**Astronomy & Astrophysics**

# Line-profile variations in radial-velocity measurements

## Two alternative indicators for planetary searches

P. Figueira[1], N. C. Santos[1,2], F. Pepe[3], C. Lovis[3], and N. Nardetto[4]

[1] Centro de Astrofísica, Universidade do Porto, Rua das Estrelas, 4150-762 Porto, Portugal
e-mail: pedro.figueira@astro.up.pt
[2] Departamento de Física e Astronomia, Faculdade de Ciências, Universidade do Porto, Portugal
[3] Observatoire Astronomique de l'Université de Genève, 51 Ch. des Maillettes, 1290 Sauverny, Versoix, Suisse
[4] Laboratoire Lagrange, UMR 7293, UNS/CNRS/OCA, 06300 Nice, France

**ABSTRACT**

*Aims.* We introduce two methods to identify false-positive planetary signals in the context of radial-velocity exoplanet searches. The first is the bi-Gaussian cross-correlation function fitting (and monitoring of the parameters derived from it), and the second is the measurement of asymmetry in radial-velocity spectral line information content, $V_{asy}$. We assess the usefulness of each of these methods by comparing their results with those delivered by current indicators.
*Methods.* We make a systematic analysis of the most used common line profile diagnosis, Bisector Inverse Slope and Velocity Span, along with the two proposed ones. We evaluate all these diagnosis methods following a set of well-defined common criteria and using both simulated and real data. We apply them to simulated cross-correlation functions that are created with the program SOAP and which are affected by the presence of stellar spots. We consider different spot properties on stars with different rotation profiles and simulate observations as obtained with high-resolution spectrographs. We then apply our methodology to real cross-correlation functions, which are computed from HARPS spectra, for stars with a signal originating in activity (thus spots) and for those with a signal rooted on a planet.
*Results.* We demonstrate that the bi-Gaussian method allows a more precise characterization of the deformation of line profiles than the standard bisector inverse slope. The calculation of the deformation indicator is simpler and its interpretation more straightforward. More importantly, its amplitude can be up to 30% larger than that of the bisector span, allowing the detection of smaller-amplitude correlations with radial-velocity variations. However, a particular parametrization of the bisector inverse slope is shown to be more efficient on high-signal-to-noise data than both the standard bisector and the bi-Gaussian. The results of the $V_{asy}$ method show that this indicator is more effective than any of the previous ones, being correlated with the radial-velocity with more significance for signals resulting from a line deformation. Moreover, it provides a qualitative advantage over the bisector, showing significant correlations with RV for active stars for which bisector analysis is inconclusive.
*Conclusions.* We show that the two indicators discussed here should be considered as standard tests to check for the planetary nature of a radial-velocity signal. We encourage the usage of different diagnosis as a way of characterizing the often elusive line profile deformations.

**Key words.** planetary systems – techniques: radial velocities – line: profiles – methods: data analysis

## 1. Introduction

The detection of extrasolar planets relies heavily on the radial velocity (RV) method, which accounts for the vast majority of the planetary discoveries up to now. Yet, it has been known since the very first announcement by Mayor & Queloz (1995) that this approach is vulnerable to false positives. The potential parasitic signals are often rooted in stellar photospheric effects that deform spectral line profiles and introduce a RV signal which is not rooted on a Doppler shift of the spectral lines.

The relatively high frequency of false positives led to the establishment of studies of line profile properties for detected RV signals as a way of proving – or better said, disproving – their planetary nature. The first in-depth study was done by Queloz et al. (2001), whom established the bisector and the bisector inverse slope (BIS) as the paradigm for its genre. The authors showed that there was a clear correlation between BIS and the measured RV for the case of a signal created by a photospheric spot. Numerous studies have made use of this principle

(e.g. Martínez Fiorenzano et al. 2005; Niedzielski et al. 2009; Hébrard et al. 2010; Figueira et al. 2010a, just to cite a few) to reject the planetary nature of a signal, and planet-hunting surveys now routinely search for a correlation between RV and BIS variations as a way of pinpointing a parasite signal.

Several alternative methodologies exist for identifying these signals. For stabilized spectrographs, the monitoring of the full-width at half-maximum (FWHM) of the cross-correlation function (CCF) can reveal profile variations (e.g. Dumusque et al. 2011), and the $V_{span}$ was introduced by Boisse et al. (2011) as an alternative to BIS to handle cases of low signal-to-noise ratio (S/N) data. Moreover, several campaigns compared RV measurements made in different wavelengths to exclude planetary candidates, since a cold spot will induce a signal whose amplitude depends on the contrast relative to the surrounding stellar disk (e.g. Martín et al. 2006; Huélamo et al. 2008; Figueira et al. 2010b). When the range of possible applications is wider, an indicator is more used, and the bisector, making very few assumptions and being easy to calculate, is by far the most popular.

In a general sense, the efficiency of a diagnosis criteria in identifying RV signals created by line deformations depends mostly on two properties. First, the indicator should be proportional to the RV signal created (if possible linearly proportional, as it eases its identification and interpretation), and second this proportionality should be characterized by a proportionality constant as large as possible (when compared with the average error bars of the indicator).

Even though BIS is undoubtedly useful in many situations, studies such as those of Saar & Donahue (1997), Santos et al. (2003b), and Desort et al. (2007) pointed to an important limitation. For many cases of line profile deformations, characterized by certain stellar and spot properties, the BIS is intrinsically less sensitive to the presence of a deformation than the RV induced by the deformation itself. In other words, the proportionality constant between BIS and RV for these configurations has an absolute value $\ll 1$. On top of this, it is well known that the relation between BIS and RV is not expected to be monotonic (see e.g. Boisse et al. 2011). While for many practical cases the correlation can be assumed to be a straight line, the real shape is similar to that of a tilted "8"; the proportionality constant between the two variables can even change from positive to negative or vice-versa for some cases. These two properties can preclude the detection of a correlation between BIS and RV and render the diagnostic ineffective.

As we expand our searches, including more varied stellar hosts and pushing the precision of our surveys further to detect lower-amplitude signals, there is an increasing demand for alternative and more effective ways of identifying false-positive signals. There have been few attempts at finding alternative methods of characterizing line profile variations in a way useful for planetary RV searches (with the work of Dall et al. 2006, being a rare exception).

In this paper, we consider two different diagnosis methods and compare their performance with those of current line-profile indicators for a wide variety of cases. We start by presenting the definition and properties of BIS and $V_{\rm span}$, our comparison references, in Sect. 2. Then we describe the data sets used for this study and the principles of our comparison in Sect. 3. We employ the standard line-profile indicators and evaluate their results in Sect. 4. We move on to apply the first of our two methods, the Bi-Gaussian fitting, in Sect. 5 and the second, $V_{\rm asy}$, in Sect. 6. We discuss the results in Sect. 7 and conclude on the usefulness of these new indicators and how they help us understand the impact of line-profile variation on RV in Sect. 8.

## 2. Current diagnosis methods

The spectral information content of a stellar spectrum can be condensed in a CCF, which is then used to measure precise RVs (Baranne et al. 1996; Pepe et al. 2002). The RV is extracted by fitting a Gaussian function to the CCF with the center of the Gaussian delivering the star's RV; this quantity is represented throughout the paper as $RV_{\rm G, center}$. We apply our methods to the CCF and its derived quantities and compare the results with those obtained for known line-profile indicator variations. For reference line-profile indicators, we consider the BIS, the most common diagnosis method for the presence of parasitic signals, and $V_{\rm span}$. In the next sections we describe these methods in detail and how they were implemented.

### 2.1. The BIS

In the BIS analysis, we used the definition provided by Queloz et al. (2001). Consider a CCF and calculate its mid-point (i.e., the

point at equal distance between its blue and red wing) for each flux level between the continuum level and the bottom of the line. Since the flux level will only coincidentally fall on a pixel value, the flux values are calculated by interpolation from the neighboring CCF pixels. The BIS value is then simply the difference between the velocity of the top and bottom sections of the line, with the top being the average mid-point of the flux levels located between 60 and 90% and the bottom of those between 10 and 40% of the line full depth. We refer the reader to the Fig. 5 of Queloz et al. (2001) for a helpful schematic representation. Since we are working in the velocity space, the line mid points are measured in km s$^{-1}$ (or any multiple of it), as is BIS.

The interpretation of the BIS is quite straightforward. Any RV signal associated with a variation between the top and bottom sections of the line is not rooted in the coordinated movement of it but rather on a line deformation, and a plot of BIS-RV will reveal a correlation.

Even though the vast majority of works follow this prescription (see references in the introduction), the top and bottom regions of the BIS can be defined differently. In particular, one can consider narrower bands located at different heights, leading to different slopes when the line is subjected to deformations. This different leverage capability will lead to different indicator amplitude variations at the expense of using only a smaller section of the line, and thus working at a lower S/N when applying it to real spectra. To explore this posibility, we considered two different parametrizations, representative of extreme cases: a maximum leverage case BIS$^+$ in which the limits of top and bottom regions are 80−90% and 10−20% of the total flux level, respectively, and a minimum leverage case BIS$^-$ with top at 60−70% and bottom at 30−40%, respectively.

### 2.2. The $V_{\rm span}$

We employ the $V_{\rm span}$ indicator as defined in Boisse et al. (2011). This indicator represents the difference $RV_{\rm high} - RV_{\rm low}$ between two Gaussian fits of the CCF which consider exclusively the upper and lower part of the CCF, respectively. For the first, $RV_{\rm high}$, one takes into account only the top of the CCF for the fit by considering only the points in the range $[-\infty:-\sigma][+\sigma:+\infty]$ as measured relative to the center of the Gaussian fit; for the second $RV_{\rm low}$ one considers the points in the range $[\infty:-3\sigma] \vee [\sigma:+\sigma] \vee [+3\sigma:+\infty]$. The former quantity is sensitive to variations on the top of the line and the latter on the bottom.

The construction of the $V_{\rm span}$ indicator was motivated by the analysis of line-profile variations for cases of low S/N and was shown to be advantageous relative to the BIS for one of these cases. We refer the interested reader to Boisse et al. (2011) for details.

## 3. Choice of datasets and comparison principles

### 3.1. The datasets

We simulated CCFs as affected by the presence of a stellar spot using the program SOAP, and refer the interested reader to Boisse et al. (2012) for details on the program. All simulations share two basic assumptions: we acquire our spectra with a spectrograph with a resolution of 100 000 (which defines the FWHM of a non-rotating star), and we observe a star with a single dark spot (i.e., with zero emissivity). We consider different $v \sin i$ for the star, different inclinations of the stellar rotation axis relative to the line of sight, different spot latitudes, and different spot filling factors. When performing our simulations,

we use a grid resolution of 2000 and a spot circumference resolution of 50; the orbital phase is homogeneously sampled by 100 points. For the steps in RV for the CCF calculation, we considered a very fine grid with steps of $10\,\mathrm{m/s}$ from $-30\,\mathrm{km\,s^{-1}}$ to $+30\,\mathrm{km\,s^{-1}}$. On real data the CCF is calculated on a much coarser grid, usually of $\sim100\,\mathrm{m/s}$ or higher. However, a coarse discretization can introduce an appreciable deviation in the fit and on the fit parameters. Since these synthetic data have, by definition, no noise, any fit instability steming from the small number of points can introduce a detectable difference relative to high S/N spectra with a coarser grid. Since computational power is not an issue for such a small number of simulations, we choose to calculate the CCF on a very fine grid to evaluate the application of our indicators in the most precise way possible.

We started by simulating the properties of standard well-studied cases, with the first being that of HD 166435, already discussed both in Queloz et al. (2001) and in Desort et al. (2007). The star has a $v\sin i = 7.0\,\mathrm{km\,s^{-1}}$ and is observed at an inclination of $30°$; its spot has a filling factor of 1% and is located at a latitude of $60°$. This is the most extensively studied spot configuration in the literature, and we included it here for that reason[1].

We then moved to more general cases with the objective of studying the impact of key stellar and spot parameters on the amplitude of the measured RVs and the associated indicator's response. First we studied the impact of equatorial spots (latitude of $0°$) with different filling factor values of 1.0, 2.0, 5.0, and 10.0% on a edge-on star ($I = 0°$) with $v\sin i = 7.0\,\mathrm{km\,s^{-1}}$. Then, we repeated the exercise for a 1% filling factor equatorial spot on a edge-on star with rotational velocities of 2, 5, and $10\,\mathrm{km\,s^{-1}}$. These tests were also inspired from those of Desort et al. (2007), who chose these inclination and spot latitudes to induce the highest RV amplitude for a given filling factor and $v\sin i$. The spot filling factors were chosen to depict typical spot sizes on main-sequence stars, and the $v\sin i$ were chosen to illustrate several cases ranging from a broadening inferior to that introduced by the resolution of the spectrograph, to a star rotating so fast that the RV precision is significantly degraded.

We finally tested our approach on real data, and to do so we used HARPS spectra and CCFs as delivered by the last version of the DRS pipeline (Mayor et al. 2003). We selected both active stars ($\log(R'_{\mathrm{HK}}) \sim -4.46$ or higher), for which a correlation between RV and the indicators ($\Delta V$ and BIS) is expected, and a planet-host star HD 216770, for which it is not (or is not expected at the level of the measured RV signal amplitude). The planetary orbits measured on this star were discussed in Mayor et al. (2004), and the reader is referred to the paper for details. The stars are listed in Table 1, along with the $\log(R'_{\mathrm{HK}})$, $v\sin i$, and the number of data points taken at the time of the publication and considered for the analysis. The $v\sin i$ was derived in Santos et al. (2003a) and $\log(R'_{\mathrm{HK}})$ was derived in Santos et al. (2000). The only exception was HD 216770, of which $\log(R'_{\mathrm{HK}})$ was taken from Henry et al. (1996).

When the activity of the star is larger, the impact on both measured RV and line-profile variation indicators is larger. To maximize its impact on RV, one should measure it at least over one full activity cycle. In practice, this is very difficult, and recent works have shown that several longer-term activity cycles are present in a star on top of the typical rotation-modulated cycle, with a strong impact on RV (e.g. Dumusque et al. 2011; Lovis et al. 2011). While keeping in mind that the

---

[1] We note that the spot latitude definition used in *SOAP* is not the same as that used by Desort et al. (2007); in the former, the latitude angle is used, while the colatitude angle is used in the latter.

**Table 1.** HARPS star properties.

| Star | $R'_{\mathrm{HK}}$ | $v\sin i$ (km s$^{-1}$) | # points | Comments |
|------|------|------|------|------|
| HD 224789 | −4.46 | 3.01 | 36 | active star |
| HD 36051 | −4.55 | 5.07 | 13 | active star |
| HD 103720 | −4.46 | 2.82 | 49 | active star |
| HD 200143 | −4.48 | 3.02 | 42 | active star |
| BD-213153 | −4.51 | 3.63 | 29 | active star |
| HD 216770 | −4.84 | 1.4 | 29 | planet-host |

characterization of the activity-induced effects is still work in progress, we concentrate our analysis on the *relative* variations between RV and line-profile indicators and how the different indicators compare with each other.

### 3.2. Comparing the datasets

We aim to compare the different indicators in the most systematic way possible. To do so, we apply them to the same datasets and evaluate them according to the same criteria.

We first compare the amplitude of the indicator's variation (relative to its average value) with those of the RV and other indicators. We pay particular attention to how the indicators perform compared to BIS, the most commonly used line-profile indicator. If one assumes that the error bars are similar, a larger amplitude of the indicator's variation results in a greater capability to pinpoint a signal. Yet, the indicator's capabilities must be assessed through a rigorous quantitative analysis. To compare the correlation of an indicator with RV, we also consider:

- the slope resulting from the fit of a linear function by least-square minimization ($m_{\mathrm{ind}}$);
- the Pearson's correlation coefficient value for the data sets ($\rho_{\mathrm{ind}}$); and
- the probability of obtaining an equal or larger correlation coefficient absolute value by chance.

The slope of the correlation between an indicator and the RV is a different way of relating a variation in the indicator with that of the RV. However, it does not evaluate the strength of the correlation betwen the two data sets. To do this, we use the Pearson's correlation coefficient, which is a measure of linear dependence between the two variables. To test if the correlation coefficient's value is significant, we test if it can be obtained by a fortuitous pair-matching between the indicator and the RV. To do so, we bootstrap 100 000 times the data pairs and calculate the correlation coefficient $\rho_{\mathrm{ind}}$ for each of these synthetic uncorrelated data sets. We then calculate the average and standard deviation ($\sigma_{\rho,\mathrm{ind}}$) of these $\rho_{\mathrm{ind}}$ values; by assuming a Gaussian distribution, one can calculate the probability that the $\rho_{\mathrm{ind}}$ of the original dataset was obtained by pure chance. While the probability value is conditioned by the hypothesis that the indicators follow a Gaussian distribution, its distance in $\sigma$ is still a good indicator on how the measured value compares with that of an uncorrelated distribution.

When using real data, we have two important additional information elements. The first of these is the RV measurement uncertainty, as estimated from the RV information content of the spectra (e.g. Bouchy et al. 2001). Moreover, we know if the RV signal introduced is of planetary or stellar origin and, as a consequence, if a line profile deformation is present or not; we also know then if the indicator is expected to be correlated with

**Table 2.** Comparison between correlations of BIS and $V_{span}$ with RV for simulated data.

| Model/star | $RV_{max}$ | $BIS_{max}$ | $V_{span,max}$ | $m_{BIS}$ | $m_{V_{span}}$ | $\rho_{BIS}$ | $\rho_{V_{span}}$ | $\sigma_{\rho,BIS}$ | $\sigma_{\rho,V_{span}}$ |
|---|---|---|---|---|---|---|---|---|---|
| HD 166435' | 9.28e+01 | 1.03e+02 | 9.04e+01 | −1.068 | −0.917 | −0.990 | −0.983 | 9.86 | 9.73 |
| $F_r = 1.0\%$ | 9.98e+01 | 1.03e+02 | 8.95e+01 | −0.686 | −0.569 | −0.827 | −0.802 | 8.23 | 7.99 |
| $F_r = 2.0\%$ | 2.01e+02 | 2.10e+02 | 1.82e+02 | −0.679 | −0.563 | −0.820 | −0.793 | 8.18 | 7.88 |
| $F_r = 5.0\%$ | 5.20e+02 | 5.50e+02 | 4.87e+02 | −0.655 | −0.540 | −0.795 | −0.760 | 7.90 | 7.58 |
| $F_r = 10.0\%$ | 1.10e+03 | 1.22e+03 | 1.10e+03 | −0.599 | −0.493 | −0.741 | −0.693 | 7.37 | 6.90 |
| $v \sin i = 2.0$ | 2.21e+01 | 1.94e+00 | 1.42e+00 | −0.057 | −0.042 | −0.801 | −0.800 | 7.97 | 7.94 |
| $v \sin i = 5.0$ | 6.43e+01 | 3.72e+01 | 2.95e+01 | −0.378 | −0.293 | −0.815 | −0.805 | 8.09 | 8.02 |
| $v \sin i = 10.0$ | 1.57e+02 | 2.54e+02 | 2.53e+02 | −1.109 | −1.001 | −0.847 | −0.793 | 8.43 | 7.86 |

**Notes.** When not explicitly stated, the probability of obtaining the presented $\rho$ values from the distribution obtained by bootstrapping is lower than 1.0e-4. RV, BIS and $V_{span}$ in units of m/s.

**Table 3.** Comparison between BIS and $V_{span}$ correlations with RV for HARPS data for the parameters that do not take into consideration the error bars.

| Star | $RV_{max}$ | $BIS_{max}$ | $V_{span}$ | $\rho_{BIS}$ | $\rho_{V_{span}}$ | $\sigma_{\rho,BIS}$ | $\sigma_{\rho,V_{span}}$ |
|---|---|---|---|---|---|---|---|
| HD 224789 | 3.53e+01 | 2.77e+01 | 2.10e+01 | −0.818 | −0.818 | 4.84 (<0.00%) | 4.84 (<0.00%) |
| HD 36051 | 2.16e+01 | 3.29e+01 | 2.03e+01 | −0.275 | −0.286 | 0.95 (34.11%) | 0.99 (32.09%) |
| HD 103720 | 1.10e+02 | 2.41e+01 | 1.87e+01 | −0.243 | −0.314 | 1.68 (9.24%) | 2.17 (2.99%) |
| HD 200143 | 4.70e+01 | 1.78e+02 | 1.53e+02 | −0.302 | −0.260 | 1.93 (5.35%) | 1.66 (9.61%) |
| BD-213253 | 5.33e+01 | 3.33e+01 | 2.45e+01 | −0.130 | −0.135 | 0.69 (49.18%) | 0.71 (47.70%) |
| HD 216770 | 2.67e+01 | 4.95e+00 | 4.30e+00 | 0.142 | 0.178 | 0.75 (45.11%) | 0.94 (34.85%) |

**Notes.** RV, BIS and $V_{span}$ in m/s.

**Table 4.** Comparison between BIS and $V_{span}$ correlations with RV for HARPS data for the parameters that consider the error bars, i.e. the slope fit parameters.

| Star | $(m, b)_{BIS}$ | $(m, b)_{V_{span}}$ |
|---|---|---|
| HD 224789 | $(-0.705^{+0.012}_{-0.011}$ or $^{+1.73\%}_{-1.57\%}, 26.083^{+0.409}_{-0.451})$ | $(-0.525^{+0.011}_{-0.010}$ or $^{+2.13\%}_{-1.99\%}, 19.424^{+0.386}_{-0.413})$ |
| HD 36051 | $(-0.476^{+0.144}_{-0.126}$ or $^{+30.16\%}_{-26.51\%}, 0.795^{+0.204}_{-0.232})$ | $(-0.309^{+0.124}_{-0.112}$ or $^{+40.27\%}_{-36.46\%}, 0.517^{+0.182}_{-0.200})$ |
| HD 103720 | $(-0.049^{+0.006}_{-0.005}$ or $^{+11.39\%}_{-11.26\%}, -0.287^{+0.036}_{-0.036})$ | $(-0.045^{+0.006}_{-0.005}$ or $^{+12.14\%}_{-12.02\%}, -0.276^{+0.036}_{-0.036})$ |
| HD 200143 | $(-0.239^{+0.022}_{-0.021}$ or $^{+9.26\%}_{-8.61\%}, -4.887^{+0.452}_{-0.421})$ | $(-0.158^{+0.021}_{-0.020}$ or $^{+13.40\%}_{-12.71\%}, -3.233^{+0.433}_{-0.410})$ |
| BD-213253 | $(-0.225^{+0.034}_{-0.032}$ or $^{+15.09\%}_{-14.09\%}, 5.527^{+0.774}_{-0.829})$ | $(-0.172^{+0.032}_{-0.030}$ or $^{+18.53\%}_{-17.67\%}, 4.220^{+0.742}_{-0.778})$ |
| HD 216770 | $(0.014^{+0.006}_{-0.006}$ or $^{+43.20\%}_{-42.97\%}, -0.439^{+0.193}_{-0.194})$ | $(0.013^{+0.006}_{-0.006}$ or $^{+47.62\%}_{-47.49\%}, -0.401^{+0.192}_{-0.193})$ |

the measured RV variation or not. It follows then that we can add the following criteria when evaluating the real data sets:

– the error bars on the correlation slope $(m)_{ind}$ which result from bootstrapping the RV data and indicator with respect to the respective error bars, performed by drawing 100 000 times values around their error bars;
– the relationship of the slope and correlation values to the stellar activity or presence of planets around the star.

We note that the bootstrapping is a non-parametric test, depending on very few assumptions, and thus being very robust. However, to take full advantage of the real data, one has to estimate the error bars on the indicators themselves. In spite of the abundant literature on BIS, there are virtually no studies on the error distribution of this estimator. It is commonly reported that the error bar on this difference is on the order of $\sqrt{2}$ times the photon noise on the CCF because BIS corresponds to a difference of velocities between the upper and the lower sections of the CCF. Even though these error bars are ill-justified, our objective is to compare the performance among indicators, and and so we choose to use the commonly accepted uncertainties for BIS for the other indicators. To allow a more straightforward interpretation of our results, we separate those that depend on

the assumptions on the error bars from those which do not and present them in different tables.

## 4. Standard profile indicators: BIS and $V_{span}$

### 4.1. BIS vs. $V_{span}$

We start by applying the two standard line-profile indicators to our datasets. These have been discussed in their presentation papers and an exhaustive analysis is not the objective of this paper, so we focus instead on their relative strenghts and weaknesses.

Both the analysis of synthetic (Table 2) and real data (Tables 3 and 4) show that the BIS and $V_{span}$ yields similar results for most of the cases. Their variation is proportional to both the spot filling factor and the star rotational velocity. For most of the considered cases, the results are of the same order of magnitude for the data and the different evaluation criteria considered. BIS exhibits higher-amplitude variations and more significant correlations in a consistent way for all synthetic datasets but does not prove to be advantageous when applied to real data. $V_{span}$ yielded poorer results than those delivered by the BIS on what concerns correlation slopes and associated errors, but we must stress that the latter depends on the assumed error bars for the bootstrapping.

**Table 5.** Comparison between correlations of BIS$^+$ and BIS$^-$ with RV for simulated data.

| Star | $RV_{max}$ | $BIS^+_{max}$ | $BIS^-_{max}$ | $m_{BIS+}$ | $m_{BIS-}$ | $\rho_{BIS+}$ | $\rho_{BIS-}$ | $\sigma_{\rho,BIS+}$ | $\sigma_{\rho,BIS-}$ |
|---|---|---|---|---|---|---|---|---|---|
| HD 166435' | 9.28e+01 | 1.50e+02 | 5.69e+01 | −1.548 | −0.593 | −0.988 | −0.993 | 9.85 | 9.90 |
| $F_r = 1.0\%$ | 9.98e+01 | 1.50e+02 | 5.72e+01 | −0.993 | −0.383 | −0.824 | −0.833 | 8.20 | 8.27 |
| $F_r = 2.0\%$ | 2.01e+02 | 3.03e+02 | 1.17e+02 | −0.982 | −0.380 | −0.816 | −0.827 | 8.15 | 8.22 |
| $F_r = 5.0\%$ | 5.20e+02 | 7.92e+02 | 3.19e+02 | −0.941 | −0.370 | −0.788 | −0.804 | 7.86 | 8.01 |
| $F_r = 10.0\%$ | 1.10e+03 | 1.68e+03 | 7.79e+02 | −0.849 | −0.346 | −0.727 | −0.745 | 7.20 | 7.41 |
| $v \sin i = 2.0$ | 2.21e+01 | 2.89e+00 | 1.05e+00 | −0.085 | −0.031 | −0.801 | −0.799 | 7.98 | 7.95 |
| $v \sin i = 5.0$ | 6.43e+01 | 5.29e+01 | 2.15e+01 | −0.541 | −0.216 | −0.818 | −0.811 | 8.17 | 8.06 |
| $v \sin i = 10.0$ | 1.57e+02 | 4.10e+02 | 1.26e+02 | −1.699 | −0.557 | −0.818 | −0.879 | 8.17 | 8.74 |

**Notes.** When not explicitly stated, the probability of obtaining the presented $\rho$ values from the distribution obtained by bootstrapping is lower than 1.0e-4. RV, BIS$^+$, and BIS$^-$ in units of m/s.

**Table 6.** Comparison between BIS$^+$ and BIS$^-$ correlations with RV for HARPS data for the parameters that do not consider the error bars.

| Star | $RV_{max}$ | $BIS^+_{max}$ | $BIS^-_{max}$ | $\rho_{BIS+}$ | $\rho_{BIS-}$ | $\sigma_{\rho,BIS+}$ | $\sigma_{\rho,BIS-}$ |
|---|---|---|---|---|---|---|---|
| HD 224789 | 3.53e+01 | 4.15e+01 | 1.51e+01 | −0.819 | −0.795 | 4.84 (<0.00%) | 4.71 (<0.00%) |
| HD 36051 | 2.16e+01 | 7.83e+01 | 3.37e+01 | −0.459 | 0.116 | 1.59 (11.16%) | 0.40 (68.64%) |
| HD 103720 | 1.10e+02 | 3.86e+01 | 1.55e+01 | −0.214 | −0.290 | 1.49 (13.66%) | 2.00 (4.49%) |
| HD 200143 | 4.70e+01 | 1.78e+02 | 4.45e+01 | −0.343 | −0.379 | 2.20 (2.82%) | 2.44 (1.51%) |
| BD-213253 | 5.33e+01 | 6.96e+01 | 2.01e+01 | 0.031 | −0.307 | 0.16 (87.10%) | 1.62 (10.46%) |
| HD 216770 | 2.67e+01 | 6.33e+00 | 4.83e+00 | 0.127 | 0.137 | 0.67 (50.09%) | 0.73 (46.83%) |

**Notes.** RV, BIS$^+$ and BIS$^-$ in m/s.

**Table 7.** Comparison between BIS$^+$ and BIS$^-$ correlations with RV for HARPS data for the parameters that consider the error bars, i.e. the slope fit parameters.

| Star | $(m, b)_{BIS^+}$ | $(m, b)_{BIS^-}$ |
|---|---|---|
| HD 224789 | $(-1.018^{+0.014}_{-0.013}$ or $^{+1.38\%}_{-1.24\%}, 37.647^{+0.466}_{-0.519})$ | $(-0.379^{+0.011}_{-0.010}$ or $^{+2.83\%}_{-2.69\%}, 14.014^{+0.376}_{-0.395})$ |
| HD 36051 | $(-1.164^{+0.209}_{-0.163}$ or $^{+17.98\%}_{-14.01\%}, 1.924^{+0.263}_{-0.338})$ | $(0.003^{+0.117}_{-0.119}$ or $^{+3711.23\%}_{-3773.88\%}, 0.008^{+0.192}_{-0.189})$ |
| HD 103720 | $(-0.072^{+0.006}_{-0.006}$ or $^{+7.76\%}_{-7.72\%}, -0.422^{+0.036}_{-0.036})$ | $(-0.029^{+0.006}_{-0.005}$ or $^{+19.25\%}_{-19.14\%}, -0.172^{+0.036}_{-0.036})$ |
| HD 200143 | $(-0.317^{+0.024}_{-0.022}$ or $^{+7.55\%}_{-6.79\%}, -6.486^{+0.488}_{-0.440})$ | $(-0.149^{+0.021}_{-0.020}$ or $^{+14.13\%}_{-13.32\%}, -3.051^{+0.431}_{-0.406})$ |
| BD-213253 | $(-0.169^{+0.039}_{-0.037}$ or $^{+23.06\%}_{-22.00\%}, 4.173^{+0.908}_{-0.951})$ | $(-0.203^{+0.032}_{-0.030}$ or $^{+15.82\%}_{-14.72\%}, 4.955^{+0.728}_{-0.781})$ |
| HD 216770 | $(0.014^{+0.006}_{-0.006}$ or $^{+44.67\%}_{-44.38\%}, -0.418^{+0.193}_{-0.194})$ | $(0.011^{+0.006}_{-0.006}$ or $^{+56.89\%}_{-56.18\%}, -0.336^{+0.191}_{-0.193})$ |

### 4.2. Different parametrizations of BIS

The first aspect coming out of Table 5 is that the variation of BIS$^+$ is larger than that of BIS$^-$ by a factor of 2−3 for all the cases presented. This is associated with stronger correlations and correlation slopes for the former indicator. Moreover, one concludes that the BIS$^+$ leads to a variation that is larger than that of BIS by ∼50% for all the configurations when compared with the corresponding values of the tables in the previous section; the correlations slopes are larger, with the correlation coefficient being very similar.

The analysis of the results on real data, presented in Tables 6 and 7, are not easy to interpret. The amplitude variations are larger for BIS$^+$ than for BIS$^-$; the correlation coefficients and correlation slopes are similar or larger for BIS$^+$ than for BIS$^-$ with one clear exception, BD-213253. For the latter, the BIS$^-$ correlation coefficient even approaches the 10% probability level, hinting at the existence of a correlation. This is very interesting in itself, because neither the BIS nor the $V_{span}$ alone could reach this level. However, the high probability of obtaining these correlations by chance prevent us from taking this discussion further.

It is also interesting to compare the results of BIS with those of BIS$^+$ for the real data. The amplitude variation of BIS$^+$ can be larger by a factor of more than 2 than that of the BIS for some cases, but they deliver essentially the same results when it comes to the detection of correlations. Interestingly, an increase in amplitude can even be accompanied by a decrease in the correlation factor; the results on the planet-host star are an interesting example of this. The same happened on BD-213253, decreasing the correlation coefficient to completely negligible levels. This behavior underlines the fact that the variation in amplitude cannot be interpreted straightforwardly. Moreover, it is important to note that that the spectra considered here are high S/N spectra and that the CCF has then a very high S/N; BIS$^+$ is expected to be much less advantageous when the S/N decreases.

The main conclusion from the analysis of the BIS$^+$ is that it tends to be advantageous for the cases considered here, but it presents a complex behavior, ranging from clearly advantageous to slightly disadvantageous, and it did not allow to detect any new correlations in the data.

## 5. The bi-Gaussian fitting

### 5.1. The method

The bi-Gaussian is a simple function that considers a Gaussian with wings characterized by two different HWHM (half-width

**Table 8.** Comparison between correlations of BIS and $\Delta V$ with RV for simulated data.

| Star | $RV_{max}$ | $BIS_{max}$ | $\Delta V_{max}$ | $m_{BIS}$ | $m_{\Delta V}$ | $\rho_{BIS}$ | $\rho_{\Delta V}$ | $\sigma_{\rho,BIS}$ | $\sigma_{\rho,\Delta V}$ |
|------|-----------|-------------|------------------|-----------|----------------|--------------|-------------------|---------------------|--------------------------|
| HD 166435' | 9.28e+01 | 1.03e+02 | 1.33e+02 | −1.068 | 1.364 | −0.990 | 0.987 | 9.82 | 9.82 |
| $F_r = 1.0\%$ | 9.98e+01 | 1.03e+02 | 1.32e+02 | −0.686 | 0.871 | −0.827 | 0.822 | 8.20 | 8.20 |
| $F_r = 2.0\%$ | 2.01e+02 | 2.10e+02 | 2.66e+02 | −0.679 | 0.859 | −0.820 | 0.812 | 8.17 | 8.08 |
| $F_r = 5.0\%$ | 5.20e+02 | 5.50e+02 | 6.94e+02 | −0.655 | 0.811 | −0.795 | 0.777 | 7.92 | 7.74 |
| $F_r = 10.0\%$ | 1.10e+03 | 1.22e+03 | 1.47e+03 | −0.599 | 0.701 | −0.741 | 0.703 | 7.38 | 7.00 |
| $v \sin i = 2.0$ | 2.21e+01 | 1.94e+00 | 2.30e+00 | −0.057 | 0.067 | −0.801 | 0.802 | 7.98 | 7.96 |
| $v \sin i = 5.0$ | 6.43e+01 | 3.72e+01 | 4.39e+01 | −0.378 | 0.448 | −0.815 | 0.817 | 8.10 | 8.15 |
| $v \sin i = 10.0$ | 1.57e+02 | 2.54e+02 | 3.83e+02 | −1.109 | 1.596 | −0.847 | 0.817 | 8.46 | 8.12 |

**Notes.** When not stated, the probability of obtaining the presented $\rho$ values from the distribution obtained by bootstrapping is lower than 1.0e-4. RV, BIS, and $\Delta V$ in units of m/s.
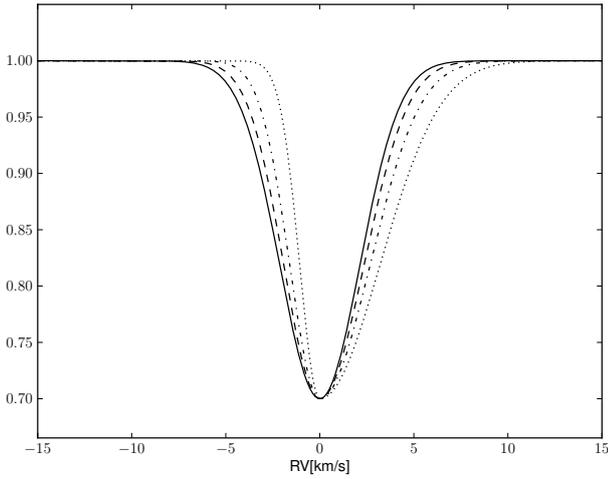


**Fig. 1.** Comparison of bi-Gaussian functions with different asymmetry values: 0% (Gaussian function – solid line), 10% (dashed line), 25% (dashed-dotted line), and 50% (dotted line). These normalized functions have a contrast of 30%, a FWHM of 5 km s$^{-1}$, and are all centered at 0 m/s.

at half-maximum), or, equivalently, an asymmetry parameter associated with the FWHM. It was first used in the context of RV measurements by Nardetto et al. (2006) as a way of simultaneously measuring the asymmetry and the center of spectral lines on pulsating stars. It can be represented by

$$bG(\text{RV}) = -D \exp\left(-\frac{2\ln 2(RV - RV_{\text{center}})^2}{(FWHM \times (1 + A))^2}\right)$$
$$+C \quad \text{if } RV > RV_{\text{center}} \qquad (1)$$

and

$$bG(\text{RV}) = -D \exp\left(-\frac{2\ln 2(RV - RV_{\text{center}})^2}{(FWHM \times (1 - A))^2}\right)$$
$$+C \quad \text{if } RV < RV_{\text{center}} \qquad (2)$$

in which $D$ is the depth of the line, and $C$ the continuum level, as measured in photoelectrons[2]. Since we apply it to the CCF, we work in the RV space, and both the bi-Gaussian center $RV_{\text{center}}$ and FWHM are expressed in km s$^{-1}$. The asymmetry $A$ is given as a percentage (of the FWHM). For illustration, we plot the effect of introducing an asymmetry on a Gaussian function in Fig. 1.

_____

[2] It is equivalent to work with a normalized function, as Nardetto et al. (2006) chose to do.

This function is fitted by minimizing the $\chi^2_{\text{red}}$, just like the Gaussian function; by imposing $A = 0$ one recovers the typical Gaussian fitting. We denote the $RV_{\text{center}}$ parameter obtained by Gaussian fitting by $RV_{G,\text{center}}$. As stated by Nardetto et al. (2006), the indicator $\Delta V = RV_{G,\text{center}} - RV_{\text{center}}$ represents the RV shift that can be explained by the asymmetry alone, and we consider it as our proxy for the line deformation. The interpretation is once again very simple. The extra degree of freedom in the bi-Gaussian fitting allows to accommodate a line profile variation and is thus more insensitive to the line deformations than a simple Gaussian fitting. The operational similarities between BIS and $\Delta V$ then follow very naturally. In this work we considered $\Delta V = RV_{\text{center}} - RV_{G,\text{center}}$, so that a plot $\Delta V$-RV displays a positive correlation.

### 5.2. $\Delta V$: application to simulated data

The results for the simulated data are presented in Table 8. In Fig. 2 we plot the phase dependence of the most important parameters for the case of the 1% filling factor and BIS and $\Delta V$ as a function of $RV_{G,\text{center}}$ in Fig. 3.

It is easy to notice that it is advantageous to use $\Delta V$ relative to the BIS as an indicator of line-profile deformations for the simulated data. The amplitude of its variation is always higher with the improvement being of 10−30% (with the most significant improvement being obtained for the largest projected rotational velocity). The value of the correlation slope is also larger, with a variable improvement, depending on the case considered, but reaches above 50%. These two factors suggest that the $\Delta V$-RV correlation and its slope will emerge more easily in cases where the BIS-RV slope is small when compared with the photon noise level. The absolute value of Pearson's correlation coefficient is very high and very similar for both correlations, and the probability that these happen by chance is smaller than 0.01%. In a nutshell, we can note that the application of the indicator is advantageous when compared with that of BIS for the whole range of spot filling factors and $v \sin i$ considered in the synthetic data, since the amplitude variations are higher with all other parameters being similar. From what we discussed in Sect. 4.2 we note that BIS$^+$ variations are higher than those of $\Delta V$ by ∼30%, while displaying similar correlation coefficients.

### 5.3. $\Delta V$: Application to HARPS data

At this point, it is important to recall the discussion on the uncertainties of the estimators, and, in particular, of BIS, as presented in Sect. 3.2. Yet, the error bars on $\Delta V$ are a different issue: these result from the subtraction of the RV's obtained by fitting the same CCF with two different functions. If one assumes that the
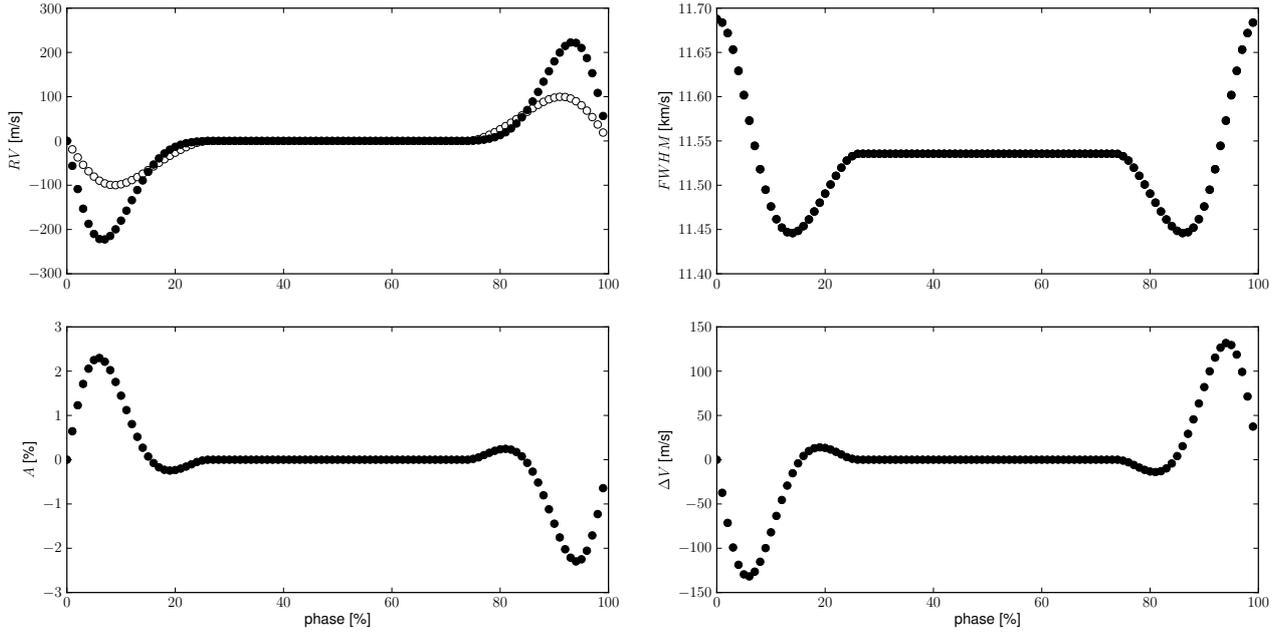
**Fig. 2.** Most important properties for the two fitted functions: Gaussian as open circles (only two upper panels, partially superposed) and *Bi-Gaussian* as filled circles, as a function of orbital phase (with an arbitrary zero-point) for the case of an equatorial spot with a a filling factor of 1% on an edge-on star with $v \sin i$ of $7 \, \mathrm{km \, s^{-1}}$. *Upper left*: center of the two fitted functions as a function of orbital phase; *Upper right*: FWHM of the two functions; on the two *lower panels*, we have the asymmetry $A$ (*bottom left*) and the $\Delta V$ indicator (*bottom right*) for the *Bi-Gaussian* fitting.
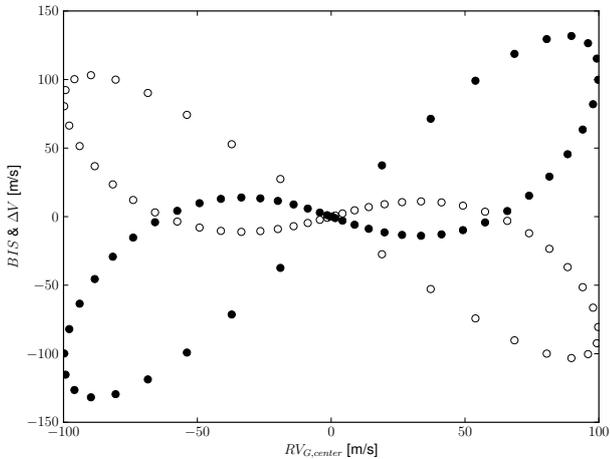


**Fig. 3.** Two indicators considered, BIS (as open circles) and $\Delta V$ (as filled circles), as a function of RV, for the case of an equatorial spot with a filling factor of 1% on an edge-on star with $v \sin i$ of $7 \, \mathrm{km \, s^{-1}}$.

fitting error is negligible and the final error is dominated by the photon-noise contribution as for BIS when fitting a function to the data, the error bars on the difference are the same as those of either fitting. In other words, the error of both fitting procedures is expected to be completely correlated. To test the two approaches on an even footing, we consider however the error bars on $\Delta V$ as being also $\sqrt{2}$ times the photon noise, just like for BIS.

As done before, we separate the results into two tables: The first contains the quantities that do not depend on the error values and their assumptions, as seen in Table 9. The second lists the fitted correlation slopes and uncertainties drawn from bootstrapping the indicators, which depend on the indicator's errors,

as seen in Table 10. The reader thus should keep in mind the assumptions behind the results of this last Table.

When applying $\Delta V$ to real data, the situation is slightly different than that of the application to simulated data. The amplitude of $\Delta V$ and the absolute value of Pearsons's correlation coefficient are higher than those of BIS, but for both planet-hosting and non-planet hosting stars. Consequently, the probability that the correlation happens by chance tends to be higher with the exception on both of these properties being that of BD-213253. In particular, a more pronounced correlation is obtained for the planet-host star HD 216770. As expected, the statistical test however shows that one has a non-negligible probability of attaining the measured $\rho$ by chance, which shows that the correlation is not to be trusted, just as for the BIS. We plot the application of the indicator to the star with the most significant correlation, HD 224789, and compare it to BIS in Fig. 4. Analyzing the error bars on the slope $m$ is more difficult, and the results are more heterogeneous. As discussed before, these results stand on additional assumptions. The first point to note is that the relative error bars (in %) for the $\Delta V$ are always smaller than those for the BIS, except for the case of the star for which no correlation was detected, BD-213253. As for the others, the absolute values for $m$ are always larger, as expected, but the results are statistically very similar. We also note that all the slopes are non-compatible with zero at several sigma, if the error bars are trusted, even those from BD-213253 and our planet-host, HD 216770. As a concluding remark on the topic, we note that the results from this indicator are very similar to those obtained using BIS$^+$ on the same stars.

In a nutshell, the $\Delta V$ indicator proves to have a quantitative advantage over the BIS in the sense that the same results are obtained with higher significance, but no qualitative difference in the sense that the same conclusions on each signal's nature were reached.

**Table 9.** Comparison between BIS and $\Delta V$ correlations with RV for HARPS data for the parameters that do not consider the error bars.

| Star | $RV_{max}$ | $BIS_{max}$ | $\Delta V_{max}$ | $\rho_{BIS}$ | $\rho_{\Delta V}$ | $\sigma_{\rho,BIS}$ | $\sigma_{\rho,\Delta V}$ |
|------|-----------|-------------|------------------|--------------|-------------------|---------------------|--------------------------|
| HD 224789 | 3.53e+01 | 2.77e+01 | 7.73e+01 | −0.818 | 0.824 | 4.82 (<0.00%) | 4.87 (<0.00%) |
| HD 36051 | 2.16e+01 | 3.29e+01 | 7.38e+01 | −0.275 | 0.388 | 0.95 (34.31%) | 1.34 (18.05%) |
| HD 103720 | 1.10e+02 | 2.41e+01 | 6.15e+01 | −0.243 | 0.252 | 1.68 (9.23%) | 1.74 (8.17%) |
| HD 200143 | 4.70e+01 | 1.78e+02 | 1.80e+02 | −0.302 | 0.305 | 1.94 (5.30%) | 1.96 (5.06%) |
| BD-213253 | 5.33e+01 | 3.33e+01 | 6.16e+01 | −0.130 | 0.046 | 0.69 (49.17%) | 0.24 (80.69%) |
| HD 216770 | 2.67e+01 | 4.95e+00 | 1.42e+01 | 0.142 | −0.267 | 0.754 (45.10%) | 1.408 (15.93%) |

**Notes.** RV, BIS, and $\Delta V$ in m/s.

**Table 10.** Comparison between BIS and $\Delta V$ correlations with RV for HARPS data for the parameters that consider the error bars, i.e. the slope fit parameters.

| Star | $(m, b)_{BIS}$ | $(m, b)_{\Delta V}$ |
|------|----------------|---------------------|
| HD 224789 | $(-0.705^{+0.012}_{-0.011}$ or $^{+1.73\%}_{-1.57\%}, 26.083^{+0.408}_{-0.451})$ | $(0.831^{+0.012}_{-0.013}$ or $^{+1.40\%}_{-1.56\%}, -30.715^{+0.479}_{-0.428})$ |
| HD 36051 | $(-0.476^{+0.143}_{-0.125}$ or $^{+30.14\%}_{-26.33\%}, 0.795^{+0.202}_{-0.232})$ | $(0.635^{+0.128}_{-0.153}$ or $^{+20.18\%}_{-24.09\%}, -1.053^{+0.247}_{-0.207})$ |
| HD 103720 | $(-0.049^{+0.006}_{-0.005}$ or $^{+11.35\%}_{-11.24\%}, -0.287^{+0.036}_{-0.036})$ | $(0.061^{+0.006}_{-0.006}$ or $^{+9.04\%}_{-9.14\%}, 0.367^{+0.036}_{-0.037})$ |
| HD 200143 | $(-0.239^{+0.022}_{-0.021}$ or $^{+9.27\%}_{-8.60\%}, -4.887^{+0.453}_{-0.420})$ | $(0.259^{+0.021}_{-0.023}$ or $^{+7.98\%}_{-8.75\%}, 5.303^{+0.422}_{-0.462})$ |
| BD-213253 | $(-0.225^{+0.034}_{-0.032}$ or $^{+15.25\%}_{-14.16\%}, 5.527^{+0.779}_{-0.838})$ | $(0.184^{+0.033}_{-0.035}$ or $^{+17.90\%}_{-19.07\%}, -4.524^{+0.857}_{-0.804})$ |
| HD 216770 | $(0.014^{+0.006}_{-0.006}$ or $^{+42.81\%}_{-42.95\%}, -0.439^{+0.193}_{-0.192})$ | $(-0.028^{+0.006}_{-0.006}$ or $^{+22.20\%}_{-22.15\%}, 0.858^{+0.192}_{-0.193})$ |
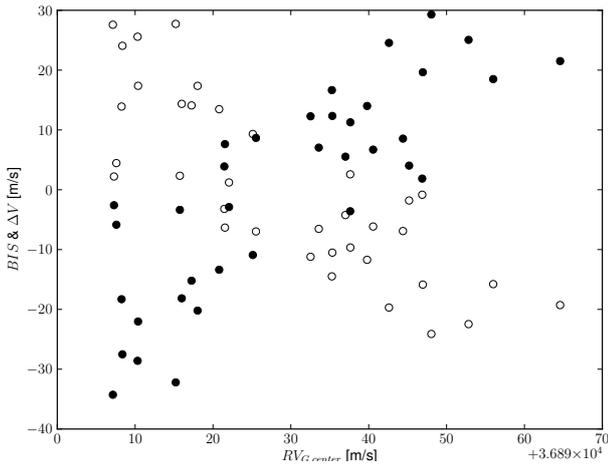


**Fig. 4.** Dependence between the indicator $\Delta V$ (filled circles) and BIS (open circles) with RV for HD 224789.

The $\Delta V$, in spite of a very different construction from that of BIS, shares important traits with this other indicator; in particular, both display an 8'-like shape in their correlation with RV (see Fig. 3). This suggested that we might be in presence of a common property of profile analysis indicators that have been created so far and it encouraged us to devise a completely different approach, as described in the next section.

# 6. The analysis of the asymmetry in the RV calculation: $V_{asy}$

## 6.1. The $V_{asy}$ indicator

We consider now an indicator that evaluates the RV information content of a spectral line and compares the information on the blue wing (i.e., for wavelengths shorter than that of the line center) to that on the red wing (i.e., for wavelengths longer than the center). For every flux level, the spectral information difference

is calculated, and the results for all fluxes are condensed onto a weighted average value for the line in which the weight is the weight attributed to the flux level. The average spectral information between the red and blue wing is used as a proxy for the weight of the undeformed line, i.e., as a weight for the flux level. We apply this procedure to the CCF and mathematically define this indicator, $V_{asy}$, as

$$V_{asy} = \frac{\sum_{flux}(W_i(red) - W_i(blue)) \times \overline{W_i}}{\sum_{flux} \overline{W_i}} \quad (3)$$

in which $W_i$ are the weights of the point calculated at the flux level $i$, as defined in Eq. (8) of Bouchy et al. (2001), where

$$W_i = \frac{\lambda^2(i)(\partial A_0(i)/\partial \lambda(i))^2}{A_0(i) + \sigma_D^2} \quad (4)$$

with $\lambda(i)$ and $A_0(i)$ representing the wavelength and flux of the spectra, respectively, and $\sigma_D^2$ the detector readout noise. Since we are working in the velocity space, the wavelength is a velocity. This is standard procedure and has been applied extensively to the CCF. Moreover we are working in a very high S/N domain with the CCF and as such, $A_0(i) + \sigma_D^2 \approx A_0(i)$. We can then rewrite the previous equation in the form of

$$W_i = \frac{RV^2(i)(\partial A_0(i)/\partial RV(i))^2}{A_0(i)} \quad (5)$$

in which RV($i$) is the RV value of each pixel (i.e. $x$-axis) of the CCF (not to be mistaken by the center of the CCF obtained by Gaussian fitting, $RV_{G,center}$).

For every flux level $i$, we consider three quantities: the blue wing information, $W_i(blue)$, the red wing information, $W_i(red)$, and the average between the two, $\overline{W_i} = (W_i(blue) + W_i(red))/2$, as representative of the weight of the information content at this flux level for an undeformed line.

We calculate the flux levels $i$ by quadratic interpolation between the CCF flux levels. We considered 100 equidistant flux levels and excluded the first and last 5 from the sum, to remove

**Table 11.** Comparison between correlations of BIS and $V_{asy}$ with RV for simulated data.

| Star | $RV_{max}$ | $BIS_{max}$ | $V_{asy,max}$ | $m_{BIS}$ | $m_{V_{asy}}$ | $\rho_{BIS}$ | $\rho_{V_{asy}}$ | $\sigma_{\rho,BIS}$ | $\sigma_{\rho,V_{asy}}$ |
|---|---|---|---|---|---|---|---|---|---|
| HD 166435' | 9.28e+01 | 1.03e+02 | 1.62e+02 | −1.068 | 1.704 | −0.990 | 0.999 | 9.80 | 9.98 |
| $F_r = 1.0\%$ | 9.98e+01 | 1.03e+02 | 1.79e+02 | −0.686 | 1.636 | −0.827 | 0.991 | 8.22 | 9.91 |
| $F_r = 2.0\%$ | 2.01e+02 | 2.10e+02 | 3.59e+02 | −0.679 | 1.620 | −0.820 | 0.991 | 8.17 | 9.85 |
| $F_r = 5.0\%$ | 5.20e+02 | 5.50e+02 | 8.94e+02 | −0.655 | 1.570 | −0.795 | 0.990 | 7.90 | 9.87 |
| $F_r = 10.0\%$ | 1.10e+03 | 1.22e+03 | 1.79e+03 | −0.599 | 1.474 | −0.741 | 0.989 | 7.38 | 9.86 |
| $v \sin i = 2.0$ | 2.21e+01 | 1.94e+00 | 3.14e+01 | −0.057 | 1.391 | −0.801 | 0.998 | 7.95 | 9.96 |
| $v \sin i = 5.0$ | 6.43e+01 | 3.72e+01 | 1.10e+02 | −0.378 | 1.576 | −0.815 | 0.988 | 8.14 | 9.84 |
| $v \sin i = 10.0$ | 1.57e+02 | 2.54e+02 | 3.02e+02 | −1.109 | 1.603 | −0.847 | 0.958 | 8.46 | 9.52 |

**Notes.** When not stated, the probability of obtaining the presented $\rho$ values from the distribution obtained by bootstrapping is lower than 1.0e-4. RV and BIS in m/s.
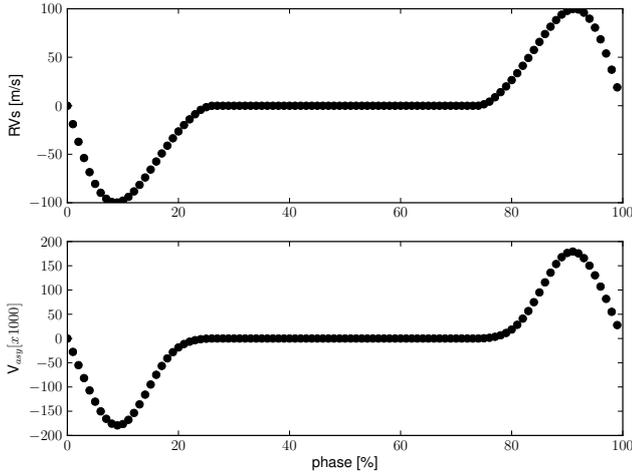


**Fig. 5.** RV (*top*) and $V_{asy}$ (*bottom*) as a function of orbital phase (with an arbitrary zero point) for the case of an equatorial spot with a a filling factor of 1% on an edge-on star with $v \sin i$ of $7 \, \mathrm{km \, s^{-1}}$.



**Fig. 6.** Two indicators considered, BIS (as open circles) and $V_{asy}$ (as filled circles), as a function of RV for the case of an equatorial spot with a a filling factor of 1% on an edge-on star with $v \sin i$ of $7 \, \mathrm{km \, s^{-1}}$.
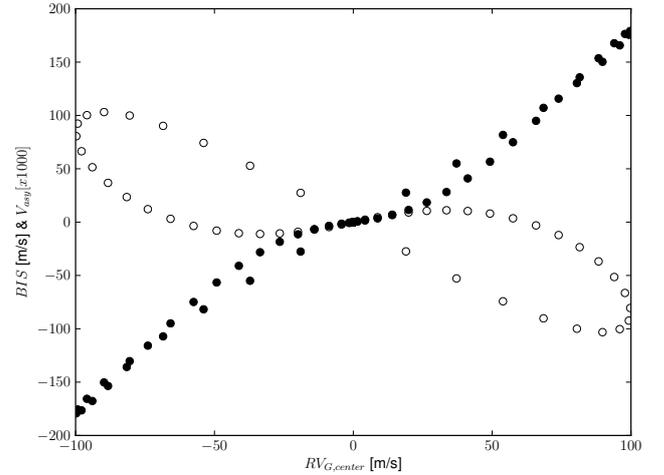
the points closer to the continuum and to the bottom of the line, where the spectral information is small compared with the noise.

In a nutshell, this indicator, $V_{asy}$, is simply the (dimensionless) average information content difference between the left and the right wing of the line. It is much closer to the RV calculation principle than line profile indicators such as BIS or $\Delta V$. Just as for the previously discussed indicators, our indicator's variation is expected to be correlated with that of RV for the case of a signal rooted in a line deformation.

### 6.2. $V_{asy}$: application to simulated data

The results of the application of this indicator to the simulated data are presented in Table 11. In Fig. 5, we plot the variation of RV and $V_{asy}$ as a function of phase for the reference case of an equatorial spot with a filling factor of 1% on a edge-on star with a $v \sin i$ of $7 \, \mathrm{km \, s^{-1}}$, and in Fig. 6, we plot how the correlation between BIS and RV compares to that between $V_{asy}$ and RV.

The amplitude variation of the indicator and how it compares with that of BIS are now meaningless, since these indicators are not calculated in the same unit. Of greater importance and insight is that the Pearson's correlation coefficients $\rho$ for $V_{asy}$ are always larger than those for the BIS or BIS$^+$; the improvement is often in the range of 10−20%. It then follows naturally that the offset in $\sigma$ relative to an uncorrelated datasets are very high and that the probabilities of obtaining these correlation coefficient values by chance are very low (even though one has to stress that they are negligibly low for all cases considered, as before).

An inspection of Fig. 6 shows that our objective was accomplished in the sense that the relation between $V_{asy}$ and RV is much closer to a linear one than that with the BIS. There is still a departure from linearity for small RV variations, but the usefulness of the method is clear.

### 6.3. $V_{asy}$: application to HARPS data

In Sect. 5.3, we applied $\Delta V$ to real data and we now repeat the same procedure and analysis for $V_{asy}$. The only difference is that we refrain from attempting to calculate the error bars on the indicator. As we saw in the mentioned section, this was the weakest point of the analysis, requiring more assumptions and making the interpretation of results more difficult. As a consequence, we also chose not to calculate the error bars on the correlation slope $m$. The results are presented in Table 12 and in Fig. 7, we depict the correlation between $V_{asy}$ and RV and how it compares with that of BIS for the case of HD 224789.

As before, we skip the analysis of the variation of the coefficient, and we note that while the values of the correlation slopes are higher this does not mean the method is more efficient in itself. These properties are in line with what was measured for the synthetic data, but they do not allow us to reach any conclusion in themselves. However, it is interesting to note that for 4 of the 5 active stars the Pearson's correlation coefficient was significantly improved, in one case the improvement was by a factor of two, and in the case of the planet-hosting star, it decreased.

**Table 12.** Comparison between correlations of BIS and $V_{asy}$ with RV for HARPS data.

| Star | $RV_{max}$ | $BIS_{max}$ | $V_{asy,max}$ | $m_{BIS}$ | $m_{V_{asy}}$ | $\rho_{BIS}$ | $\rho_{V_{asy}}$ | $\sigma_{\rho,BIS}$ | $\sigma_{\rho,V_{asy}}$ |
|------|-----------|-------------|---------------|-----------|----------------|--------------|------------------|---------------------|-------------------------|
| HD 224789 | 3.53e+01 | 2.77e+01 | 4.40e+03 | −0.756 | 116.351 | −0.818 | 0.823 | 4.83 (<0.00%) | 4.86 (<0.00%) |
| HD 36051 | 2.16e+01 | 3.29e+01 | 1.25e+02 | −0.376 | 3.439 | −0.275 | 0.690 | 0.95 (34.07%) | 2.40 (1.66%) |
| HD 103720 | 1.10e+02 | 2.41e+01 | 2.88e+02 | −0.043 | 1.409 | −0.243 | 0.688 | 1.69 (9.11%) | 4.77 (<0.00%) |
| HD 200143 | 4.70e+01 | 1.78e+02 | 5.19e+04 | −0.443 | 59.314 | −0.302 | 0.153 | 1.94 (5.30%) | 0.98 (32.64%) |
| BD-213153 | 5.33e+01 | 3.33e+01 | 1.59e+03 | −0.103 | 9.432 | −0.130 | 0.257 | 0.69 (48.93%) | 1.35 (17.54%) |
| HD 216770 | 2.67e+01 | 4.95e+00 | 8.86e+02 | 0.018 | −0.013 | 0.142 | −0.001 | 0.75 (45.18%) | 0.00 (99.68%) |

**Notes.** RV and BIS in m/s.
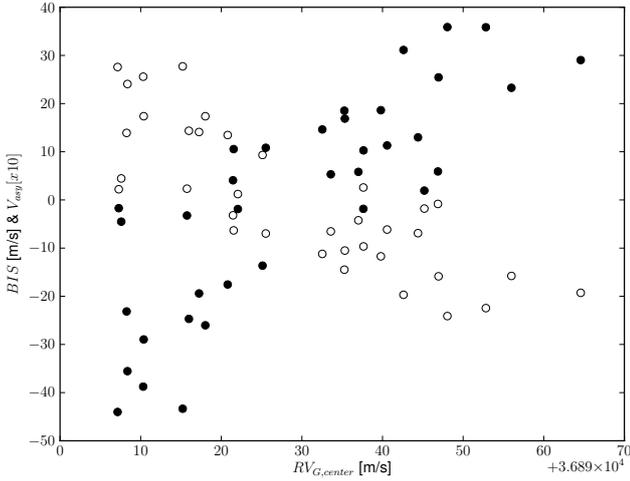


**Fig. 7.** Dependence between the indicator $V_{asy}$ (filled circles) and BIS (open circles) with RV for HD 224789.

Consequently, the deviation in $\sigma$ increased for the four active stars and the probability of the correlation arising from a fortuitous alignment of the data decreased. The opposite was obtained for the planet-host star. This provides a strong argument for the usefulness of the indicator and illustrates very well how advantageous it can be when compared with the BIS.

## 7. Discussion

We have shown that the $\Delta V$ indicator provides a more efficient characterization of line deformations than the BIS, and very similar to that of BIS$^+$. This property can make the difference when the correlation slope is comparable to the photon noise and/or the number of observations is very small. The indicator proved to be advantageous over the standard BIS for all the simulated datasets and slightly so for the real ones for which a correlation had been found between the RV and the BIS. In this sense, we can then conclude that the $\Delta V$ provides a quantitative advantage over the BIS. Yet, it is important to note that some of the active stars show RV variations without being correlated with BIS or BIS$^+$, and our new indicator did not allow to identify new correlations (for the meager number of stars tested).

Furthermore, the error bars on $\Delta V$ and BIS can be used to assess the significance of the results in another way. If one assumes that the activity signal is the only signal present in the RV data, the residuals should follow the noise distribution of the indicator once the slope is subtracted. This is expected to be a Gaussian distribution with a center at zero and a scatter dictated by the photon noise or, to be more generic, proportional to it. We consider different multiplicative factors $\alpha$ for the relation between the observed scatter and the one dictated by photon noise

**Table 13.** The $\alpha$ correction factor on photon noise used to obtain the observed noise distribution.

| Star | $\alpha(BIS)$ | $\alpha(\Delta V)$ |
|------|---------------|--------------------|
| HD 224789 | 14.5 | 20.0 |
| HD 36051 | 5.5 | 7.0 |
| HD 103720 | 7.5 | 6.0 |
| HD 200143 | 7.5 | 5.5$^a$ |
| BD-213153 | 5.0 | 5.5 |

**Notes.** $^{(a)}$ Note that the probability of correspondence for $\Delta V$ is smaller than 20%.

$\sigma_{obs} = \alpha \times \sigma_{ph}$. We evaluate them through the Kolmogorov-Smirnoff analysis (as implemented in Press et al. 1992), and evaluate which of these postulated distributions $G(\alpha\,\sigma_{ph})$ has the highest probability of coming from the same parent distribution as the observed one, and thus which $\alpha$ is more appropriate to characterize $\sigma_{obs}$. We considered $\alpha$ over an interval ranging from 0 to 30 with a step of 0.5. We list the highest-probability $\alpha$ obtained for both correlations in Table 13.

We can see from the Table that neither of the methods/correlations is intrinsically superior when it comes to noise distribution. In other words, the correction factors applied to the error bars on the indicators are larger for the BIS-RV correlation for some of the cases, and larger for the $\Delta V$-RV correlations for some other cases. They are significantly higher than unity in both situations, which should send a warning to all those who use the BIS indicator routinely. Yet, one has to bear in mind that the fit of a straight line on the data is ill-justified, since we know that the real shape is closer to an 8 rather than to a straight line. While the results still hold for a relative comparison, an $\alpha$ larger than unity was always expected.

We stress again that non-parametric tests are particularly important when the properties of the distributions are unknown or ill-defined, and we recall the usefulness of the permutation test to the reader as a way of assessing the significance of a correlation.

As discussed by Nardetto et al. (2006), the fitting makes use of the full information of the line and will still deliver precise results at low S/N, which is a level at which BIS is significantly more vulnerable. The reason for this is two-fold: BIS requires interpolation to calculate the flux levels and does not make use of all the spectral information content of the line. It is calculated only over a restricted range of flux. This is particularly important for those who use CCFs constructed from a relatively low number of lines or aim at measuring RV from low-S/N spectra. Neither is common for typical RV planet searches, but the latter is frequent when using RV measurements to confirm transit candidates from space observatories, such as CoRot (Csizmadia et al. 2011) or *Kepler* (Endl et al. 2011; Santerne et al. 2012). Interestingly, a significant source of false positives in transit searches is the blending of the spectral lines with those of a

foreground or background star. The usage of indicators, such as the BIS or those presented in this paper, can again be of help, but this is a complex issue and beyond the reach of this paper. We will discuss it in a separate paper, using meaningful false positive examples from transit searches (Santerne et al., in prep.).

The results on $V_{asy}$ are even more promising than those of $\Delta V$, with the indicator being constructed in a completely different way. The results point for more secure correlations between $V_{asy}$ and the RV for most of the analyzed active stars (4 out of 5) compared to those obtained using the BIS. It provides both a qualitative and quantitative advantage over the bisector in the sense that new correlations arise for active stars, and those previously identified are detected with higher significance. Yet, it has to be noted that even though the correlation is much closer to a linear one than that of the BIS, a departure from the linear relationship is apparent for the smallest amplitude RV variations. This attests for an undersensitivity of the indicator in this regime and warrants further investigation.

In this work we made a distinction and a separated analysis of synthetic spectra, without noise, and real spectra, with noise. This was a choice prompted by the availability of the SOAP model, which delivers noise-free CCFs, but allowed us to explore and distinguish expected theoretical results and observed ones. While one could have attempted to introduce noise in the CCFs and compare it with the real data, this would have to be done *before* the CCF computation, i.e., at the spectral generation level, using a completely different and by far more complicated procedure, which is highly dependent on the particular assumptions on the spectra. By construction, the calculation of the CCF transforms white noise into correlated noise, simply because the correlation step for the calculation is smaller than the pixel size. Moreover, a real spectra is expected to have its own noise correlated from one pixel to the other; this is a consequence of atmospheric extinction being a function of wavelength, just like instrumental transmission and instrumental profile. As a consequence, any model of the spectra which does not fully characterize the spectrograph, would fall short in describing its correlated noise, and thus lead to an incorrect evaluation of its impact on the RV and associated indicator's calculation.

It has been suggested recently that there is a significant difference between the results derived from a simple spot model (such as SOAP) and comprehensive RV campaigns (e.g. Crockett et al. 2012). In this work, the parametrized synthetic cases show similar order-of-magnitude variations in RV and indicators compared to the real ones. Yet, one cannot exclude that there might be physical mechanisms that are not represented and might produce an appreciable effect. Works such as those of Meunier et al. (2010) showed the importance of considering the reduction of convective blueshift at the spot's location on the measurement of precise stellar RV. This is only one of the many points not at all adressed by SOAP, since it emulates a very simple stellar model, and does not even consider emission of any kind at the spot's location, nor the nuances in its emitted spectra. While a cold-spot model might be indeed an oversimplification, we note that the methods discussed here proved to be effective not only for simulated data but for real data too.

It is also important to stress that the RV center obtained for a deformed line depends on the RV calculation method. For the implications of this characteristic on the case of Rossiter-MacLaughlin measurements, the interested reader is referred to Boué et al. (2013). This is a point to keep in mind when comparing the results obtained here with those obtained for RVs acquired with the iodine cell technique (e.g. Buchhave et al. 2011).

We conclude this discussion on a different note: instead of considering these different methods as alternative options, one should look at them as independent ways of characterizing the same phenomena and use them simultaneously, while bearing in mind their relative strengths and weaknesses. The active star BD-123153 illustrates this very well: it exhibits a non-significant correlation between BIS and RV or $\Delta V$ and RV but shows a more significant one between $V_{asy}$ and RV or BIS$^-$ and RV. The use of different tests is then very important, especially because while the presence of a correlation disproves the planetary nature of a signal, its absence does not prove its planetary nature. We tried to do a systematic assessment of the potential of the presented methods, but the work presented here should be seen merely as the first step toward a more detailed characterization of these diagnosis methods and the development of complementary approaches.

## 8. Conclusion

In this paper, we introduced two diagnosis methods for detecting line-profile variations in the context of exoplanet searches. The objective is to evaluate if these variations can be pinpointed as being at the root of detected RV signals. We attempted at a systematic assessment of the efficiency and sensitivity of our presented methods when compared with that of BIS, assuming different parametrizations, and $V_{span}$, and following very clear criteria. Our conclusions can be outlined as follows:

– The standard BIS delivers very similar results to those of $V_{span}$ for the cases considered. Different parametrizations of BIS can be explored to increase the leverage of the diagnosis at the expense of considering a lower fraction of the line and thus working at lower S/N. Yet, statisticaly more significant results can be obtained, showing that this is an option to consider.

– The bi-Gauss/$\Delta V$ method is superior to the standard BIS in a multitude of cases. It presents no qualitative advantage, identifying the same signals than BIS as activity-rooted, but provides more significant results with the characterization parameters being affected by smaller error bars. This point can make the difference for cases where the correlation slope between the indicator and RV is close to the noise level and puts the indicator in even footing with the more effective BIS$^+$ parametrization.

– The $V_{asy}$ indicator evaluates the unbalance between the red and blue wings of a line in terms of spectral information, being by construction very different from the other profile indicators. It shows a significant advantage over different parametrizations of BIS and $\Delta V$, when applied to simulated spectra and to real spectra. The indicator leads to more significant correlations, and identifies new signals rooted in activity, providing a clear qualitative advantage. It also gives rise to a correlation with RV which is much closer to a straight line than any of the other profile-based indicators.

Finally we argue for the usage of independent techniques to characterize stellar line profile variations and to identify their origins, creating a coherent view of the stellar phenomena. This is particularly important in cases where the signal and/or potential correlation slopes approach the RV uncertainty level. A more detailed knowledge of the diagnosis methods themselves is required to meet the exigent requirements of today's scientific programs. The work presented here should be considered as a first step in this direction; a more significant amount of resources

should be put into characterizing these indicators and devising new and independent ones.

## References

Baranne, A., Queloz, D., Mayor, M., et al. 1996, A&AS, 119, 373
Boisse, I., Bouchy, F., Hébrard, G., et al. 2011, A&A, 528, A4
Boisse, I., Bonfils, X., & Santos, N. C. 2012, A&A, 545, A109
Bouchy, F., Pepe, F., & Queloz, D. 2001, A&A, 374, 733
Boué, G., Montalto, M., Boisse, I., Oshagh, M., & Santos, N. C. 2013, A&A, 550, A53
Buchhave, L. A., Bakos, G. Á., Hartman, J. D., et al. 2011, ApJ, 733, 116
Crockett, C. J., Mahmud, N. I., Prato, L., et al. 2012, ApJ, 761, 164
Csizmadia, S., Moutou, C., Deleuil, M., et al. 2011, A&A, 531, A41
Dall, T. H., Santos, N. C., Arentoft, T., Bedding, T. R., & Kjeldsen, H. 2006, A&A, 454, 341
Desort, M., Lagrange, A.-M., Galland, F., Udry, S., & Mayor, M. 2007, A&A, 473, 983
Dumusque, X., Lovis, C., Ségransan, D., et al. 2011, A&A, 535, A55
Endl, M., MacQueen, P. J., Cochran, W. D., et al. 2011, ApJS, 197, 13
Figueira, P., Marmier, M., Bonfils, X., et al. 2010a, A&A, 513, L8
Figueira, P., Pepe, F., Melo, C. H. F., et al. 2010b, A&A, 511, A55
Hébrard, G., Udry, S., Lo Curto, G., et al. 2010, A&A, 512, A46
Henry, T. J., Soderblom, D. R., Donahue, R. A., & Baliunas, S. L. 1996, AJ, 111, 439
Huélamo, N., Figueira, P., Bonfils, X., et al. 2008, A&A, 489, L9
Lovis, C., Dumusque, X., Santos, N. C., et al. 2011, A&A, submitted [arXiv:1107.5325]
Martín, E. L., Guenther, E., Zapatero Osorio, M. R., Bouy, H., & Wainscoat, R. 2006, ApJ, 644, L75
Martínez Fiorenzano, A. F., Gratton, R. G., Desidera, S., Cosentino, R., & Endl, M. 2005, A&A, 442, 775
Mayor, M., & Queloz, D. 1995, Nature, 378, 355
Mayor, M., Pepe, F., Queloz, D., et al. 2003, The Messenger, 114, 20
Mayor, M., Udry, S., Naef, D., et al. 2004, A&A, 415, 391
Meunier, N., Lagrange, A.-M., & Desort, M. 2010, A&A, 519, A66
Nardetto, N., Mourard, D., Kervella, P., et al. 2006, A&A, 453, 309
Niedzielski, A., Nowak, G., Adamów, M., & Wolszczan, A. 2009, ApJ, 707, 768
Pepe, F., Mayor, M., Galland, F., et al. 2002, A&A, 388, 632
Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. 1992, Numerical recipes in FORTRAN, The art of scientific computing
Queloz, D., Henry, G. W., Sivan, J. P., et al. 2001, A&A, 379, 279
Saar, S. H., & Donahue, R. A. 1997, ApJ, 485, 319
Santerne, A., Díaz, R. F., Moutou, C., et al. 2012, A&A, 545, A76
Santos, N. C., Mayor, M., Naef, D., et al. 2000, A&A, 361, 265
Santos, N. C., Israelian, G., Mayor, M., Rebolo, R., & Udry, S. 2003a, A&A, 398, 363
Santos, N. C., Udry, S., Mayor, M., et al. 2003b, A&A, 406, 373