{2R_p}{n} \right) I_{by},$$

(9)

where $I_{by}$ is the limb darkened intensity at a given position in the aforementioned grid and $(2R_p/n)^2$ is the corresponding area.

Our aim was to try to recover $R_p$ as a function of wavelength, wherein the injected $R_p$ was only used to construct the correct light curves (acting as if we had simultaneous multi-colour photometry). Hence, when trying to recover $R_p$, we started with the broadband planet radius and then allowed it to vary by up to ±0.005 $R_p$ in steps of 0.0001 $R_p$. The recovered planet radius then corresponded to the planet radii that minimised the difference between $\beta$ and $A_{\text{loc}}/A_{\text{out}}$.

3.3. Simulated star-planet system

We used the simulated stellar grid of Cegla et al. (2015, 2016b) and injected into each grid cell a Gaussian profile with a full-width half-maximum (FWHM) of 5 km s$^{-1}$ (note this width is similar to the expected value for the stellar photosphere). In the simulated star we did not consider any astrophysical effects (i.e. granulation or starspots etc.) other than rigid body stellar rotation, which was set to the value obtained by C16, 3.25 km s$^{-1}$. We also assumed an edge-on ($i_\star = 90^\circ$) aligned orbit.

The transit was sampled in 21 equal steps in phase from −0.02 to 0.02, centred about mid-transit, with an additional sample at phase = 0.03 to serve as a completely out-of-transit reference. We simulated a transit for each of the seven passbands (from 400–700 nm) used in B16, and applied a linear limb darkening using the coefficients (for each passband) these authors provided. For each of the seven passband transits we injected a planet with a constant radius equal to the value assumed by B16 for the whole HARPS passband ($R_p = 0.1581 R_\star$), hereafter referred to as the broadband $R_p$), but varied the limb darkening accordingly.

For this set of transits, we tested the impact of the transit light curve normalisation. In the first case, we followed the procedure in B16, and in the second case we normalised the data following C16 and used the numerical approximation in Sect. 3.2 to estimate $R_p$. Examining the first case allowed us to examine any errors introduced using the $\beta$ approximation and/or the B16 normalisation. On the other hand, the second case offered a test case to ensure we could recover the model inputs.

For a second test, we repeated the above, but varied the radius of the simulated planet; for this we selected $R_p$ equal to the values reported by B16 for each passband. Again we tested two cases: first following the B16 procedure (where the light curve limb darkening varies in each passband, but the light curve radius remains fixed at the broadband $R_p$), and the second using...
the C16 normalisation (we note the assumed light curve has the correct \( R_p \) here) and the numerical approximation in Sect. 3.2.

We then examined the recovered \( R_p \) as a function of stellar disc position, and found only a slight dependence on disc position when following B16. However, if one point each at the ingress and egress regions were included then the dependence on disc position was strong, and including such data would systematically decrease the recovered \( R_p \) (as the \( \beta \) formulation is not valid in these regions).

Moreover, we found that, regardless of the B16 or C16 normalisation, using the \( \beta \) approximation always underestimated the limb darkening behind the planet and therefore also underestimated the true planet radius. This is because the analytical approximation assumes the limb darkening behind the planet is constant, and equal to the value behind the centre of the planet. In reality, the stellar photosphere behind the planet exhibits a range of limb darkening. This is why the numerical model in Sect. 3.2 is necessary to recover the \( R_p \) injected into the simulated data.

The \( R_p \) reported in B16 comes from averaging together the planet radii recovered across the stellar disc. If the limb darkening effects are sufficiently removed (and the stellar profile is constant), then this provides a good means to boost the signal-to-noise in the reported constant), then this provides a good means to boost the signal-enhancing \( \beta \) formulation.

Moreover, we found that, regardless of the B16 or C16 normalisation, using the \( \beta \) approximation always underestimated the limb darkening behind the planet and therefore also underestimated the true planet radius. This is because the analytical approximation assumes the limb darkening behind the planet is constant, and equal to the value behind the centre of the planet. In reality, the stellar photosphere behind the planet exhibits a range of limb darkening. This is why the numerical model in Sect. 3.2 is necessary to recover the \( R_p \) injected into the simulated data.

For our numerical approach and the C16 normalisation, we demonstrate accurate recovery of \( R_p \) (regardless of whether or not we include ingress and egress data), but only if the light curve normalisation is done with the correct \( R_p \) for each passband (using the broadband \( R_p \) for all passbands meant only the broadband \( R_p \) was recovered). Hence, regardless of the normalisation (i.e. B16 or C16) or the brightness formulation (i.e. \( \beta \) or our numerical approximation), we could only retrieve the parameters injected into the system via the transit light curve normalisation, as expected from Sect. 2.

3.4. Reanalysis of the HARPS data

Our application of the B16 procedure on the simulated data cannot explain the wavelength-dependent planet radii reported in B16. To further investigate this aspect, and to ensure we have applied the B16 method correctly, we have reanalysed the same three transits of HD 189733 b following their technique, but using the Levenberg-Marquardt least-squares minimisation from IDL’s MPFIT (Markwardt 2009, and references therein) rather than IDL’s GAUSSFIT. The results are plotted in Fig. 2, alongside those from B16. We demonstrate we can reproduce (red points in Fig. 2) results in 1–2σ agreement with B16 (in green); hence, we are confident we have applied their technique properly (in both the simulated and observed data). However, we argue that with the correct treatment of the uncertainties the recovered planet radii (in black) are consistent with a flat line (within 1–3σ of the mean \( R_p \) recovered in the simulated data, i.e. the solid black points in Fig. 1), as expected from Sect. 2. We believe the reason B16 report a trend with wavelength, and we do not, is largely due to differences in our error analysis and Gaussian fitting techniques.

In B16, the recovered \( R_p \) for each stellar disc position and all three transits were averaged together to provide one \( R_p \) for each passband, and the reported errors came from the rms of these individual planet radii (i.e. the standard deviation divided by the square root of the total number in the passband). In our analysis, we report the weighted mean for each passband, with the weights being the inverse square of the error for each individual planet radii (where the error was calculated by propagating the errors on the CCF areas as reported from the Gaussian fits following C16, and assuming negligible error on the limb darkening and stellar disc positions). The error on the weighted mean then was simply the square root of the inverse sum of the weights squared. If the errors on individual \( R_p \) were all exactly equal to the standard deviation, then the two approaches would yield the same result.

In addition to this slight difference in error analysis, B16 also applied their Gaussian fits to the oversampled CCF grid provided by the HARPS pipeline (Borsa, priv. comm.). We caution against such an approach, as the oversampling will lead to a significant underestimation of the errors. Hence, we also fit Gaussians to data composed of every one in four points from the original CCFs (to compensate for the original sampling rate of 0.25 km s\(^{-1}\) for a mean pixel width of 0.82 km s\(^{-1}\)); these results are shown in black in Fig. 2.

To test the significance of a trend in \( R_p \) with wavelength, we fitted the data with both a flat line and a linear regression, and calculated the reduced chi-squared, \( \chi^2 \), and the Bayesian information criterion (BIC); the results are shown in Table 1. We note that even if a wavelength-dependent \( R_p \) is found, it does not confirm the B16 technique is valid, as we have already shown it is not mathematically possible to retrieve radius variations. Rather, it would serve as a red flag that we do not fully characterise the interplay of the various complexities present in the observations.

Table 1. Best fits to observed data.

<table>
<thead>
<tr>
<th>Data</th>
<th>Function</th>
<th>BIC</th>
<th>( \chi^2 )</th>
<th>Function</th>
<th>BIC</th>
<th>( \chi^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 in 4 pts</td>
<td>Flat</td>
<td>13.9</td>
<td>2.0</td>
<td>Linear</td>
<td>11.7</td>
<td>1.6</td>
</tr>
<tr>
<td>Oversamp.</td>
<td>Flat</td>
<td>65.4</td>
<td>10.6</td>
<td>Linear</td>
<td>29.1</td>
<td>5.0</td>
</tr>
<tr>
<td>B16</td>
<td>Flat</td>
<td>13.1</td>
<td>1.9</td>
<td>Linear</td>
<td>5.2</td>
<td>0.3</td>
</tr>
</tbody>
</table>

1 MPFIT did not produce significantly different results compared to IDL’s GAUSSFIT, but it did allow us to propagate our errors more thoroughly (see C16 for details).
2 We note that we used the same transit parameters as B16, but there is a slight difference in the template mask used to obtain the CCFs. B16 used the archival data available from the ESO website, where 2 nights used the G2 mask and 1 night used the K5 mask, whereas our data always used the K5 mask. However, this difference is unlikely to impact the analysis since each night had its own master \( CCF_{\text{out}} \).
wavelength-dependent fit, which indicates the model is overfitting the data. We note these tests were only performed on our reanalysis of the oversampled data for comparison with B16; for our conclusions on the best-fit, we refer the reader to the analysis on the CCFs sampled every one in four points.

For the properly sampled dataset, we found only a marginal improvement in the fit for the wavelength-dependent model, and do not deem this improvement to be statistically significant (see Table 1). Moreover, the best-fit flat model ($R_p = 0.1569 \pm 0.0003$) lies within 3$\sigma$ of the mean $R_p$ predicted by the simulations. The slight improvement for the sloped model is also heavily influenced by only a couple data points from a single transit, in August 2007, as shown in the subplot of Fig. 2. If the best-fit model is robust, it should withstand removing the August transit; however, doing so means the data is then best-fit by a flat line ($\chi^2 = 1.6$, BIC = 11.1 and $\chi^2 = 1.7$, BIC = 12.6 for flat and sloped line, respectively). Consequently, we believe B16 likely report a wavelength-dependent trend in $R_p$ due to insufficient error analysis, and that its agreement with the literature may be purely coincidental.

4. Conclusions

We outline our conclusions on a point-by-point basis below.

– The technique presented in B16 using the ratio of the areas of the local (starlight behind the planet) to the out-of-transit CCF cannot be used to determine $R_p$, as $R_p$ must be known a priori for the transit light curve normalisation required for ground-based spectra. This is shown both analytically and using a simulated star-planet system.

– The analytical $\beta$ approximation used in B16 also introduces (slight) systematic trends with planet position due to inadequately accounting for limb darkening and fractional area occultation effects, and underestimates the value of $R_p$.

– We postulate that the $R_p$ variations reported in B16 are likely due to underestimated errors (largely originating from use of oversampled CCFs), as our reanalysis of the HD189733b transits further demonstrates that the only $R_p$ recoverable is that injected into the transit light curve normalisation.

– Chromatic RM measurements from ground-based spectra are not possible without taking the Doppler information into account. Hence, for future measurements, we advise readers to either follow the works of Snellen (2004), Di Gloria et al. (2015) or to apply the line-profile tomography of Collier Cameron et al. (2010) directly on each spectral passband.

Acknowledgements. We thank the referees, I. A. G. Snellen and S. Albrecht, for their careful reading and constructive comments, which improved the clarity of the manuscript. We also thank F. Borsa for useful discussions. Additionally, H.M.C. thanks E. de Mooij for suggesting we use simulated stars to test the work herein. H.M.C. and C.A.W. gratefully acknowledge support from the Leverhulme Trust (grant RPG-249). H.M.C., V.B., C.L. and A.W. acknowledge the financial support of the National Centre for Competence in Research “Planets” supported by the Swiss National Science Foundation (SNSF). C.A.W. also acknowledges support from STFC grant ST/L000709/1, and A.W. acknowledges additional financial support directly from the SNSF. This research has made use of NASA’s Astrophysics Data System Bibliographic Services.

References