

A HERMITE–GAUSSIAN BASED EXOPLANET RADIAL VELOCITY ESTIMATION METHOD

BY PARKER H. HOLZER¹, JESSI CISEWSKI-KEHE², DEBRA FISCHER^{3,*} AND LILY ZHAO^{3,†}

¹Department of Statistics and Data Science, Yale University, parker.holzer@yale.edu

²Department of Statistics, University of Wisconsin–Madison, jjkehe@wisc.edu

³Department of Astronomy, Yale University, * debra.fischer@yale.edu; † lily.zhao@yale.edu

As the first successful technique used to detect exoplanets orbiting distant stars, the radial velocity method aims to detect a periodic Doppler shift in a stellar spectrum due to the star’s motion along the line sight. We introduce a new, mathematically rigorous approach to detect such a signal that accounts for the smooth functional relationship of neighboring wavelengths in the spectrum, minimizes the role of wavelength interpolation, accounts for heteroskedastic noise and easily allows for accurate calculation of the estimated radial velocity standard error. Using Hermite–Gaussian functions, we show that the problem of detecting a Doppler shift in the spectrum can be reduced to linear regression in many settings. A simulation study demonstrates that the proposed method is able to accurately estimate an individual spectrum’s radial velocity with precision below 0.3 m s^{-1} , corresponding to a Doppler shift much smaller than the size of a spectral pixel. Furthermore, the new method outperforms the traditional cross-correlation function approach for estimating the radial velocity by reducing the root mean squared error up to 15 cm s^{-1} . The proposed method is also demonstrated on a new set of observations from the EXtreme PREcision Spectrometer (EXPRES) for the host star 51 Pegasi, and successfully recovers estimates of the planetary companion’s parameters that agree well with previous studies. The method is implemented in the R package *rvmethod*, and supplemental Python code is also available.

1. Introduction. The discovery of a planet orbiting the Sun-like star 51 Pegasi (Mayor and Queloz (1995)) launched a new subfield in astronomy, the detection and characterization of planets orbiting other main sequence stars (i.e., exoplanets). This discovery was made using the radial velocity (RV) method (also known as the Doppler method). The RV method aims to detect an oscillatory motion in an observed star, possibly due to the gravitational pull of one or more orbiting planets. At each time of observation, the relative motion of the star results in a slight rescaling of the wavelength of the stellar light; this rescaling is often referred to as a *Doppler shift*.

The data for the RV method are obtained with a spectrograph and are referred to as the stellar spectrum. The optical elements in the spectrograph disperse light from the star into component wavelengths and focus the spectrum onto an electronic detector. The pixels in the detector sample the stellar spectrum. The continuous stellar spectrum is imprinted with thousands of narrow absorption lines that form when atoms and molecules in the stellar atmosphere of the star absorb light at specific wavelengths. The wavelength of such spectral lines corresponds to the quantum mechanical energy level differences in the absorbing atoms and molecules of the stellar atmosphere. As the star moves toward us or away from us, the velocity component that is projected along our line of sight (i.e., the radial velocity) produces

Received July 2019; revised October 2020.

Key words and phrases. Exoplanets, Hermite–Gaussian functions, linear regression, astrostatistics, Doppler effect.

a wavelength rescaling in the spectrum that is described by the Doppler equation. Overall, a Doppler shift is a wavelength rescaling and is more readily observable at wavelengths where spectral lines form.

All stars orbit the galaxy and will exhibit a nearly constant RV relative to the Sun. If a star also has a planet, then the orbiting planet will exert a gravitational force on the star around a common center of mass. By measuring this varying reflex velocity in the stellar spectrum over time, the orbital parameters of a planetary companion can be derived.

The magnitude of the RV signal depends on several factors, including the mass of the star, the mass of the planet, the planet's orbital period, the shape of the orbit (i.e., eccentricity) and the orientation of the orbital plane. Since orbits that are oriented "face-on" are tangential to our line of site, they do not have a radial component and therefore cannot be detected with the RV method. Fortunately, face-on orbits are a statistically rare configuration.

In the solar system, Jupiter induces a RV with a magnitude up to about 12 m s^{-1} for the Sun while the lower mass Earth only induces a RV up to about 0.09 m s^{-1} . If observed with high spectral resolution of 1.5×10^5 , one pixel on the detector spans about 500 m s^{-1} , so these RVs would only shift the solar spectrum by 0.024 or 0.00018 pixels, for Jupiter and the Earth, respectively. Further complicating the detection, these tiny shifts are merely the semiamplitudes of nearly sinusoidal RV variations with periods of about 12 years for Jupiter and 1 year for the Earth. Because the RV signal is so tiny, compared to the speed of light, detecting such a subpixel shift in stellar absorption features is nontrivial. The state-of-the-art Doppler precision for the past decade has been about 1 m s^{-1} (Fischer et al. (2016)). This is sufficient to detect Jupiter (with 12 years of observations) but precludes the detection of Earth analogs around Sun-like stars. Because the RV amplitude increases with decreasing stellar mass (all else equal), some Earth mass planets have been detected around stars that are lower in mass than the Sun. Figure 1 shows the velocity amplitudes and orbital periods of exoplanets detected using the RV method over the past 25 years.

The RV error budget includes instrumental errors, photon noise and velocities from within the atmosphere of the star that introduce scatter to the center-of-mass RV (Dumusque et al.

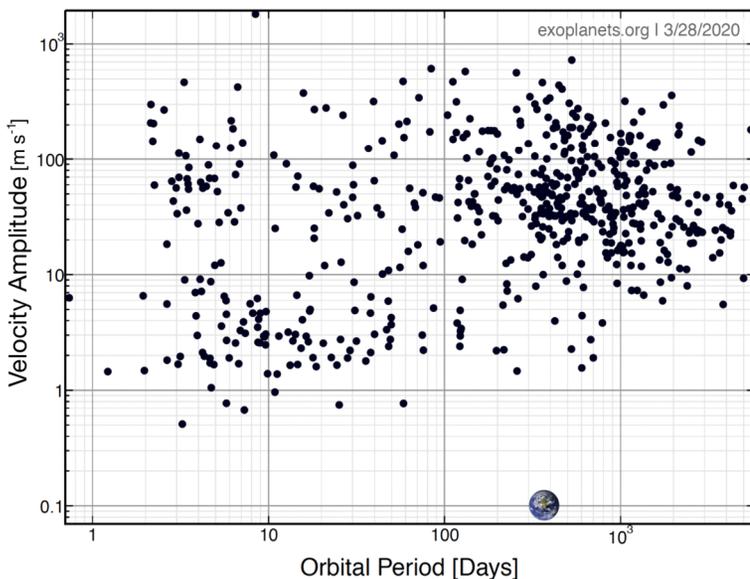


FIG. 1. Orbital period and stellar RV semiamplitude for all exoplanets discovered with the RV method. Data come from Exoplanets.org (Han et al. (2014)) on March 28, 2020, with a total of about 800 exoplanets. Note the symbol of the Earth at an orbital period of 365.25 days and a velocity amplitude of approximately 0.1 m s^{-1} and that analogs of the Earth were not detectable.

(2017), Halverson et al. (2016), Blackman et al. (2020)). The EXtreme PREcision Spectrometer (EXPRES) (Blackman et al. (2020), Jurgenson et al. (2016), Petersburg et al. (2020)) is a newly commissioned instrument that was designed to significantly reduce instrumental errors. The primary goal of the EXPRES instrument is to provide higher-fidelity data (high signal-to-noise with reduced instrumental errors) and has demonstrated intrinsic instrumental measurement precision better than 0.1 m s^{-1} (Blackman et al. (2020)). The next critical step for reaching Earth-detecting precision is the development of statistical techniques that estimate RVs with high precision and are less sensitive to atmospheric velocities (Davis et al. (2017), Dumusque, Boisse and Santos (2014), Dumusque et al. (2017), Rajpaul et al. (2015), Rajpaul, Aigrain and Buchhave (2020)).

The traditional cross-correlation function (CCF) method (Baranne et al. (1996)) has long been used to measure Doppler shifts in stellar spectra by minimizing a weighted dot product between the observed spectrum and a template (Pepe et al. (2002)). Various template matching algorithms have also been developed which minimize the (interpolated) sum of squared differences between the spectrum and a template spectrum using the Doppler shift as a free parameter (Anglada-Escudé and Butler (2012), Astudillo-Defru et al. (2015)). A variant of the template matching approach assumes the Doppler shift is small and estimates the derivative of the spectrum from the template (Bouchy, Pepe and Queloz (2001), Dumusque (2018)). The EXPRES analysis pipeline has implemented the CCF method as well as a higher precision forward-modeling (FM) code that makes use of a very high signal-to-noise (S/N) stellar template to model a Doppler shift in every $2\text{-}\text{\AA}$ segment of the observed spectrum (Petersburg et al. (2020)).

The new method we propose for estimating the RV is designed to work well in the small RV regime typical of orbiting exoplanets. Additionally, the proposed method is developed to generalize well to other stars with enough discernible absorption lines. This is because the modeling is carried out on the spectra observed for an individual star and does not require a pre-specified template. The only interpolation that takes place in the proposed method is on a high S/N, oversampled, template spectrum (derived from the data). Compared to the approach of Anglada-Escudé and Butler (2012), which requires interpolation of every (low S/N) observed spectrum, the numerical error introduced through interpolation is likely reduced in the new proposed method. Perhaps most importantly, the new method simplifies the RV estimation process to simple linear regression without the need for getting a nonparametric estimate of both the template spectrum and its derivative. This allows the method to easily account for the heteroskedastic noise in spectra. Furthermore, this simplification allows for straightforward statistical inference on the estimated RV without making assumptions regarding the validity of “propagation of error” (i.e., the delta method) or other approximate estimates of the standard error.

The proposed Hermite–Gaussian Radial Velocity (HGRV) estimation method makes use of the well-known Hermite–Gaussian functions. These functions have been used extensively in solving Schrodinger’s equation for models of quantum systems (Marhic (1978), Dai, Wang and Liu (2016)) as well as in fitting emission lines in galaxy spectroscopy (Riffel (2010)). The key contribution of this paper is that shifts of spectral lines between two spectra (e.g., due to a Doppler shift) can be well estimated with the first Hermite–Gaussian function fitted to the *difference* spectrum.

The use of the Hermite–Gaussian functions is mainly motivated by the method’s assumption that absorption features are Gaussian shaped (an assumption that can be generalized). It is important to note that large optical depth, rotational broadening, collisional broadening, stellar activity and other astrophysical effects can cause absorption features to depart from a Gaussian shape. (The model misspecification due to this Gaussian-shape assumption is explored in Section 3.4.)

In Section 2 we introduce the data commonly used in the RV method, namely, stellar spectra. We also propose an algorithm for finding absorption features in the spectrum that will be used in the HGRV method. Section 3 includes details of the proposed HGRV method, and simulation study results are discussed in Section 4. Section 5 then applies the method to recently collected data of 51 Pegasi by EXPRES. A discussion is provided in Section 6, and we conclude in Section 7.

2. Absorption feature finding algorithm. A small section of the Sun’s spectrum, as collected by the National Solar Observatory (NSO) (Rimmele and Radick (1998)), is shown in Figure 2. In general, such a spectrum gives a representation of the relative brightness (hereafter referred to as normalized flux¹) as a function of wavelength. The narrow dips in the normalized flux are spectral absorption features which have variable intensity and frequent blending with neighboring features. In the (unrealistic) situation of these absorption features not being present, the remaining spectrum is referred to as the continuum.

The blackbody spectrum (Planck (1901)), together with the instrumental effect, often referred to as the blaze function, lead to a continuum that is not flat in the raw spectrum. However, various normalization techniques have been developed to correct for these effects (e.g., Xu et al. (2019), Petersburg et al. (2020)). A spectrum where the continuum has been normalized (i.e., flattened with a maximum amplitude of one plus noise) is hereafter referred to as a *normalized spectrum*. Figure 2 is an example of such a normalized spectrum.

We define the template spectrum of a star, τ , to be its noiseless, normalized spectrum with no instrumental or astrophysical effects (e.g., activity such as spots). Furthermore, we define the difference flux to be the difference between a single observed normalized spectrum and this template. An important characteristic of the HGRV method is that, rather than modeling a Doppler shift in the spectrum as a change in the explanatory variable (wavelength) as the CCF method does, we can model the difference in normalized flux caused by the Doppler shift. This characteristic is present in various other RV detection methods (Bouchy, Pepe and Queloz (2001), Rajpaul, Aigrain and Buchhave (2020)), but it is implemented rather differently with our proposed method.

Since a Doppler shift only rescales the wavelength axis, there is little RV information in the normalized continuum. Most of the information for small Doppler shifts comes from the high-derivative regions of spectral lines, so identification of the absorption features in a given spectrum is the first step for the HGRV method.

The locations, depths and degree of blending of absorption features depend on the stellar parameters and chemical composition of the star and, therefore, vary from star to star. The HGRV method involves modeling individual absorption features, so an algorithm is needed

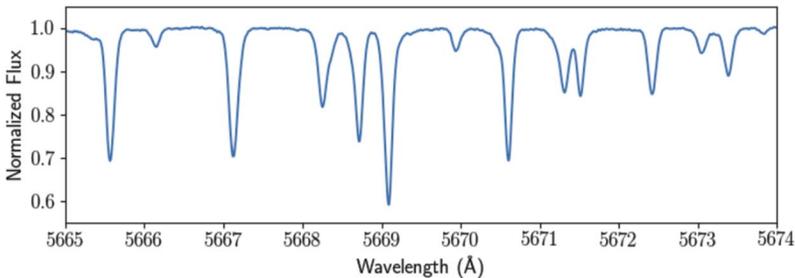


FIG. 2. A subset of the NSO spectrum of the Sun between 5665 and 5674 Å.

¹Throughout this paper we acknowledge that what we refer to as “flux” is not a true physical flux.

that not only identifies the central wavelength at which each feature occurs but also the wavelength bounds that contain the feature. Were all absorption features to be well separated, these wavelength bounds would nearly be symmetric about the central wavelengths with a nearly constant width. However, since blends are very common, this is not the case in practice.

Designing the HGRV method to generalize across stars motivates the use of an algorithm for identifying absorption feature wavelength bounds in a way that can adapt to different spectra. The proposed absorption feature finding algorithm is a statistically motivated heuristic algorithm. The overarching goal is to find wavelength windows of absorption features, not to perform any statistical inference on them.

The algorithm has two main sequential steps: (i) identify local minima that are likely to be absorption lines and (ii) proceed outward from each local minimum until the normalized flux flattens out. This algorithm is detailed in Section 1 of the Supplementary Material (Holzer et al. (2021)) as Algorithm 1 and is illustrated with the flowchart shown in Figure 3. This algorithm requires three tuning parameters: a wavelength window size m in units of pixel count and significance levels α , η where $\eta \geq \alpha$. For a more thorough motivation of this algorithm as well as a more detailed overview of the steps involved, see Section 1 of the Supplementary Material (Holzer et al. (2021)).

This algorithm was empirically evaluated using the NSO spectrum. After the step-by-step optimization of the three tuning parameters described in Section 1 of the Supplementary Material (Holzer et al. (2021)), we found that $m = 25$, $\alpha = 0.01$, and $\eta = 0.05$ found the most absorption features. Furthermore, we visually identified no false positives remaining after eliminating features with a line depth less than 0.015. A subset of the absorption features found in the NSO spectrum are shown in Figure 4.

To estimate the false-positive rate of this algorithm, we considered the NSO spectrum between 5000 and 6000 Å and replaced the normalized flux axis with a flat 500 S/N simulation 20 independent times. See Sections 3.5 and 4.1 for details on how we estimate a template spectrum with this level of S/N and which we use in the absorption feature finder (AFF) algorithm. Applying the AFF algorithm to these simulations with parameters $m = 25$, $\alpha = 0.01$ and $\eta = 0.05$ gave a total of 55 detected features. Since the spectra did not have any absorption features, this approximates the false positive rate as *one* absorption feature per 363 Å. Additionally, the line depths of these 55 false features had mean 0.0046, standard deviation 0.0018 and maximum 0.0098 so that all the false lines would be eliminated with the minimum line depth parameter set to 0.015. Note that for spectra with either different S/N or resolution m , α , η , and the minimum line depth may need to be adjusted (e.g., a lower S/N or resolution may need higher significance levels or a higher minimum line depth). We recommend setting m to be approximately $25 \times \frac{R}{2 \times 10^6}$, where R is the resolution of the spectrum and the minimum line depth to be approximately $0.015 \times \frac{500}{S/N}$. This recommended threshold is, in part, motivated by the idea that avoiding false lines, which can corrupt the RV estimate, may be more beneficial than detecting every true line. For details on this recommendation, see Section 1 of the Supplementary Material (Holzer et al. (2021)).

In addition, we applied the AFF algorithm directly to the NSO spectrum between 5000 and 6000 Å. We found that the wavelength bounds given by the algorithm contained 64.3% of the spectrum but accounted for 97.7% of the mean squared deviations from 1.0 of the normalized flux. The remaining 2.3% was mostly due to occasional absorption features whose overall shape, due to line blends, seemed to contribute to the algorithm missing them. For some additional plots associated with these results, see Section 1 of the Supplementary Material (Holzer et al. (2021)).

The proposed algorithm may have difficulty distinguishing two spectral lines that are strongly blended together because the slope of the normalized flux may not flatten out between the two lines. Depending on the S/N of the spectrum, it may not be able to find small

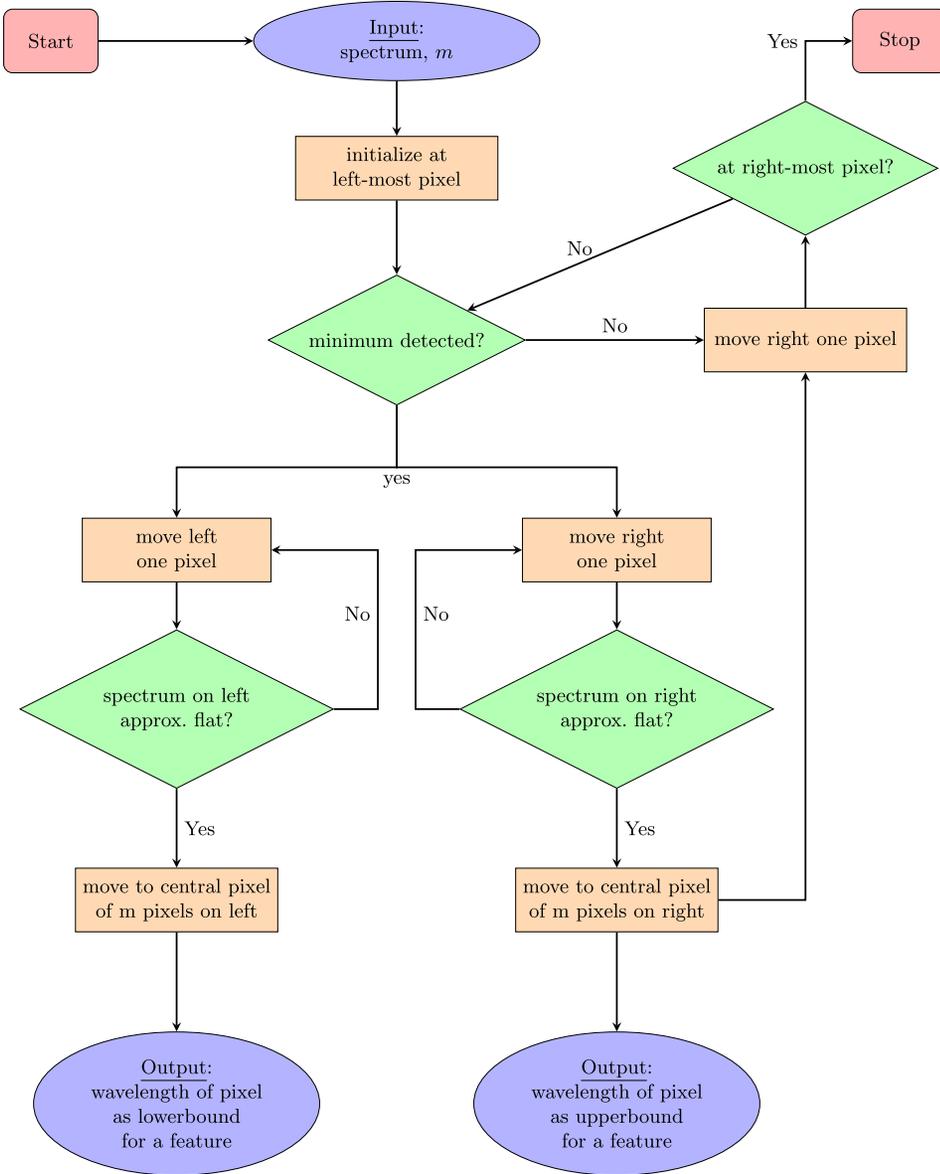


FIG. 3. Absorption feature finder flowchart for Algorithm 1 provided in Section 1 of the Supplementary Material (Holzer et al. (2021)). The algorithm takes a (potentially high S/N) normalized spectrum and a tuning parameter m that represents the number of pixels to consider at a time. The algorithm begins at the left-most pixel and proceeds to the right until it has reached a pixel that represents a statistically significant minimum (a minimum that is likely to not just be due to noise). At this point the algorithm proceeds outward in both directions until the spectrum (m pixels further outward) flattens out. The bounds of the detected absorption feature are stored, and the algorithm proceeds to the right, again looking for a statistically significant minimum.

features, as the noise would reduce the statistical significance of the left and right slopes. The lower the S/N is, the narrower the wavelength bounds will be for each detected absorption feature. This is because, as we move outward from the central wavelength of a feature, the slope eventually decreases in magnitude and becomes statistically insignificant sooner in the presence of more noise. We find that, as long as the spectrum has a S/N above 500, the results of our algorithm are stable whether or not one accounts for the heteroskedastic nature of the noise. We use the estimated template spectrum (described in Section 3.5) in the AFF

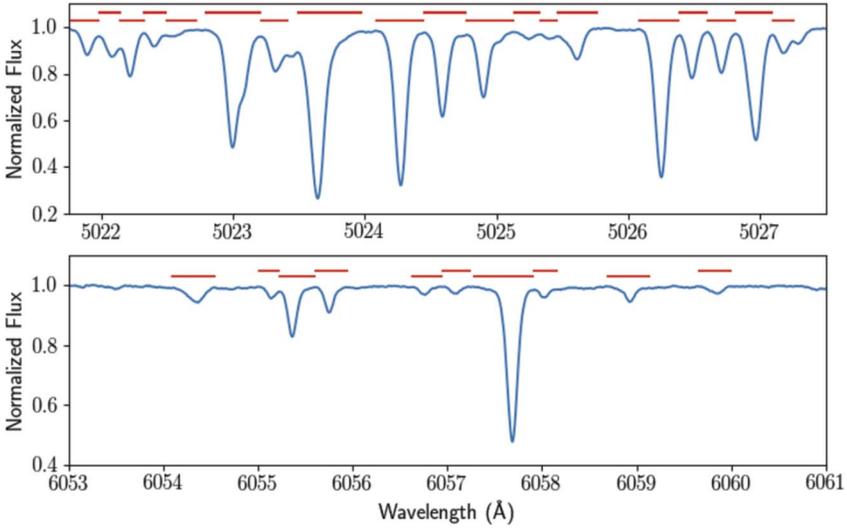


FIG. 4. Results of using the absorption feature finder algorithm on the NSO Spectrum. Red horizontal lines show the wavelength windows found to correspond to individual absorption features.

algorithm, and demonstrate in Section 4.1 that the template has a S/N above 500 as long as there are at least 11 observed spectra provided.

3. Hermite–Gaussian RV method. We now introduce the HGRV method by first considering the difference imposed on a Gaussian by a multiplicative shift of its argument. We introduce a theorem that quantifies the approximation error made by using only the first-degree Hermite–Gaussian function to model this difference and provide the proof through four lemmas (the proofs of which can be found in Section 2 of the Supplementary Material (Holzer et al. (2021))). We then show that, in the context of stellar spectroscopy, this approximation error is small and the coefficient of the first-degree Hermite–Gaussian function is nearly a constant multiple of the RV. This allows us to extend to the case of multiple absorption features and reduce the problem of estimating the Doppler shift in a spectrum to linear regression.

3.1. *Mathematics of a Doppler-shifted Gaussian.* If x represents the wavelength of light and $f(x)$ represents the normalized flux of light at that wavelength, then the normalized flux of Doppler-shifted light is represented mathematically as $f(\xi x)$ where $\frac{1}{\xi}$ is referred to as the Doppler factor (Doppler (1842)). In special relativity, ξ is given by

$$(1) \quad \xi = \frac{1 + v_r/c}{\sqrt{1 - (v/c)^2}},$$

where c is the speed of light (Einstein et al. (1905)), v is the absolute speed of the source and v_r is the velocity along the line of site of the observer. While the Earth’s rotation and revolution around the solar system barycenter often lead to relativistic effects, these motions are well understood and can be corrected for with high precision (Blackman et al. (2017), Wright and Eastman (2014), Blackman et al. (2020)). Furthermore, the velocity due only to orbiting exoplanets is well below the speed of light. Therefore, under the assumption that the barycentric corrections are applied accurately and $v \ll c$, ξ can be well approximated with the classical formula

$$(2) \quad \xi = 1 + \frac{v_r}{c}.$$

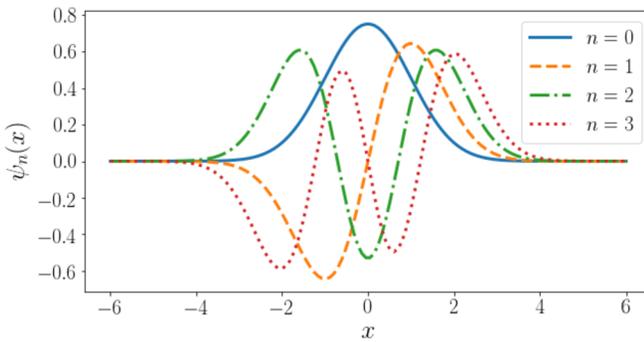


FIG. 5. The first four Hermite–Gaussian functions given by equation (3).

Consider the effect of a Doppler shift when $f(x)$ is a Gaussian like many of the inverted absorption features in a spectrum (Gray (2005)). To model this, we propose the Hermite–Gaussian functions, $\psi_n(x)$, defined as

$$(3) \quad \psi_n(x) = \frac{1}{\sqrt{2^n n! \sqrt{\pi}}} H_n(x) e^{-(x^2)/2},$$

where $H_n(x)$ represents the n th degree (physicist’s) Hermite polynomial, which can be written in closed form as

$$(4) \quad H_k(s) = k! \sum_{m=0}^{\lfloor k/2 \rfloor} \frac{(-1)^m}{m!(k-2m)!} (2s)^{k-2m},$$

with $\lfloor a \rfloor$ representing the floor function that returns the largest integer less than or equal to the real number a (Lanczos (1938)).

An illustration of the first four Hermite–Gaussian functions is shown in Figure 5.

According to Johnston (2014),

$$(5) \quad \int_{-\infty}^{\infty} H_n(x) H_m(x) e^{-x^2} dx = \sqrt{\pi} 2^n n! \mathbb{1}\{m = n\}$$

is a well-known fact about the Hermite polynomials, where $\mathbb{1}\{A\}$ represents the indicator function of the event A (which is equivalent to the Kronecker delta function).

Therefore, we have, by combining equations (3) and (5), that

$$(6) \quad \int_{-\infty}^{\infty} \psi_n(x) \psi_m(x) dx = \mathbb{1}\{m = n\}.$$

Furthermore, one can show that the set of Hermite–Gaussian functions forms a complete orthonormal basis of the set of all square-integrable real-valued functions, $L^2(\mathbb{R})$ (Johnston (2014)). One can also generalize the definition of the Hermite–Gaussian functions to have a general location, μ , and scale, σ :

$$(7) \quad \psi_n(x; \mu, \sigma) = \frac{1}{\sqrt{\sigma 2^n n! \sqrt{\pi}}} H_n\left(\frac{x - \mu}{\sigma}\right) e^{-\frac{(x - \mu)^2}{2\sigma^2}}.$$

By a simple change of variables, one can show that the set of generalized Hermite–Gaussian functions, $\psi_n(x; \mu, \sigma)$, also forms a complete orthonormal basis of $L^2(\mathbb{R})$ for any $\mu \in \mathbb{R}$ and any $\sigma \in \mathbb{R}^+$, the positive real numbers. Therefore, for such an $L^2(\mathbb{R})$ function g , we can decompose it as

$$(8) \quad g(x) = \sum_{n=0}^{\infty} c_n \psi_n(x; \mu, \sigma).$$

In this instance, let $f(x)$ be a Gaussian with center μ and width σ , and let $g(x; \xi) = f(x) - f(\xi x)$ be the difference between $f(x)$ and its Doppler-shifted version. Decomposing this $g(x; \xi)$ as in equation (8), we have Theorem 1, giving the approximation error when only $n = 1$ is used.

THEOREM 1. For any $\sigma \in \mathbb{R}^+$ and any $\mu, \xi \in \mathbb{R}$ and $g(x; \xi) = \exp(-\frac{(x-\mu)^2}{2\sigma^2}) - \exp(-\frac{(\xi x-\mu)^2}{2\sigma^2})$ decomposed in the Hermite-Gaussian basis as $g(x; \xi) = \sum_{n=0}^{\infty} c_n(\xi) \psi_n(x; \mu, \sigma)$,

$$(9) \quad \lim_{\xi \rightarrow 1} \frac{\int_{-\infty}^{\infty} (g(x; \xi) - c_1(\xi) \psi_1(x; \mu, \sigma))^2 dx}{\int_{-\infty}^{\infty} (g(x; \xi))^2 dx} = \frac{1}{1 + \frac{2\mu^2}{3\sigma^2}}.$$

Before proving Theorem 1, we interpret it in the context of stellar spectroscopy. It is well known that many absorption features in the spectrum of a star are described by the Voigt profile (Ciuryło (1998), Gray (2005)) which is well approximated by a Gaussian for many absorption features in stellar spectra. It is also the case that the central wavelength, μ , is significantly larger than the width, σ , for each of these features. As an example, a typical wavelength in the visible spectrum is 5000 Å, and the largest features near this wavelength have a width that is upper-bounded by 0.5 Å; the maximum width of absorption features detected between 4700 Å and 5300 Å by the AFF algorithm for the data collected from 51 Pegasi by EXPRES was 0.366 Å with the 88'th quantile being 0.1 Å (more details to come in Section 5). For a feature with center 5000 Å and width 0.5 Å, the limit in Theorem 1 becomes 1.5×10^{-8} . Therefore, the theorem implies that as ξ approaches 1 (i.e., at small values of RV), the proportion of the difference, $g(x; \xi)$, that remains to be modeled after using only ψ_1 with the same width and center as the original Gaussian is nearly zero. In other words, Doppler shifting a Gaussian absorption feature at a small RV is approximately the same as adding a constant multiple of ψ_1 (which is a scalar multiple of the Gaussian's derivative) to the feature. This is a special case of a well-known result: if a given function evaluates to zero at a given point, then the derivative is the leading term in the Taylor expansion about that point.

Some of the RV detection algorithms, such as the template matching method described in Bouchy, Pepe and Queloz (2001), attempt to model a Doppler shift by approximating the derivative of absorption features with a high S/N template spectrum. They then use a wavelength multiple of this derivative to create a nonlinear model of a Doppler shift with parameters to be fitted. At high wavelength values though, rescaling across a narrow wavelength window is nearly the same as an additive shift. In fact, if the Doppler shift were additive, the limit in Theorem 1 would be 0. Furthermore, an additive shift removes the nonlinearity in the Doppler shift model. While this idea is not new (Butler et al. (1996)), the approximation error of this has remained unknown. Therefore, Theorem 1 takes account of the wavelength rescaling nature of the Doppler shift, giving the value of this approximation error for assuming the shift to be additive at the limit of low values of RV.

To answer the question of how small an RV is small enough for this to be valid, we first state some Lemmas that solve for the coefficients in the decomposition shown in equation (8) with $g(x; \xi)$, as defined in Theorem 1. Lemma 1 gives a useful recursive relationship of an integral quantity that arises in solving the coefficients.

LEMMA 1. For $I_k(a, b, c) := \int_{-\infty}^{\infty} u^k e^{-(au^2+bu+c)} du$ where $a > 0$, we have that

$$(10) \quad I_0(a, b, c) = \sqrt{\frac{\pi}{a}} e^{\frac{b^2}{4a}-c},$$

$$(11) \quad I_1(a, b, c) = -\frac{\sqrt{\pi}b}{2a^{3/2}}e^{\left(\frac{b^2}{4a}-c\right)}$$

and, for all $k \geq 2$,

$$(12) \quad I_k(a, b, c) = -\frac{b}{2a}I_{k-1}(a, b, c) + \frac{k-1}{2a}I_{k-2}(a, b, c).$$

Using $I_k(a, b, c)$, as defined in Lemma 1, Lemma 2 gives the mathematical solution for the coefficients.

LEMMA 2. For $g(x; \xi)$, as defined in Theorem 1, decomposed as $g(x; \xi) = \sum_{n=0}^{\infty} c_n(\xi)\psi_n(x; \mu, \sigma)$, and $I_k(a, b, c)$, as defined in Lemma 1, we have that, for $\varepsilon = \xi - 1$,

$$(13) \quad c_0(\varepsilon) = \sqrt{\sigma\sqrt{\pi}} - \frac{1}{\sqrt{\sigma\sqrt{\pi}}}I_0\left(\frac{1 + \varepsilon + \frac{\varepsilon^2}{2}}{\sigma^2}, -\frac{2\mu + \varepsilon\mu}{\sigma^2}, \left(\frac{\mu}{\sigma}\right)^2\right)$$

and, for all $k \geq 1$,

$$(14) \quad c_k(\varepsilon) = -\sqrt{\frac{\sigma k! 2^k}{\sqrt{\pi}}} \sum_{m=0}^{\lfloor \frac{k}{2} \rfloor} \frac{(-1)^m}{4^m m! (k-2m)!} I_{k-2m}\left(1 + \varepsilon + \frac{\varepsilon^2}{2}, \frac{\varepsilon\mu}{\sigma}(1 + \varepsilon), \frac{1}{2}\left(\frac{\varepsilon\mu}{\sigma}\right)^2\right).$$

Using Lemmas 1 and 2, we numerically calculate the first seven coefficients as a function of RV and illustrate the results in Figure 6. It is not hard to notice that all the coefficients go to 0 as the RV goes to 0. This is because with no RV, $g(x; \xi)$, as defined in Theorem 1, is the zero-function. More importantly though, Figure 6 illustrates that, as the RV approaches zero, the dominating coefficient is c_1 .

When v_r has a magnitude below 100 m s^{-1} , it appears that all other coefficients besides c_1 are negligible, with c_0 and c_2 being the only possible exceptions. Furthermore, at velocities with a magnitude below 500 m s^{-1} , c_1 is approximately linear as a function of v_r . Since Figure 1 illustrates that a considerable number of currently known exoplanets exert a RV on their host star with a semi-amplitude less than 100 m s^{-1} , which is especially true for

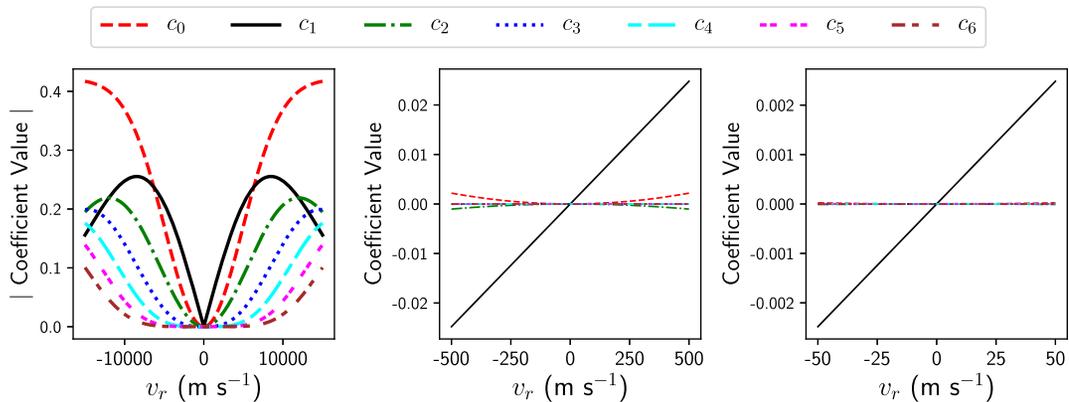


FIG. 6. The coefficient solutions that result from modeling a Doppler-shifted Gaussian with the Hermite–Gaussian basis are plotted here as a function of v_r . The left panel has the absolute value of the coefficients on the vertical axis and illustrates that, at low values of v_r , c_1 is the dominating coefficient. The middle and right panels show the exact coefficient value and illustrate that, at low values of v_r , c_1 is nearly a constant multiple of it. Only the zero'th up to the sixth coefficients are shown. The Gaussian here has the parameters of $\mu = 5000$ and $\sigma = 0.1$ which is meant to represent a typical absorption feature in a stellar spectrum.

Earth-like exoplanets, it suggests that it is not unreasonable to ignore all Hermite–Gaussian coefficients besides c_1 in modeling a Gaussian absorption feature that is Doppler-shifted due to an exoplanet.

Now that we have the coefficient solutions and have a sense that c_1 is the most dominant coefficient at values of RV that are of interest, we calculate the approximation error made by ignoring all other coefficients. To do so, we introduce a new quantity that we refer to as the standardized approximation error which appears in Theorem 1. For a function φ approximated by the function ϕ , define the standardized approximation error $D(\phi \parallel \varphi)$ as

$$(15) \quad D(\phi \parallel \varphi) = \frac{\int_{-\infty}^{\infty} (\varphi(x) - \phi(x))^2 dx}{\int_{-\infty}^{\infty} \varphi(x)^2 dx}.$$

In a sense, $D(\phi \parallel \varphi)$ gives the proportion of the squared function φ that remains to be modeled after approximating with ϕ . In our case we consider $D(g(x; \xi) \parallel c_1(\xi)\psi_1(x; \mu, \sigma))$.² Lemmas 3 and 4 help us solve for the limit as ξ approaches 1 (i.e., as v_r approaches 0).

LEMMA 3. For $g(x; \xi)$, as defined in Theorem 1, decomposed as $g(x; \xi) = \sum_{n=0}^{\infty} c_n(\xi)\psi_n(x; \mu, \sigma)$, we have that

$$(16) \quad D(g(x; \xi) \parallel c_1(\xi)\psi_1(x; \mu, \sigma)) = 1 - \frac{c_1^2(\xi)}{\int_{-\infty}^{\infty} (g(x; \xi))^2 dx}.$$

$$\text{LEMMA 4.} \quad \lim_{\xi \rightarrow 1} \frac{c_1^2(\xi)}{\int_{-\infty}^{\infty} (g(x; \xi))^2 dx} = \frac{1}{1 + \frac{3\sigma^2}{2\mu^2}}.$$

Combining Lemmas 3 and 4 completes the proof of Theorem 1. (See Section 2 of the Supplementary Material (Holzer et al. (2021)) for a more detailed proof of each.)

Theorem 1 does not explicitly give a rate at which the standardized approximation error approaches its limit. But by using Lemma 3 and equation 35 from the proof of Lemma 4 in Section 2 of the Supplementary Material (Holzer et al. (2021)), we illustrate the rate with Figure 7. Note that the standardized approximation error shown here is bounded between 0 and 1 and that the limit is actually nonzero. Figure 7 illustrates that as $\xi \rightarrow 1$, $D(g(x; \xi) \parallel c_1(\xi)\psi_1(x; \mu, \sigma))$ approaches its limit quadratically and that when $v_r < 50 \text{ m s}^{-1}$, the standardized approximation error is less than 2.5×10^{-5} away from the limiting value.

3.2. RV estimation method. Theorem 1 suggests a natural new method for detecting a Doppler shift in the spectrum of a star. As long as the magnitude of v_r is small enough, the absorption feature is approximately Gaussian and the ratio μ/σ for the feature is large enough, we can do a least-squares fitting of the first-degree Hermite–Gaussian function to the difference between a template spectrum and a Doppler-shifted spectrum and map the fitted coefficient to a RV. As illustrated in Figure 6, c_1 at low values of v_r is directly proportional to v_r .

According to Lemma 2, $c_1(\varepsilon) = \frac{\sqrt{\sqrt{\pi}}}{\sqrt{2\sigma}} \varepsilon \mu (1 + \varepsilon) \tilde{h}(\varepsilon)$, and $\lim_{\varepsilon \rightarrow 0} \frac{\partial}{\partial \varepsilon} c_1(\varepsilon) = \frac{\mu \sqrt{\sqrt{\pi}}}{\sqrt{2\sigma}}$.

Furthermore, using equation (2) with $\varepsilon = \xi - 1$, we have that the mapping from ε to RV is $v_r(\varepsilon) = c\varepsilon$ and $\lim_{\varepsilon \rightarrow 0} \frac{\partial}{\partial \varepsilon} v_r(\varepsilon) = c$. Hence, $\lim_{\varepsilon \rightarrow 0} \frac{\partial}{\partial v_r} c_1(v_r(\varepsilon)) = \frac{\mu \sqrt{\sqrt{\pi}}}{c\sqrt{2\sigma}}$ which is the

²Since $g(x; \xi)$ approaches the zero function as $\xi \rightarrow 1$ and for any $k \geq 0$ $c_k(\xi) \rightarrow 0$ as $\xi \rightarrow 1$, the ordinary approximation error of using any individual k would approach 0. This would tell us nothing about the relative magnitudes of the Hermite–Gaussian coefficients. The denominator of $D(g(x; \xi) \parallel c_1(\xi)\psi_1(x; \mu, \sigma))$ adjusts for this by standardizing the quantity.

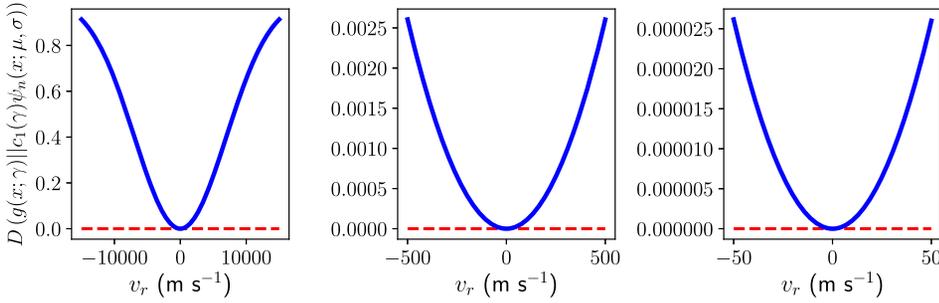


FIG. 7. The standardized approximation error $D(g(x; \xi) \parallel c_1(\xi)\psi_n(x; \mu, \sigma))$ in Theorem 1 as a function of v_r with parameters $\mu = 5000$ and $\sigma = 0.1$ is plotted in bold. The limit is also shown in the horizontal red dashed line.

desired proportionality constant. So the proportionality that is valid at low values of RV, v_r , is

$$(17) \quad c_1 = \frac{\mu\sqrt{\sqrt{\pi}}}{c\sqrt{2\sigma}}v_r.$$

The strongest assumption made when applying the theorem is that the absorption features are Gaussian shaped. Because this is never exactly true, we analyze this model misspecification further in Section 3.4 below.

3.3. Extension to multiple features. Since a single absorption feature is unable to give a RV estimate that is precise enough, we need to use as many features in the spectrum as possible. Instead of fitting only a single first-degree Hermite–Gaussian function to the difference spectrum, we fit a sum of these functions to it. To construct this sum, we note that it must take into account the fact that differing absorption features will have different centers, widths, and depths. The generalized Hermite–Gaussian functions in equation (7) can take account of the different centers and widths. Furthermore, according to equation 12 in the proof of Lemma 2 (provided in Section 2 of the Supplementary Material (Holzer et al. (2021))), Doppler-shifting a Gaussian with any amplitude simply multiplies the resulting coefficients by the same amplitude. In the case of stellar spectra, this amplitude is simply the line depth. Therefore, using equation (17), the resulting model of the difference flux at pixel i , y_i , as a function of wavelength, x_i , to be fitted becomes

$$(18) \quad y_i = v_r \sum_{j=1}^n \frac{\sqrt{\sqrt{\pi}}d_j\mu_j}{c\sqrt{2\sigma_j}}\psi_1(x_i; \mu_j, \sigma_j) + \varepsilon_i,$$

where the sum is over all n absorption features, d_j represents the line depth of the j th feature and each ε_i is independent with expectation 0.

In practice, we assume that $\varepsilon_i \sim N(0, \varrho_i^2)$ and is independent for each i . Many modern stellar spectra come with uncertainties for each pixel’s normalized flux.³ This is particularly true for the normalized spectra from EXPRES that we analyze here. EXPRES estimates the uncertainty in each pixel by assuming the unnormalized flux is Poisson, estimating the red noise, and accounting for intrinsic effects of flat-fielding (Petersburg et al. (2020)). Therefore,

³If these uncertainties are not provided, weights can be defined using the standard assumption that the raw flux is Poisson. That is, the weights can be set to $w_i = \frac{\text{cont}_i}{\hat{\tau}_i}$, where cont_i is the value of the raw continuum used for normalization at pixel i and $\hat{\tau}_i$ is the value of the estimated template.

we assume that the provided uncertainties, \hat{q}_i , are accurate estimates of each q_i and estimate v_r in equation (18) through weighted least squares with weights $w_i = 1/\hat{q}_i^2$.

To calculate the difference flux, y_i , at pixel i we need a template spectrum. Here, we use the estimated template calculated from the set of observed spectra (see Section 3.5 for more details).

Since equation (17) approximately holds for $v_r < 500 \text{ m s}^{-1}$, which well encompasses most exoplanets of interest, we have a new Hermite–Gaussian based radial velocity (HGRV) estimation method. For a spectrum of Gaussian absorption features, we can create a linear model of the difference spectrum due to a Doppler-shift as a function of the sum of ψ_1 functions as given by equation (18), the coefficient of which is the RV. Therefore, we have reduced the Doppler shift estimation problem to linear regression with no intercept. This method does not fundamentally require interpolation,⁴ treats neighboring pixels similarly, accounts for the heteroskedastic noise and easily allows for obtaining the standard error of the estimated RV for statistical inference.

3.4. Model misspecification. The HGRV method assumes that the shape of absorption features is Gaussian which does not hold exactly. Various reasons are understood to contribute to this: a line following the Voigt profile may have a nonnegligible Lorentzian component, the line may be deep enough to depart from the Voigt profile or there may be additional effects in the star’s atmosphere that are not well encompassed by current physical models.

Since the HGRV method assumes Gaussian shaped absorption features, we now investigate the effects of applying it to non-Gaussian shaped features. We consider the absorption feature in the NSO spectrum between 5243.7 and 5244.2 Å. This feature is shown in the left panel of Figure 8, along with its best-fit Gaussian. For 50 equally spaced values of RV from 1 to 100 m s^{-1} , we Doppler shift this feature according to equation (2), use cubic splines to interpolate back to the original wavelength solution (Mészáros and Prieto (2013)) and fit the difference flux with the HGRV model from equation (18) (with $n = 1$ and d , μ , and σ as the estimated parameters from the best-fit Gaussian). The ratio between the estimated and true RV is shown in the right panel of Figure 8.

Figure 8 illustrates that, for this particular absorption feature, the HGRV method slightly overestimates the RV. For example, if the true RV is 1 m s^{-1} , this bias would be approximately 0.5 cm s^{-1} . Similarly, for a true RV of 100 m s^{-1} the bias would be less than

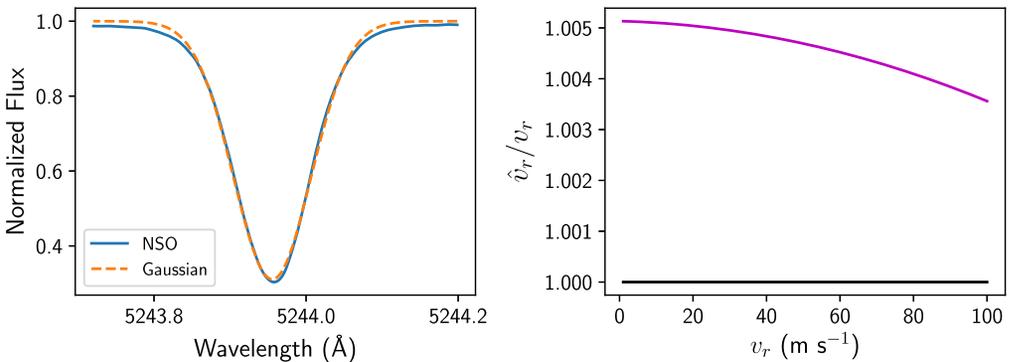


FIG. 8. Results for analyzing the effects of misspecifying the model of the absorption feature in the NSO spectrum between 5243.7 and 5244.2 Å as a Gaussian. The left panel shows the feature in solid blue and the best-fit Gaussian in dashed orange. The right panel shows the ratio of the RV estimated with equation (18) \hat{v}_r (with $n = 1$) and the true RV, v_r .

⁴Interpolation is, however, used later on a high S/N, oversampled estimate of the template spectrum to give it the same wavelength solution as each observed spectrum so that the difference flux can be calculated.

0.4 m s^{-1} . These results are consistent across other absorption features considered. For additional discussion about applying the same analysis to other NSO absorption features, see Section 3 of the Supplementary Material (Holzer et al. (2021)).

3.5. Nonparametric template estimation. Since the HGRV method models the difference in normalized flux, we need to have a template spectrum that approximates the quiet spectrum of a star with no stellar activity. In principal, if one knows the approximate effective temperature, surface gravitational acceleration, metallicity, microturbulent velocity and the elemental abundances of the star with high precision, a synthetic spectrum could be produced at the proper resolution to give such a template (Snedden et al. (2012)). However, in practice these stellar parameters and the atomic line transition data are not known well enough to make this feasible. Therefore, we take a data-driven approach.

The method we propose for estimating the template is to combine all normalized, barycentric corrected, observed spectra across time epochs and fit a smooth curve to estimate a representative spectrum. The time sampling of the spectra can affect how well the estimated template approximates the true template. For example, two of the possible extremes in the sampling are if all the observations are at the same orbital phase or if the observations are uniform across all phases. The estimated template under these extremes are not likely to affect the end result of the HGRV approach, so this template estimation method is sufficient for our purposes.⁵

All observed spectra are combined together, and we fit a local regression curve to this combined spectrum with a Gaussian kernel. We use local quadratic, instead of local linear, regression in order to better model the cores of absorption features. In practice we only fit at most *eight* Å of the combined spectrum at a time, choosing an optimal bandwidth through generalized cross-validation for each section. This allows the computation to be parallelized. It also allows the bandwidth to be locally adaptive and take account of how absorption features are narrower on the blue end of the spectrum compared to the red end. An advantage of this approach is that, when combining all observed spectra, the wavelength solutions do not need to match across epochs, further minimizing the role of interpolation.

4. Simulation studies. This section includes two simulation studies based on the proposed methodology. The first is related to the template estimation approach, and the second compares properties of the RV estimation using the HGRV method with those of the commonly used CCF method.

4.1. Template estimation. A nice feature of the HGRV approach is that no prespecified template is required because the template spectrum is estimated from the full time-series of spectra using local quadratic regression (see Section 3.5). The estimated template contains both bias and variance, and we investigate the overall root mean squared error (RMS) through simulation. Furthermore, we consider how the RMS changes with the number of spectra and the S/N. Finally, we explore how the time-sampling cadence affects the estimated template.

⁵Using this template estimation approach with time sampling that is approximately uniform across all phases of an exoplanet's orbit may lead to slightly broader features in the estimated template. However, broadening tends to be primarily an even effect and so would not significantly hinder the RV estimation using the HGRV method which fits an odd function (ψ_1) to the difference flux in equation (18). Time sampling carried out in such a way that the observations occur at approximately the same phase of an exoplanet's orbit should not have this broadening of features. However, a constant RV offset may be present between the estimated template spectrum and all observed spectra. Because the same estimated template is used for each observation and only relative RV estimates are needed, this offset should not influence the fitted orbital parameters.

For a star's true template with normalized flux τ and estimated template with normalized flux $\hat{\tau}$, we define the RMS as

$$(19) \quad \text{RMS}(\hat{\tau}) = \sqrt{\frac{1}{n} \sum_{i=1}^n (\tau_i - \hat{\tau}_i)^2}.$$

For our simulation we use a version of the NSO spectrum that we smooth through local quadratic regression that approximately represents the quiet solar spectrum with infinite S/N. We also use cubic spline interpolation to give this smoothed NSO spectrum the same wavelength solution as the 51 Pegasi spectrum observed by EXPRES on Julian Day (JD) 2,458,641.952. For a given number of observed spectra, N , each with a given S/N, our simulation consists of the following steps: (i) sample time epochs t_1, \dots, t_N where $t_k \sim \text{iid Uniform}(0, 2\pi)$, (ii) calculate RV's $v_{r,1}, \dots, v_{r,N}$ where $v_{r,k} = 10 \sin(t_k)$, (iii) simulate N observed spectra with wavelength axis Doppler-shifted using equation (2) with RV $v_{r,k}$, and normalized flux axis with independent Poisson noise at the given S/N (where the noise is added to the un-normalized flux) and (iv) apply the template estimation method described in Section 3.5, and calculate the resulting $\text{RMS}(\hat{\tau})$.

In our simulations the number of spectra, N , ranges from *one* to 31 (in steps of 2) and the S/N ranges from 100 to 250 (in steps of 10). For each pair of values, we perform the simulation 50 independent times and calculate the average and standard deviation of the RMS. Each of these 50 represents a different cadence. For computational purposes we do not use the entire spectrum for this simulation. Instead, we use the wavelength window 5240–5245 Å for our simulation. We also ran the same simulation on the wavelength window 4965–4970 Å, which has a higher density of absorption features, as well as the window 6381–6386 Å which has a lower absorption feature density. The results for these additional windows are similar to the first window. The results for the window 5240–5245 Å are summarized in Figure 9, which shows the average $\text{RMS}(\hat{\tau})$ on the left panel and the standard deviation of the $\text{RMS}(\hat{\tau})$ on the right, for each pair of S/N and number of spectra.⁶

The left plot in Figure 9 illustrates that, once the number of spectra reaches approximately 21, the average $\text{RMS}(\hat{\tau})$ of the estimated template is below approximately 0.001 (which represents a S/N of about 1000) for any S/N above 100. On the other hand, if all observed spectra had a S/N above 200 (which is often true of EXPRES spectra), one would only need about 11 spectra to reach this template estimation precision. Furthermore, by examining the differences between the true template and individual instances of an estimated template, the residuals showed no obvious systematic bias within the wavelength bounds of absorption features. The same plot also shows that the $\text{RMS}(\hat{\tau})$ is more affected by the number of spectra than the S/N in this example. Figure 10 shows how the RMS indicated on the colorbar of the left plot in Figure 9 maps to an effective S/N of the estimated template.

The right plot in Figure 9 illustrates how the $\text{RMS}(\hat{\tau})$ varies due to the differing cadences in the 50 samples used for each pair of S/N and number of spectra. The simulation suggests that, as expected, the greatest differences are found when using only one spectrum. The variation is minimal for 11 or more spectra and a S/N above 150.

We next investigate how the HGRV-estimated RV changes when using an estimated template instead of the true template. We first note that there are two main sources of a difference in the estimated RV due to template estimation for this simulation study design: (i) finite S/N

⁶The results of this simulation did not change significantly when calculating the RMS using only the portion of the spectrum that lies inside the wavelength intervals corresponding to absorption features as given by the AFF algorithm. For instance, with a S/N of 200 and 20 observed spectra, the original RMS was 0.00078 ± 0.00003 , whereas the RMS on only the portion of the spectrum pertaining to absorption features was 0.00079 ± 0.00003 .

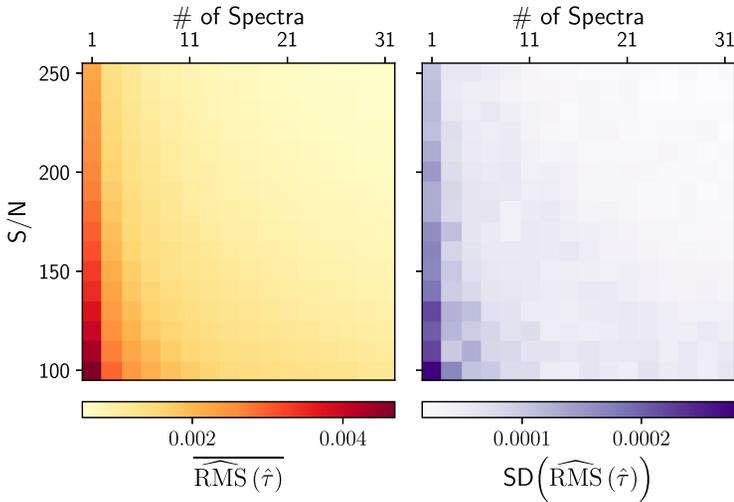


FIG. 9. Simulation study results for estimating the template spectrum between 5240 and 5245 Å. For each S/N and number of spectra, N , 50 simulations were carried out each with a different cadence. Each simulation involved estimating the template with local quadratic regression and calculating the RMS. The left plot shows the average, and the right plot shows the standard deviation, of the RMS across the 50 simulations for each pair of S/N and N . The plots share the same vertical-axis.

of the estimated template and (ii) shape changes of absorption features due to the nonparametric smoothing and the changing cadence. The second source is what we investigate here. Since each of the observed spectra has a different RV, they all will be slightly shifted with respect to the true template and each other. Depending on the cadence, this could lead the estimated template to have some shift relative to the true template and have slightly broader absorption features. Furthermore, the bias of nonparametric smoothing may also broaden the shape of absorption features. Since the same estimated template is used for each observation, a shift with respect to the true template would only result in a constant offset of all RV's. Since this constant offset can be accounted for and removed, it is not problematic.

To see how such shape changes might influence the variance of the estimated RV, we apply a procedure similar to the one used to produce Figure 9. With the number of observed spectra set to 20 and a S/N of 80,000 (essentially infinite), we simulate 50 sets of observations, each with a randomly sampled cadence. The final result of this is a set of 50 different

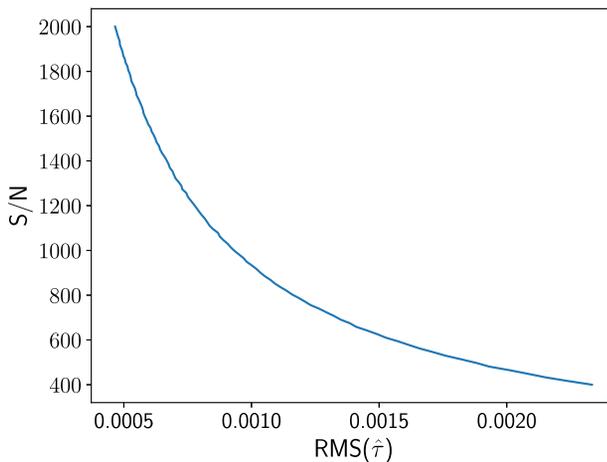


FIG. 10. The estimated effective S/N of a spectrum as a function of the RMS as defined in equation (19).

estimated templates with the dominant difference from the true template being the possible shape changes described above.

For each of these 50 estimated templates we use the true template to generate an observed spectrum with S/N of 300 and a 10 m s^{-1} Doppler shift. To understand how the shape changes affect the variance of the RV estimate, we calculate the difference flux using the true template and this observed spectrum. The HGRV method then returns an estimated RV for each of these 50. Using the same set of 50 observed spectra, this process is then repeated with the estimated templates in place of the true template. Each observed spectrum then has two RV estimates, one from the true template and one from an estimated template. We find that the standard deviation of the RV estimates that used the true template differs only by approximately 0.001 m s^{-1} from the standard deviation of the RV estimates that used the estimated templates. This suggests that the effect of shape changes on the variance of the estimated RV due to nonparametric smoothing and differing cadences is negligible. Furthermore, we find that the estimated relative RV differs in magnitude by approximately 0.006 m s^{-1} on average when using the estimated templates instead of the true template.

4.2. RV estimation. To investigate the accuracy of the HGRV method, especially at low velocities, we simulate spectra with a known RV and estimate the RMS of \hat{v}_r . By design, this simulation ignores astrophysical effects on RV-precision from stellar activity, analyzing the error contribution from modeling alone. To estimate this RMS, we use

$$(20) \quad \widehat{\text{RMS}}(\hat{v}_r) = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{v}_{r,i} - v_r)^2},$$

where n is the number of simulations at RV v_r . The square of $\widehat{\text{RMS}}(\hat{v}_r)$ can be decomposed into the sum of the variance and squared bias of \hat{v}_r as well. To get a more detailed summary of our simulation, we also estimate the standard deviation (SD) with

$$(21) \quad \widehat{\text{SD}}(\hat{v}_r) = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{v}_{r,i} - \bar{v}_r)^2},$$

where \bar{v}_r is the average estimated velocity, and estimate the bias with

$$(22) \quad \widehat{\text{Bias}}(\hat{v}_r) = \bar{v}_r - v_r.$$

We explore how the $\text{RMS}(\hat{v}_r)$, $\text{Bias}(\hat{v}_r)$ and $\text{SD}(\hat{v}_r)$ vary with S/N and v_r . Our simulation takes *five* equally spaced values of S/N 100, 150, ..., 300 and *four* values of v_r equally spaced on a log scale from 0.01 to 100 m/s. For each pair of S/N and v_r values, we use the estimated template spectrum for 51 Pegasi to simulate 2000 independent spectra with the proper Doppler shift given by equation (2). Each such simulation consists of using cubic splines to interpolate the shifted, oversampled and high S/N template to the same wavelength solution as the observed 51 Pegasi spectrum from EXPRES on JD 2,458,639.958 (see Section 5 for more details) and including Poisson noise of the specified S/N.

To approximately account for the uncertainty of the estimated template in this simulation, we use the results from Section 4.1. Figure 9 illustrates that the RMS (as defined in equation (19)) for a given S/N does not change significantly when increasing the number of spectra beyond about 21. Therefore, we take the column of the left plot in Figure 9 pertaining to 21 spectra and the results shown in Figure 10 to map the S/N of observed spectra to an effective S/N of the estimated template spectrum. For each of the 2000 independent spectra in this simulation and for each pair of S/N and v_r , we add noise to the true template with the appropriately mapped effective S/N to approximate the effect of using the estimated template.

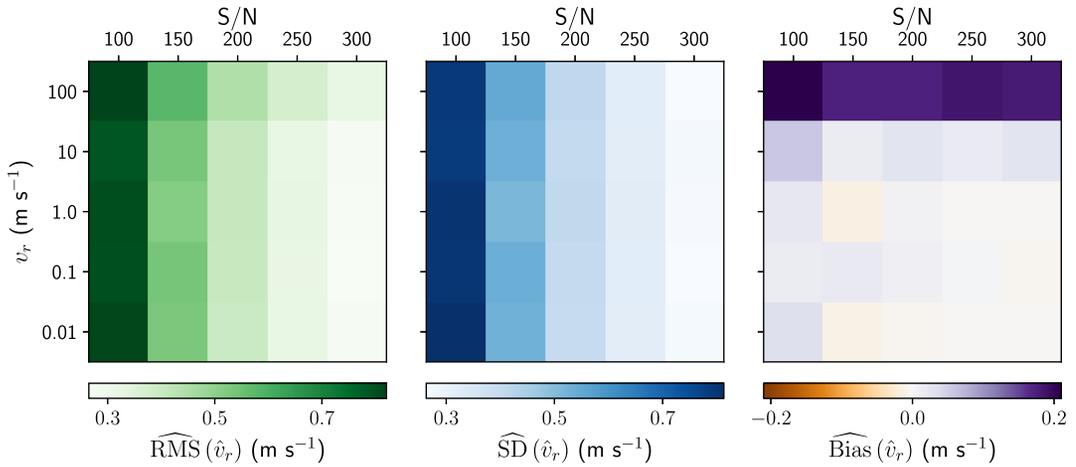


FIG. 11. *The results for applying the HGRV method to spectra simulated from the estimated 51 Pegasi template spectrum. The left, middle and right panels show the estimated RMS, SD and bias of the estimated RV, respectively. All three panels share the same vertical axis that represents the true RV with which each spectrum was simulated. The S/N of the simulated spectra are given by the horizontal axis on top of each panel. The color scale for each panel is represented by the colorbar below it. Each pair of S/N and v_r involved 2000 independent simulations to estimate the three quantities.*

The results for obtaining each \hat{v}_r with the HGRV method are shown in Figure 11. The left panel of Figure 11 illustrates that the HGRV method is able to obtain a precision less than 0.3 m s^{-1} when the S/N is approximately 250 or higher, at least in the small RV regime. Additionally, the right panel of Figure 11 builds upon the model misspecification simulation done in Section 3.4 and informs us that combining many (non-Gaussian) absorption features in the HGRV method does not lead to an amplified systematic bias. We also find that the bias is somewhat proportional to the true RV. Furthermore, the SD contributes significantly more to the overall RMS than whatever bias may be present at the RV and S/N considered here.⁷

We also run the same simulation, estimating the RV with the CCF method, as used in the EXPRES pipeline (Petersburg et al. (2020)), with the commonly used HARPS G2 mask. Such a mask is a collection of wavelengths and weights (based on line depths) for many of the absorption features in a spectrum; the mask is designed to align with the absorption features of a star at rest (i.e., a star with zero RV.) Since the CCF method returns an absolute RV, rather than a relative RV, we first calculate the RV given for the estimated 51 Pegasi template with no noise ($-33,168.5399 \text{ m s}^{-1}$) and subtract this offset from all estimated RV's from the simulation. We then compare the estimated bias, SD and RMS of the two methods at each pair of S/N and v_r . Figure 12 shows the difference in RMS between the HGRV and CCF methods. Since every pair of S/N and v_r in Figure 12 shows a negative RMS difference, this suggests that the HGRV method has higher RV-precision than the CCF approach in this regime.

As a more detailed summary of the RMS improvement of the HGRV, as demonstrated by Figure 12, the difference in the estimated SD and absolute bias (the sum of squares of which equal the squared RMS) is shown in Figure 13.

⁷We also performed the same simulation with a S/N of 1000 and a RV of 1 m s^{-1} (again using the estimated 51 Pegasi template spectrum and simulating 2000 independent spectra). This simulation gave an estimated RMS of 0.077 m s^{-1} , an estimated SD of 0.077 m s^{-1} and an estimated bias of $2.5 \times 10^{-3} \text{ m s}^{-1}$. This demonstrates that the HGRV method has the capability of obtaining a RV precision less than 0.1 m s^{-1} .

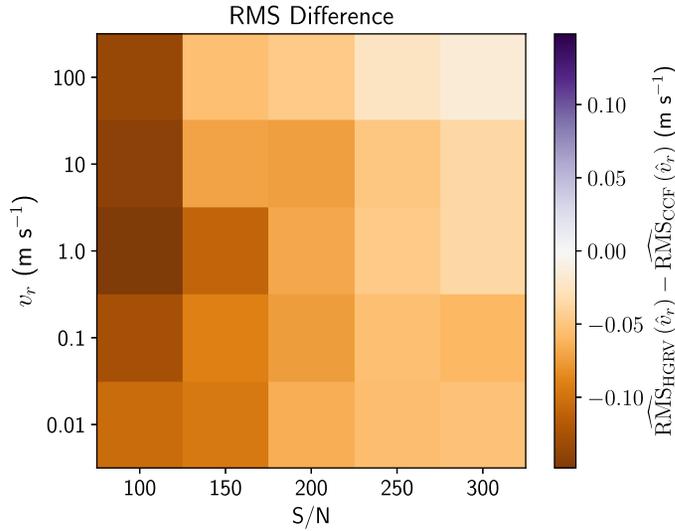


FIG. 12. The difference between the HGRV and CCF RMS for each pair of S/N and true v_r . Each pair consisted of 2000 independent simulations for each method. The difference is indicated on the right by the colorbar, which is centered at 0.0 m s^{-1} , and demonstrates the higher RV-precision of the HGRV method.

Figures 12 and 13 inform us that the HGRV method is an example of the statistical phenomenon where a small increase in bias reduces the overall RMS. The greatest difference in RMS between the HGRV and CCF methods appears to be at low S/N .

We suspect that the HGRV method's reduced RMS may be connected to the pixel sampling of the wavelength. In this simulation we found that, when using the wavelength sampling of the NSO spectrum, the RMS between the CCF and HGRV methods was nearly the same at most pairs of S/N and v_r . However, when we instead interpolated to the wavelength solution of a 51 Pegasi spectrum observed by EXPRES, the obvious difference in RMS, as demonstrated in this simulation study, emerged. This suggests that the HGRV method may be more robust against the pixel sampling that differs across spectrographs.

To check the stability of this simulation, we used the wavelength solution for the 51 Pegasi spectrum from EXPRES observed on JD 2,458,804.588 instead of the wavelength solution from JD 2,458,639.958 used above. Running the HGRV and CCF approach, each

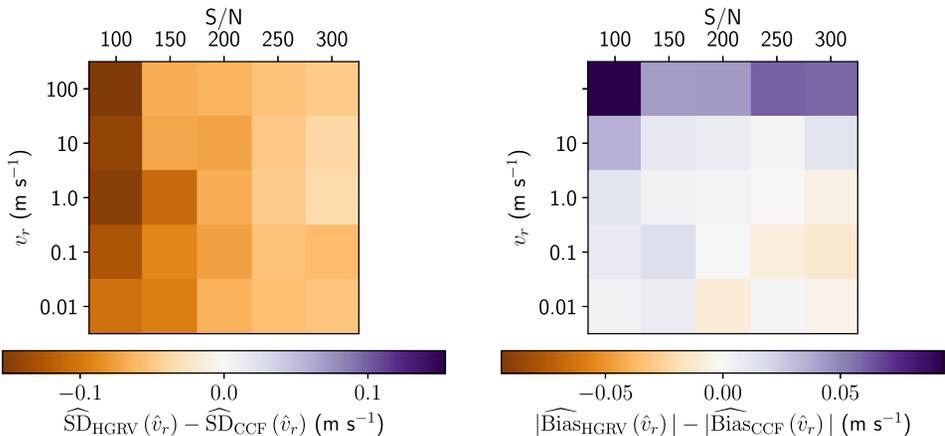


FIG. 13. The difference between the HGRV and CCF standard deviation and absolute bias for each pair of S/N and true v_r . Each pair consisted of 2000 independent simulations for each method. The differences are indicated below each panel by the colorbars which are centered at 0.0 m s^{-1} .

with 2000 independent simulations, with $v_r = 1 \text{ m s}^{-1}$ and a S/N of 200 produced an RMS difference of -0.088 m s^{-1} (previously -0.067 m s^{-1} , as shown in Figure 12. All estimated RVs from the CCF and HGRV methods for these simulations are provided in the source code of the Supplementary Material (Holzer (2021)) and in the repository https://github.com/parkerholzer/hgrv_method.

5. Applications to 51 Pegasi data. 51 Pegasi is the first main-sequence star similar to the Sun discovered to possess an exoplanet (Mayor and Queloz (1995)). The exoplanet has been found to have a RV semiamplitude of $55.57 \pm 2.22 \text{ m s}^{-1}$ and orbital period of 4.2292 ± 0.0003 days (Marcy et al. (1997), Mayor and Queloz (1995), Wang and Ford (2011), Bedell et al. (2019)). To test the proposed HGRV method, we use data recently collected for 51 Pegasi by EXPRES (Jurgenson et al. (2016), Petersburg et al. (2020)). The recent spectrograph of EXPRES corrects for many of the instrumental effects that prior observations of 51 Pegasi were unable to avoid, allowing for greater precision of derived RV. Our dataset consists of 56 observed spectra from JD 2,458,639 to 2,458,805 (June 5, 2019 to November 18, 2019). The S/N of these spectra ranges from 89 to as high as 385, but most are close to 200 (see Table 1 in Section 4 of the Supplementary Material (Holzer et al. (2021)) for more details). These spectra have wavelength solutions that differ and do not consist of equally spaced pixels.

5.1. Data corrections. The raw data collected by the spectrograph do not have a flat continuum. This is, in part, due to the star's temperature causing more photons to be emitted at certain wavelengths than others. It is also due to instrumental effects, such as the theoretical blaze function (Barker (1984), Xu et al. (2019)). To correct for these effects, we adopt the normalization from the EXPRES pipeline provided with each spectrum (Petersburg et al. (2020)).

We also correct for the effects of the Earth's motion around the Sun by adopting the barycentric corrected wavelength solution provided with each observed spectrum by the EXPRES pipeline (Blackman et al. (2017), Blackman et al. (2020), Petersburg et al. (2020)). Without the barycentric wavelengths provided by the EXPRES team, our derivation of RV would incur errors at the level of tens of cm s^{-1} .

Finally, we correct for absorption features due to the Earth's atmosphere, often referred to as tellurics. Since the spectrograph is ground-based, the light from the star passes through the Earth's atmosphere, causing the presence of additional absorption features in the spectrum that are not representative of the target star. To correct for these tellurics, we use the model provided by the EXPRES team with each spectrum that was created using the approach of Leet, Fischer and Valenti (2019). Although one could potentially divide out shallow tellurics to approximately correct for them with such a model, we take a more conservative approach and mask out all pixels with a telluric model normalized flux less than 1.0.

While this telluric correction approach may not remove all tellurics, the HGRV method has two advantages that minimize the effect tellurics have on the estimated RV. First, because the tellurics occur at the same wavelength over time in the raw spectra and, therefore, occur at different wavelengths in the barycentric corrected spectra, the nonparametric smoothing used to estimate the template averages out any tellurics missed by the telluric model. Second, any tellurics that persist after the smoothing in the estimated template and are detected by the AFF algorithm will have a very small depth. Therefore, they will have a relatively low leverage in the HGRV regression, minimally affecting the estimated coefficient.

Because a spectrum covers over 3000 \AA of wavelength, the spectrograph collects the data in (partially overlapping) wavelength orders stacked onto the rectangular detector. Therefore, we begin by stitching all orders of a given epoch together to create a single array of wavelength and normalized flux. To stitch two neighboring orders together in their overlapping

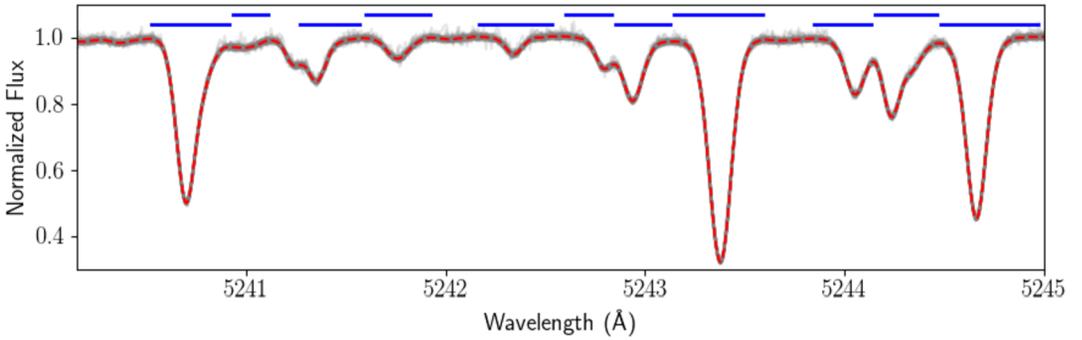


FIG. 14. A subset of the estimated template spectrum calculated from 51 Pegasi data is shown in the red dashed line on top of all observed spectra used in the calculation (shown in gray). The feature bounds that result from running the AFF algorithm on the estimated template spectrum are also shown in blue horizontal lines. The full spectrum goes from 4470–6800 Å, but for visualization only 5240–5245 Å are displayed. The error bars of the estimated template between 4850 and 6800 Å (i.e., the wavelengths used in the analysis) have a median of 5.2×10^{-4} and a 99th percentile of 1.1×10^{-3} .

region, we use cubic-spline interpolation to give the same wavelength solution to both orders in the overlap region (Mészáros and Prieto (2013)). We then take the (pointwise) weighted average of the normalized flux in the overlap region of the two orders. Since the signal decreases at the edge of each order due to the instrumental blaze function, we set the weights for this averaging to decrease linearly for a given order as we get closer to the edge of the order. After applying this stitching to all neighboring orders, we have a full observed spectrum for each epoch.

We then proceed to estimate the template spectrum by way of local quadratic regression, as described in Section 3.5. A small wavelength window of the estimated template spectrum that is calculated from the 51 Pegasi data is shown in Figure 14.

Once we have the high S/N estimated template spectrum, we can use it in the AFF algorithm to find absorption feature wavelength bounds. The tuning parameters of the algorithm that were found through the optimization process, described in Section 1 of the Supplementary Material (Holzer et al. (2021)), were $m = 7$, $\alpha = 0.05$, and $\eta = 0.07$ while eliminating any features with a line depth less than 0.015. The algorithm finds a total of 4190 features between wavelengths 4470 Å and 6800 Å. The results of this are also indicated in Figure 14 for the section of the spectrum displayed. Note that, when neighboring features are strongly blended together, the AFF algorithm may either count both as a single feature or only pick out one of the two.

5.2. Absorption feature parameters. In order to use equation (18) and estimate the RV, we need to get estimates of the Gaussian parameters d_i , μ_i and σ_i for each absorption feature i using the high S/N estimated template spectrum. To do so, we use the trust-region reflective algorithm (Branch, Coleman and Li (1999)) which allows for initialization and bounds for each parameter to be fitted in nonlinear least-squares. For absorption feature i we initialize the Gaussian amplitude d_i at one minus the minimum flux attained by the estimated template spectrum within the wavelength bounds of feature i , the Gaussian center μ_i is initialized at the wavelength for which this minimum flux is attained and the Gaussian spread σ_i is initialized at one-fifth the width of the wavelength window for feature i . The bounds on the Gaussian amplitude are set to be $[0, 1]$, the Gaussian center is restricted to be within the wavelength bounds for feature i and the Gaussian spread is lower-bounded by 0 and upper-bounded by the width of the wavelength window for feature i .

For computational purposes we do not optimize the Gaussian parameters for all absorption features simultaneously. Instead, we estimate the parameters of one absorption feature by

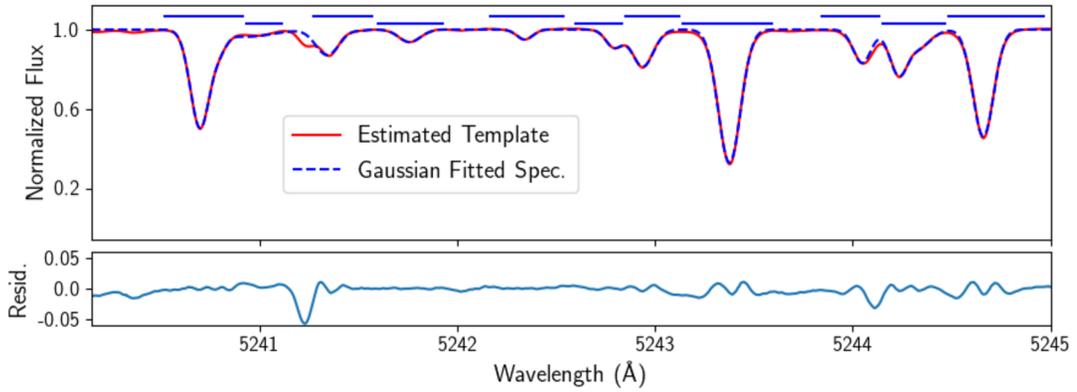


FIG. 15. The estimated template spectrum for 51 Pegasi is shown in solid red with the spectrum that approximates it as a sum of Gaussians shown in dashed blue. The full spectra go from 4470–6800 Å, but for visualization only 5240–5245 Å are displayed. All absorption features in this wavelength range were well fitted with Gaussians within the feature wavelength bounds. Portions of the spectrum that are poorly fitted with the sum of Gaussians are not contained within wavelength bounds of detected features, indicated with horizontal blue solid lines. The residual difference is shown below the main plot with the same “Wavelength” axis and a magnified vertical axis.

simultaneously optimizing that feature with its two neighboring features. If the resulting fit has a MSE within the wavelength bounds of the feature that is high,⁸ which particularly happens when two strongly blended spectral lines are counted as one absorption feature, we try fitting a sum of two Gaussians to it. If this still does not give a good fit, we eliminate the respective feature so as to minimize the effects of model misspecification analyzed in Section 3.4. Out of the 4174 absorption features detected by the AFF algorithm, 3868 were well fitted with one or two Gaussians. An example of the fit model spectrum is shown in Figure 15. Most of the features that were eliminated at this stage were strongly blended with one or more neighboring features.

5.3. *Results.* To derive the RV for each epoch, we first limit the spectrum to the wavelength region 4850–6800 Å. While the wavelength solution is excellent from 5000 to 7000 Å, due to the laser frequency comb of EXPRES spanning that region (Blackman et al. (2020), Petersburg et al. (2020)) and increasingly poor outside that window, we find that the spectra are acceptable for our purposes down to about 4850 Å.⁹ Below 4850 Å the noise of the spectra increases, and wavelengths above 6800 Å have too many strong telluric features. Limiting to this wavelength region reduces the number of absorption features from 3868 to 2796. Furthermore, we eliminate any pixels in the spectrum that are not contained in the wavelength windows of these 2796 features.

After using cubic-splines to interpolate the high S/N, oversampled, estimated template spectrum to the wavelength solution of the observed spectrum for a given epoch¹⁰ (Mészáros and Prieto (2013)), we calculate the difference spectrum between the two. We then compute a new variable, which can be thought of as a transformation of the wavelength, using the sum $\sum_{j=1}^n \frac{\sqrt{\pi} d_j \mu_j}{c \sqrt{2\sigma_j}} \psi_1(x_i; \mu_j, \sigma_j)$ from equation (18). This transformation uses all fitted Gaussian parameters, after which we model the difference flux across the full stitched spectrum as

⁸We consider a MSE to be high if it is greater than four multiples of the median MSE.

⁹The RV estimation error for the CCF and FM approaches increases when extending to wavelengths below 5000 Å. The HGRV approach does not appear to be negatively impacted when using wavelengths down to 4850 Å; the reason for this is not yet fully understood.

¹⁰This is the only time in the proposed method that interpolation takes place.

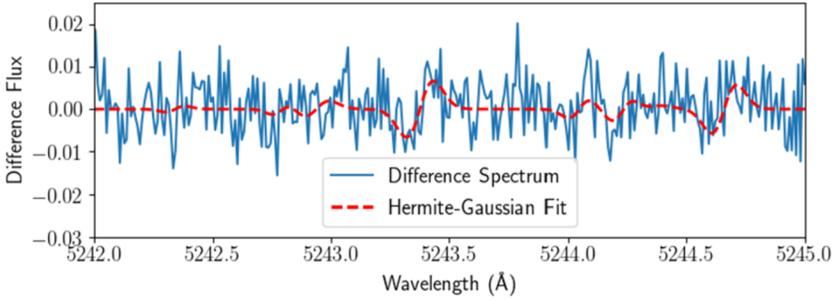


FIG. 16. The difference spectrum between the estimated template and the spectrum observed on June 7, 2019, (JD 2,458,641.452) by EXPRES is shown in solid blue. The curve fitted according to equation (18) is shown in dashed red. For visualization, only 5242–5245 Å is shown.

a function of this new variable using weighted least-squares regression without an intercept to get the single RV estimate, \hat{v}_r .¹¹ The standard error of \hat{v}_r is also easily estimated by the usual least-squares approach. On average across the epochs, this standard error is approximately 0.52 m s^{-1} . An example of what the difference spectrum looks like in the interval 5242–5245 Å, together with the fitted Hermite–Gaussian model, is shown in Figure 16.

For our analysis we used the same 47 observations that were analyzed by the EXPRES team in Petersburg et al. (2020) to estimate the orbital parameters. Several available observations were excluded by the EXPRES team due to low S/N or failure of the laser frequency comb (see Petersburg et al. (2020) for details). The estimated RV’s for all available 51 Pegasi EXPRES spectra using the proposed HGRV method are given in Table 1 of Section 4 of the Supplementary Material (Holzer et al. (2021)). Using the noted 47 EXPRES observations and the RV’s estimated from the HGRV method, we compare the orbital parameters and the overall RV curve fit to those of the CCF method and the FM approach of Petersburg et al. (2020).

The exoplanet orbiting 51 Pegasi has been found to have an eccentricity that is nearly zero (Bedell et al. (2019), Marcy et al. (1997), Wang and Ford (2011), Petersburg et al. (2020)), implying an orbit that is nearly circular. For a nearly circular planetary orbit, the host star’s RV will behave approximately as a sine curve over time. Therefore, we use the Levenberg–Marquardt optimization algorithm (Moré (1978)) to fit a sine curve to the derived RV using

$$(23) \quad v_r(t) = K \sin\left(\frac{2\pi}{P}t + \phi\right) + b.$$

The semiamplitude (K) is initialized at 55.5 m s^{-1} and the period (P) at 4.23 days. The phase (ϕ), representing a horizontal shift of the sine curve, and the RV offset (b), giving the vertical shift, are both initialized at 0. To account for instrumental changes to EXPRES, b is allowed to be different before and after August, 2019. The optimization converges to the fit parameters given in Table 1,¹² and the results of this fitting are shown in Figure 17. Therefore, the HGRV estimation method recovers the well-known parameters for 51 Pegasi. The only pair of parameters that had a significant correlation were the phase, $\hat{\phi}$, and the period, \hat{P} , with correlation -0.813 . All other pairs had correlation magnitudes less than 0.25.

Table 1 also gives the fit parameters from using the RV’s estimated from the CCF and FM methods in Petersburg et al. (2020) for the 47 observations. Similar to the simulation study in

¹¹The usual regression diagnostics should be considered here (e.g., investigating extreme outliers or points with high leverage). No issues were found in this application to 51 Pegasi.

¹²The fitted values of the two offsets are not given in Table 1 since they are expected to differ significantly between the three methods. The HGRV and FM methods give the RV relative to an estimated template, whereas the CCF method gives the RV relative to a prespecified mask.

TABLE 1
Fit parameters of equation (23) for 51 Pegasi

	HGRV	CCF	FM
\hat{K}	$56.48 \pm 0.16 \text{ m s}^{-1}$	$56.20 \pm 0.19 \text{ m s}^{-1}$	$56.17 \pm 0.18 \text{ m s}^{-1}$
\hat{P}	$4.2308 \pm 0.0001 \text{ days}$	$4.2304 \pm 0.0002 \text{ days}$	4.2306 ± 0.0002
$\hat{\phi}$	-1.333 ± 0.006	-1.326 ± 0.007	-1.331 ± 0.007
RMS	0.774 m s^{-1}	0.936 m s^{-1}	0.902 m s^{-1}

Section 4.2, the reduced RMS demonstrates the ability of the HGRV method to outperform the traditional CCF approach.

Including all 56 available spectra gives an estimated $\hat{K} = 56.38 \pm 0.16 \text{ m s}^{-1}$, $\hat{P} = 4.2308 \pm 0.0001 \text{ days}$, $\hat{\phi} = -1.327 \pm 0.005$ and an $\text{RMS} = 0.858 \text{ m s}^{-1}$.

6. Discussion. In this paper we introduce a new approach to estimate the RV in stellar spectra for exoplanet detection that we call the HGRV method. This method works by modeling the differences between observed normalized spectra and an estimated template spectrum. Even though this difference spectrum visually appears to be nothing more than noise (e.g., see Figure 16), there is still an important Doppler signal present. By assuming that absorption features are approximately Gaussian and that $v_r < 500 \text{ m s}^{-1}$, the HGRV method is able to identify this small signal. The application to 51 Pegasi, using spectra from EXPRES, provides an example of how the HGRV-estimated RV's produce a lower RMS in the overall Keplerian fit than the classical CCF approach. Furthermore, the simulation study of Section 4.2 demonstrates that, at low RV, characteristic of Earth-like exoplanets orbiting Sun-like stars, the HGRV approach has higher RV-precision than the CCF.

Theorem 1 implies that the difference flux, imposed on a Gaussian absorption feature by a planetary Doppler shift, can almost entirely be explained as a constant multiple of ψ_1 . This

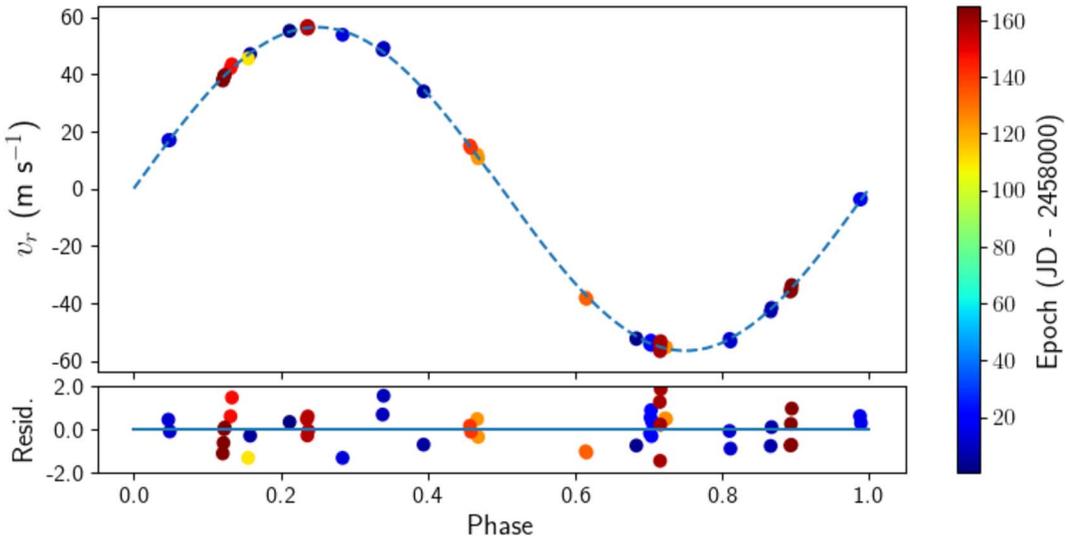


FIG. 17. The RV's derived for 51 Pegasi by the HGRV method, plotted as a function of orbital phase with solid points whose color indicates the epoch, according to the colorbar on the right. All error bars are smaller than the size of the points. The fitted sine curve from equation (23) is also shown in a blue-dashed curve using the HGRV values from Table 1. The residuals are shown in the magnified window at the bottom and have the same units (m s^{-1}) as the plotted RV's.

reduces RV estimation to linear regression with no intercept, where the estimated coefficient is the estimated RV. Therefore, the RV can be interpreted as a proportionality constant between the difference flux and an explanatory variable expressed as a linear combination of first-degree generalized Hermite–Gaussian functions (see equation (18)).

One of the benefits of the HGRV method is the simplification to linear regression, allowing for straightforward statistical inference on the estimated RV. Additionally, linear regression allows heteroskedasticity to be easily addressed with weighted least squares.

Interpolation is only used for stitching together the orders of each observed spectrum and for getting the estimated template spectrum on the same wavelength solution as each observed spectrum. However, the interpolation for stitching orders can be fully avoided by taking each order out to the midpoint of the overlapping regions rather than using weighted averages. Alternatively, each order could be considered on its own as a way to fully avoid stitching orders. Furthermore, the template can be produced with the same wavelength solution, as any observed spectrum, by making these wavelengths the target in the local quadratic regression, therefore removing the need for later interpolation.

We also observed in the 51 Pegasi example that the HGRV method is relatively robust to inaccurate normalization. For example, the difference flux between the observation at JD 2,458,639.958 and the estimated template has a visually identifiable offset from zero, but including this observation’s estimated RV in the orbital parameter estimation of equation (23) slightly reduced the model’s RMS. This robustness may be due to how, on the scale of individual absorption features, inaccurate normalization is approximately an even effect. More work is needed, however, to confirm this general robustness.

An important feature of the HGRV method that also arises from its use of linear regression is its potential to be extended for disentangling Keplerian velocities due to exoplanets from atmospheric velocities due to the star itself. The convective motion and magnetic activity of stars lead to stellar activity in the form of starspots, granulation, faculae, etc. which add red noise to the spectra of stars that can hide a true Doppler shift or temporarily mimic a RV (Desort et al. (2007), Queloz et al. (2001), Saar and Donahue (1997), Meunier, Desort and Lagrange (2010)). Stellar activity can impose a false RV of approximate magnitude 1 m s^{-1} for quiet stars (Hatzes (2002), Lagrange, Desort and Meunier (2010), Isaacson and Fischer (2010)) to hundreds of m s^{-1} for the most active (Saar and Donahue (1997), Paulson, Cochran and Hatzes (2004)). While efforts have been made to model this activity (e.g., Rajpaul et al. (2015), Tuomi et al. (2013), Delisle et al. (2018)) as well as use alternative forms of the cross-correlation method to correct for activity (e.g., Queloz et al. (2001), Simola, Dumusque and Cisewski-Kehe (2019)), these have had limited success in disentangling it from a true Doppler shift at RV’s below 1 m s^{-1} (Dumusque et al. (2017)).

One way the HGRV method could potentially be utilized for disentangling stellar activity from Keplerian Doppler shifts is by approximately orthogonalizing these two effects. The general idea behind this is to find a way by which stellar activity affects absorption features and a Doppler shift does not. Davis et al. (2017) use principal components analysis to show that, at least according to simplified models of starspots and faculae on the Sun, the signals of stellar activity and a Doppler shift are distinguishable. Therefore, stellar activity would change a Gaussian absorption feature in a way that requires more Hermite–Gaussian terms than just ψ_1 , whereas Theorem 1 states that (at least at low RV) a Doppler shift would not. One could then use observations from either the Sun (e.g., Dumusque, Boisse and Santos (2014)) or a star with high stellar activity levels (e.g., Giguere et al. (2016)) to model c_1 in equation (8) as a function of the higher-degree coefficients and remove the RV component that is due only to stellar activity. This is possible because the Hermite–Gaussian functions are orthogonal, and, therefore, as long as the blending between neighboring absorption features is small, a sum of higher-degree Hermite–Gaussian functions would be approximately

orthogonal to the sum of first-degree Hermite–Gaussian functions. These ideas are the topic of future work.

The proposed method does have the limitation that, at high values of RV, c_1 in equation (8) is no longer the only coefficient that is significantly nonzero (see Figure 6), therefore, the HGRV method would not work well. Fortunately, a small fraction of detected exoplanets, none of which are Earth-like, exert such a large RV on their host star. But values of RV well above 500 m s^{-1} easily arise when considering binary star systems.

An improvement that could potentially be made to the proposed method is to relax the assumption of absorption features being Gaussian shaped. The advantage of using this assumption is that its derivative is a constant multiple of a basis function in the well known orthonormal Hermite–Gaussian basis. It is this orthogonality that potentially will allow us to identify signals that are unique to stellar activity and, ultimately, correct for it in the RV estimation. For M-dwarfs, however, where there is essentially no continuum and absorption features overlap significantly, this orthogonality may not be possible. Furthermore, this assumption allows us to quantify with Theorem 1 the approximation error of our model. In order to replace the Gaussian assumption with a more general shape and potentially still model out stellar activity, one may need to have the derivative of the new shape be a basis function in another orthonormal basis.

Data and Python3 code associated with this work is provided as source code in the Supplementary Material (Holzer (2021)) and can be found at https://github.com/parkerholzer/hgrv_method. The HGRV method is also implemented in the open source R package *rvmethod*.

7. Conclusion. By using the mathematical property that Doppler-shifting a Gaussian is nearly the same as adding a first-degree Hermite–Gaussian function, we propose a new method for estimating a Doppler shift in the spectrum of a star. Under the assumptions that the spectrum’s absorption features can be well approximated by a sum of Gaussians and that the true RV is not too large in magnitude, the problem of estimating a RV in the spectrum can be simplified to weighted linear regression with no intercept. By testing this new method on recently collected, high-resolution spectra from EXPRES for the star 51 Pegasi we recover the well-known orbital parameters with an overall RMS (0.774 m s^{-1}) below that of the traditional CCF method (0.936 m s^{-1}). This is only possible because the barycentric corrected wavelengths were provided by the EXPRES team. Furthermore, simulation studies demonstrate the ability of the HGRV method to outperform the CCF approach, giving an RV-precision RMS that is up to approximately 15 cm s^{-1} lower than the CCF. This includes at the level of RV that is characteristic of Earth-like exoplanets orbiting Sun-like stars (i.e., 0.1 m s^{-1}). Unlike many other RV estimation algorithms, the HGRV method easily allows for statistical inference on the estimated RV, does not rely heavily on interpolation, takes account of the functional relationship in neighboring pixels and has a natural extension that could potentially be used to model out the effects of stellar activity.

Acknowledgments. The authors gratefully acknowledge support through NSF-AST 1616086 and NASA XRP 80NSSC18K0443. This work used the EXtreme PREcision Spectrograph (EXPRES) that was designed and commissioned at Yale with financial support by the U.S. National Science Foundation under MRI-1429365 and ATI-1509436 (PI D. Fischer). The authors also gratefully acknowledge the EXPRES team for building this high-fidelity instrument, providing the stellar spectra of 51 Pegasi and the benchmark radial velocities derived with their CCF and FM codes and for helpful discussions. We also thank the Associate Editor and the two referees who provided valuable feedback and suggestions while reviewing this paper. LLZ gratefully acknowledges support from the National Science Foundation Graduate Research Fellowship under Grant No. DGE1122492. These results made use of

the Lowell Discovery Telescope at Lowell Observatory. Lowell is a private, nonprofit institution dedicated to astrophysical research and public appreciation of astronomy and operates the LDT in partnership with Boston University, the University of Maryland, the University of Toledo, Northern Arizona University and Yale University. We thank the Yale Center for Research Computing for guidance and use of the research computing infrastructure.

SUPPLEMENTARY MATERIAL

Supplement to “A Hermite–Gaussian based exoplanet radial velocity estimation method” (DOI: [10.1214/20-AOAS1406SUPPA](https://doi.org/10.1214/20-AOAS1406SUPPA); .pdf). We give the details and additional analysis of the AFF algorithm, together with the mathematical proof of Theorem 1. The model misspecification details in Section 3.4 are applied to additional absorption features, and the estimated RV’s for 51 Pegasi are provided.

Source code for “A Hermite–Gaussian based exoplanet radial velocity estimation method” (DOI: [10.1214/20-AOAS1406SUPPB](https://doi.org/10.1214/20-AOAS1406SUPPB); .zip). Python3 source code for the HGRV theory, the AFF algorithm, and simulation studies. Data from 51 Pegasi, and code that applies the HGRV method to it, is also provided.

REFERENCES

- ANGLADA-ESCUDE, G. and BUTLER, R. P. (2012). The harps-terra project. I. Description of the algorithms, performance, and new measurements on a few remarkable stars observed by harps. *Astrophys. J., Suppl. Ser.* **200** 15.
- ASTUDILLO-DEFRU, N., BONFILS, X., DELFOSSE, X., SÉGRANSAN, D., FORVEILLE, T., BOUCHY, F., GILLON, M., LOVIS, C., MAYOR, M. et al. (2015). The harps search for southern extra-solar planets-xxxvi. Planetary systems and stellar activity of the m dwarfs gj 3293, gj 3341, and gj 3543. *Astron. Astrophys.* **575** A119.
- BARANNE, A., QUELOZ, D., MAYOR, M., ADRIANZYK, G., KNISPEN, G., KOHLER, D., LACROIX, D., MEUNIER, J.-P., RIMBAUD, G. et al. (1996). Elodie: A spectrograph for accurate radial velocity measurements. *Astron. Astrophys. Suppl. Ser.* **119** 373–390.
- BARKER, P. (1984). Ripple correction of high-dispersion iue spectra-blazing echelles. *Astron. J.* **89** 899–903.
- BEDELL, M., HOGG, D. W., FOREMAN-MACKEY, D., MONTET, B. T. and LUGER, R. (2019). Wobble: A data-driven method for precision radial velocities. *Astron. J.* **158** 164.
- BLACKMAN, R. T., SZYMKOWIAK, A. E., FISCHER, D. A. and JURGENSON, C. A. (2017). Accounting for chromatic atmospheric effects on barycentric corrections. *Astrophys. J.* **837** 18.
- BLACKMAN, R. T., FISCHER, D. A., JURGENSON, C. A., SAWYER, D., MCCracken, T. M., SZYMKOWIAK, A. E., PETERSBURG, R. R., ONG, J. M. J., BREWER, J. M. et al. (2020). Performance verification of the EXtreme PREcision Spectrograph. *Astron. J.* **159** 238.
- BOUCHY, F., PEPE, F. and QUELOZ, D. (2001). Fundamental photon noise limit to radial velocity measurements. *Astron. Astrophys.* **374** 733–739.
- BRANCH, M. A., COLEMAN, T. F. and LI, Y. (1999). A subspace, interior, and conjugate gradient method for large-scale bound-constrained minimization problems. *SIAM J. Sci. Comput.* **21** 1–23. [MR1722103 https://doi.org/10.1137/S1064827595289108](https://doi.org/10.1137/S1064827595289108)
- BUTLER, R. P., MARCY, G. W., WILLIAMS, E., MCCARTHY, C., DOSANJH, P. and VOGT, S. S. (1996). Attaining Doppler precision of 3 m s⁻¹. *Publ. Astron. Soc. Pac.* **108** 500.
- CIURYŁO, R. (1998). Shapes of pressure- and Doppler-broadened spectral lines in the core and near wings. *Phys. Rev. A* **58** 1029.
- DAI, C.-Q., WANG, Y. and LIU, J. (2016). Spatiotemporal Hermite–Gaussian solitons of a (3 + 1)-dimensional partially nonlocal nonlinear Schrödinger equation. *Nonlinear Dynam.* **84** 1157–1161. [MR3486561 https://doi.org/10.1007/s11071-015-2560-9](https://doi.org/10.1007/s11071-015-2560-9)
- DAVIS, A. B., CISEWSKI, J., DUMUSQUE, X., FISCHER, D. A. and FORD, E. B. (2017). Insights on the spectral signatures of stellar activity and planets from pca. *Astrophys. J.* **846** 59.
- DELISLE, J.-B., SÉGRANSAN, D., DUMUSQUE, X., DIAZ, R., BOUCHY, F., LOVIS, C., PEPE, F., UDRY, S., ALONSO, R. et al. (2018). The harps search for southern extra-solar planets-xxliii. A compact system of four super-Earth planets orbiting hd 215152. *Astron. Astrophys.* **614** A133.
- DESORT, M., LAGRANGE, A.-M., GALLAND, F., UDRY, S. and MAYOR, M. (2007). Search for exoplanets with the radial-velocity technique: Quantitative diagnostics of stellar activity. *Astron. Astrophys.* **473** 983–993.

- DOPPLER, C. (1842). *Über Das Farbige Licht der Doppelsterne und Einiger Anderer Gestirne des Himmels*. Calve.
- DUMUSQUE, X. (2018). Measuring precise radial velocities on individual spectral lines-I. Validation of the method and application to mitigate stellar activity. *Astron. Astrophys.* **620** A47.
- DUMUSQUE, X., BOISSE, I. and SANTOS, N. (2014). Soap 2.0: A tool to estimate the photometric and radial velocity variations induced by stellar spots and plages. *Astrophys. J.* **796** 132.
- DUMUSQUE, X., BORSA, F., DAMASSO, M., DÍAZ, R. F., GREGORY, P., HARA, N., HATZES, A., RAJPAUL, V., TUOMI, M. et al. (2017). Radial-velocity fitting challenge-ii. First results of the analysis of the data set. *Astron. Astrophys.* **598** A133.
- EINSTEIN, A. et al. (1905). On the electrodynamics of moving bodies. *Ann. Phys.* **17** 50.
- FISCHER, D. A., ANGLADA-ESCUDE, G., ARRIAGADA, P., BALUEV, R. V., BEAN, J. L., BOUCHY, F., BUCHHAVE, L. A., CARROLL, T., CHAKRABORTY, A. et al. (2016). State of the field: Extreme precision radial velocities. *Publ. Astron. Soc. Pac.* **128** 066001.
- GIGUERE, M. J., FISCHER, D. A., ZHANG, C. X., MATTHEWS, J. M., CAMERON, C. and HENRY, G. W. (2016). A combined spectroscopic and photometric stellar activity study of epsilon eridani. *Astrophys. J.* **824** 150.
- GRAY, D. F. (2005). *The Observation and Analysis of Stellar Photospheres*. Cambridge Univ. Press.
- HALVERSON, S., TERRIEN, R., MAHADEVAN, S., ROY, A., BENDER, C., STEFANSSON, G. K., MONSON, A., LEVI, E., HEARTY, F. et al. (2016). A comprehensive radial velocity error budget for next generation Doppler spectrometers. In *Ground-Based and Airborne Instrumentation for Astronomy VI 9908* page 99086P. International Society for Optics and Photonics.
- HAN, E., WANG, S. X., WRIGHT, J. T., FENG, Y. K., ZHAO, M., FAKHOURI, O., BROWN, J. I. and HANCOCK, C. (2014). Exoplanet orbit database. II. Updates to exoplanets.org. *Publ. Astron. Soc. Pac.* **126** 827.
- HATZES, A. P. (2002). Starspots and exoplanets. *Astron. Nachr.* **323** 392–394.
- HOLZER, P. H. (2021). Source code for “A Hermite–Gaussian based exoplanet radial velocity estimation method.” <https://doi.org/10.1214/20-AOAS1406SUPPB>.
- HOLZER, P. H., CISEWSKI-KEHE, J., FISCHER, D. and ZHAO, L. (2021). Supplement to “A Hermite–Gaussian based exoplanet radial velocity estimation method.” <https://doi.org/10.1214/20-AOAS1406SUPPA>.
- ISAACSON, H. and FISCHER, D. (2010). Chromospheric activity and jitter measurements for 2630 stars on the California planet search. *Astrophys. J.* **725** 875.
- JOHNSTON, W. (2014). The weighted Hermite polynomials form a basis for $L^2(\mathbb{R})$. *Amer. Math. Monthly* **121** 249–253. MR3168998 <https://doi.org/10.4169/amer.math.monthly.121.03.249>
- JURGENSON, C., FISCHER, D., MCCRACKEN, T., SAWYER, D., SZYMKOWIAK, A., DAVIS, A., MULLER, G. and SANTORO, F. (2016). Expres: A next generation rv spectrograph in the search for Earth-like worlds. In *Ground-Based and Airborne Instrumentation for Astronomy VI 9908* page 99086T. International Society for Optics and Photonics.
- LAGRANGE, A.-M., DESORT, M. and MEUNIER, N. (2010). Using the sun to estimate Earth-like planets detection capabilities-I. Impact of cold spots. *Astron. Astrophys.* **512** A38.
- LANCZOS, C. (1938). Trigonometric interpolation of empirical and analytical functions. *J. Math. Phys.* **17** 123–199.
- LEET, C., FISCHER, D. A. and VALENTI, J. A. (2019). Towards a self-calibrating, empirical, light-weight model for tellurics in high-resolution spectra. *Astron. J.* **157** 187.
- MARCY, G. W., BUTLER, R. P., WILLIAMS, E., BILDSTEN, L., GRAHAM, J. R., GHEZ, A. M. and JERNIGAN, J. G. (1997). The planet around 51 pegasi. *Astrophys. J.* **481** 926.
- MARHIC, M. E. (1978). Oscillating Hermite–Gaussian wave functions of the harmonic oscillator. *Lett. Nuovo Cimento* (2) **22** 376–378. MR0495931
- MAYOR, M. and QUELOZ, D. (1995). A Jupiter-mass companion to a solar-type star. *Nature* **378** 355.
- MÉSZÁROS, S. and PRIETO, C. A. (2013). On the interpolation of model atmospheres and high-resolution synthetic stellar spectra. *Mon. Not. R. Astron. Soc.* **430** 3285–3291.
- MEUNIER, N., DESORT, M. and LAGRANGE, A.-M. (2010). Using the sun to estimate Earth-like planets detection capabilities-ii. Impact of plages. *Astron. Astrophys.* **512** A39.
- MORÉ, J. J. (1978). The Levenberg–Marquardt algorithm: Implementation and theory. In *Numerical Analysis (Proc. 7th Biennial Conf., Univ. Dundee, Dundee, 1977)*. *Lecture Notes in Math.* **630** 105–116. MR0483445
- PAULSON, D. B., COCHRAN, W. D. and HATZES, A. P. (2004). Searching for planets in the hyades. V. Limits on planet detection in the presence of stellar activity. *Astron. J.* **127** 3579.
- PEPE, F., MAYOR, M., GALLAND, F., NAEF, D., QUELOZ, D., SANTOS, N., UDRY, S. and BURNET, M. (2002). The coralie survey for southern extra-solar planets vii-two short-period saturnian companions to hd 108147 and hd 168746. *Astron. Astrophys.* **388** 632–638.

- PETERSBURG, R. R., ONG, J. M. J., ZHAO, L. L., BLACKMAN, R. T., BREWER, J. M., BUCHHAVE, L. A., CABOT, S. H. C., DAVIS, A. B., JURGENSON, C. A. et al. (2020). An extreme precision radial velocity pipeline: First radial velocities from EXPRES. *Astron. J.* **159** 187.
- PLANCK, M. (1901). On the law of distribution of energy in the normal spectrum. *Ann. Phys.* **4** 1.
- QUELOZ, D., HENRY, G., SIVAN, J., BALIUNAS, S., BEUZIT, J., DONAHUE, R., MAYOR, M., NAEF, D., PERRIER, C. et al. (2001). No planet for hd 166435. *Astron. Astrophys.* **379** 279–287.
- RAJPAUL, V. M., AIGRAIN, S. and BUCHHAVE, L. A. (2020). A robust, template-free approach to precise radial velocity extraction. *Mon. Not. R. Astron. Soc.* **492** 3960–3983.
- RAJPAUL, V., AIGRAIN, S., OSBORNE, M. A., REECE, S. and ROBERTS, S. (2015). A Gaussian process framework for modelling stellar activity signals in radial velocity data. *Mon. Not. R. Astron. Soc.* **452** 2269–2291.
- RIFFEL, R. A. (2010). Profit: A new alternative for emission-line profile fitting. *Astrophys. Space Sci.* **327** 239–244.
- RIMMELE, T. R. and RADICK, R. R. (1998). Solar adaptive optics at the national solar observatory. In *Adaptive Optical System Technologies* **3353** 72–82. International Society for Optics and Photonics.
- SAAR, S. H. and DONAHUE, R. A. (1997). Activity-related radial velocity variation in cool stars. *Astrophys. J.* **485** 319.
- SIMOLA, U., DUMUSQUE, X. and CISEWSKI-KEHE, J. (2019). Measuring precise radial velocities and cross-correlation function line-profile variations using a skew normal density. *Astron. Astrophys.* **622** A131.
- SNEDEN, C., BEAN, J., IVANS, I., LUCATELLO, S. and SOBECK, J. (2012). Moog: Lte line analysis and spectrum synthesis. Astrophysics Source Code Library.
- TUOMI, M., ANGLADA-ESCUDE, G., GERLACH, E., JONES, H. R., REINERS, A., RIVERA, E. J., VOGT, S. S. and BUTLER, R. P. (2013). Habitable-zone super-Earth candidate in a six-planet system around the k2. 5v star hd 40307. *Astron. Astrophys.* **549** A48.
- WANG, J. and FORD, E. B. (2011). On the eccentricity distribution of short-period single-planet systems. *Mon. Not. R. Astron. Soc.* **418** 1822–1833.
- WRIGHT, J. and EASTMAN, J. (2014). Barycentric corrections at 1 cm s⁻¹ for precise Doppler velocities. *Publ. Astron. Soc. Pac.* **126** 838.
- XU, X., CISEWSKI-KEHE, J., DAVIS, A. B., FISCHER, D. A. and BREWER, J. M. (2019). Modeling the echelle spectra continuum with alpha shapes and local regression fitting. *Astron. J.* **157** 243.