Submitted to the Annals of Applied Statistics

A HERMITE-GAUSSIAN BASED RADIAL VELOCITY ESTIMATION METHOD

By Parker Holzer* and Jessi Cisewski-Kehe* and Debra Fischer† and Lily Zhao†

*Department of Statistics & Data Science, Yale University**

Department of Astronomy, Yale University†

Abstract 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 star's spectrum. We introduce a new, mathematically rigorous, approach to detect such a signal that accounts for functional relationships of neighboring wavelengths, minimizes the role of wavelength interpolation, accounts for heteroskedastic noise, and easily allows for statistical inference. 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}. Furthermore, the new method outperforms the traditional Cross-Correlation Function approach by achieving an estimation error that is on average 10 cm s^{-1} lower. The proposed method is also demonstrated on a new set of observations from the EXtreme PREcision Spectrometer (EXPRES) for the star 51 Pegasi, and successfully recovers estimates that agree well with previous studies of this planetary system. Data and Python3 code associated with this work can be found at *https : //github.com/parkerholzer/hgrv_method*. The method is also implemented in the open source R package *rvmethod*.

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 stars, or exoplanets. This discovery was made using the radial velocity (RV) method (also known as the Doppler technique). The RV method makes use of stellar spectra to derive the radial component of stellar velocity over time. Orbiting planets will tug the star around a common center of mass, producing a cyclical variation in the velocity of the target star with the same period as the orbiting planet.

The data for the RV method are obtained with a spectrograph. 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 outer atmosphere (hereafter referred to as the photosphere) of the star absorb specific wavelengths of light, corresponding to the quantum mechanical energy level differences in the absorbing atoms. As the star moves toward us or away from us, the velocity component that is projected along our line of site, i.e., the radial velocity, produces a wavelength rescaling in the spectrum that is described by the Doppler equation.

All stars orbit the galaxy and will exhibit a nearly constant radial velocity relative to the Sun. If a star also has a planet, then the orbiting planet will tug 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 radial velocity signal depends on several factors, including the mass of the star, the mass of the planet, the orbital period, the shape (eccentricity), and the orientation of the orbit. 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 radial velocity of about 12 m s^{-1} in the Sun while the lower mass Earth only induces a velocity of 0.09 m s^{-1}. If observed with very high spectral resolution, one pixel on the detector spans about 500 m s^{-1}, so these radial velocities 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 semi-amplitudes of nearly sinusoidal variations with periods of about 12 years for Jupiter and 1 year for the Earth. Detecting such a tiny sub-pixel shift in stellar absorption features is non-trivial. 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, 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 over the past 25 years.

The RV error budget includes instrumental errors, photon statistics (shot noise), and velocities from within the photosphere of the star that introduce scatter to the Keplerian velocities (Halverson et al., 2016; Dumusque et al., 2017; Blackman et al., 2020). The EXtreme PREcision Spectrometer (EXPRES) (Jurgenson et al., 2016; Blackman et al., 2020; 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 the instrument 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 velocities with high precision and

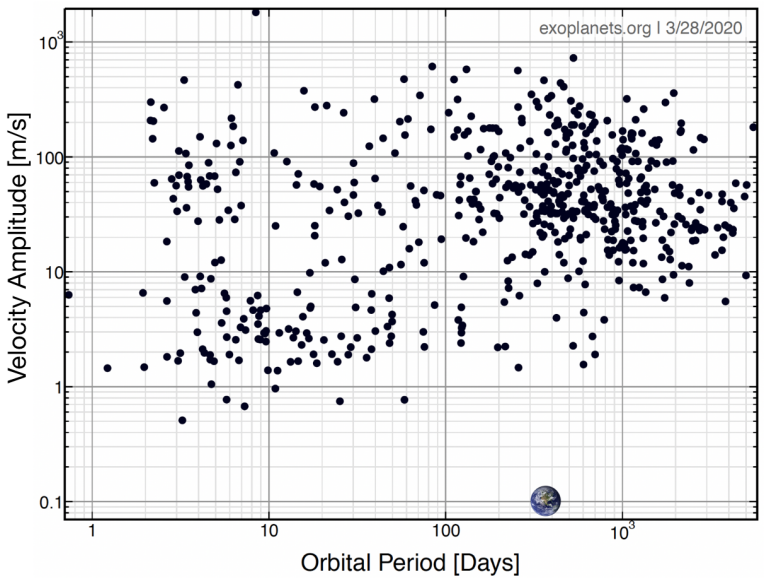

FIGURE 1. *Orbital period and stellar RV semi-amplitude for all exoplanets discovered with the RV Method. Data come from Exoplanets.org (Han et al., 2014) on March 28, 2020 with a total of about* 800 *exoplanets. Note that with an orbital period of* 365.25 *days and a semi-amplitude of approximately*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

are less sensitive to photospheric velocities (Dumusque et al., 2014; Rajpaul et al., 2015; Dumusque et al., 2017; Davis et al., 2017; Rajpaul et al., 2020).

The traditional cross correlation function (CCF) (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 et al., 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-Å 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 different types of stars because the modeling is carried out on the

spectra observed for an individual star, and it 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. 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. This allows the method to easily account for the heteroskedastic noise in spectra. Furthermore, this simplification allows for straight-forward statistical inference on the estimated RV without making assumptions regarding the validity of propagation error 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 modeling with Schrodinger's Equation (Marhic, 1978; Dai et al., 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 partially a consequence of the method's assumption that absorption features are Gaussian shaped. While the traditional CCF approach is designed to not depend on the individual shapes of absorption features by its use of a mask, and the template matching approaches take full account of absorption feature shapes, the HGRV approach can be thought of as between these two extremes in that it assumes the features are Gaussian-shaped. 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 flux) 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 astrophysical blackbody effect (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 (Xu et al., 2019; Petersburg et al., 2020). A spectrum where the continuum has been normalized by dividing out the instrumental blaze and the blackbody curve is hereafter referred to as a normalized spectrum. Figure 2 is an example of such a normalized spectrum.

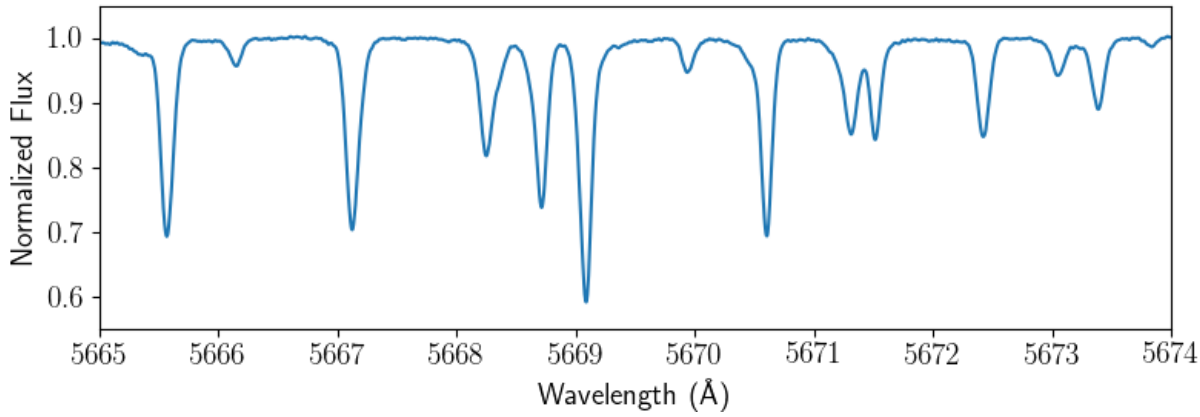

FIGURE 2. *A subset of the NSO spectrum of the Sun between* 5665 *and* 5674 *Å.*

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 et al., 2001; Rajpaul et al., 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 slopes 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 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 presented in Algorithm 1 and requires three tuning parameters: a wavelength window size m in units of pixel count, and significance levels $\alpha,\ \eta$ where $\eta \geq \alpha$. For a more thorough motivation of this algorithm, as well as a more detailed overview of the steps involved, see Appendix A.

Algorithm 1: Absorption Feature Finder

Data: ordered wavelengths $\Lambda = (x_0, x_1, ..., x_n)$ and corresponding flux values $\tau = (\tau_0, \tau_1, ..., \tau_n)$
Initialize tuning parameters $m \in \mathbb{N}$, $\alpha \in (0,1)$, and $\eta \in (\alpha, 1)$
for $x_i \in \Lambda$ **do**
  set $\Lambda_{l,i} = (x_{i-m+1}, x_{i-m+2}, ..., x_i)^T$, $\Lambda_{r,i} = (x_i, x_{i+1}, ..., x_{i+m-1})^T$, $\tau_{l,i} = (\tau_{i-m+1}, \tau_{i-m+2}, ..., \tau_i)^T$, and $\tau_{r,i} = (\tau_i, \tau_{i+1}, ..., \tau_{i+m-1})^T$
  model $\tau_{l,i} = \beta_{0,l}\mathbb{1}_m + \beta_{1,l}\Lambda_{l,i} + \varepsilon$ and $\tau_{r,i} = \beta_{0,r}\mathbb{1}_m + \beta_{1,r}\Lambda_{r,i} + \varepsilon'$ where $\varepsilon, \varepsilon' \sim N\left(0, \varsigma^2 I_m\right)$ and $\mathbb{1}_m = (1, 1, ..., 1)^T$ with length m
  get p-values $p_{l,i}$ for testing $\beta_{1,l} = 0$ against $\beta_{1,l} < 0$ and $p_{r,i}$ for testing $\beta_{1,r} = 0$ against $\beta_{1,r} > 0$
end
Initialize index $j = m$ and upperbound $u = 0$
while $j \leq$ length$(\Lambda) - m + 1$ **do**
  if $p_{l,j} < \alpha/2$ and $p_{r,j} < \alpha/2$ **then**
    set $k_{\max} = \max\{k \in \{u, u+1, ..., j\} : p_{l,k} \geq \eta\}$
    set $k_{\min} = \min\{k \in \{j, j+1, ..., \text{length}(\Lambda)\} : p_{r,k} \geq \eta\}$
    save $\left(\frac{x_{k_{\max}} + x_{k_{\max}-m}}{2}, \frac{x_{k_{\min}} + x_{k_{\min}+m}}{2}\right)$ as absorption feature wavelength bounds
    $j \leftarrow \lfloor (k_{\min} + m/2) \rfloor$
    $u \leftarrow j$
  else
    $j \leftarrow j + 1$
  end
end

Algorithm 1 was empirically evaluated using the NSO spectrum. After the step-by-step optimization of the three tuning parameters described in Appendix A, 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 3.

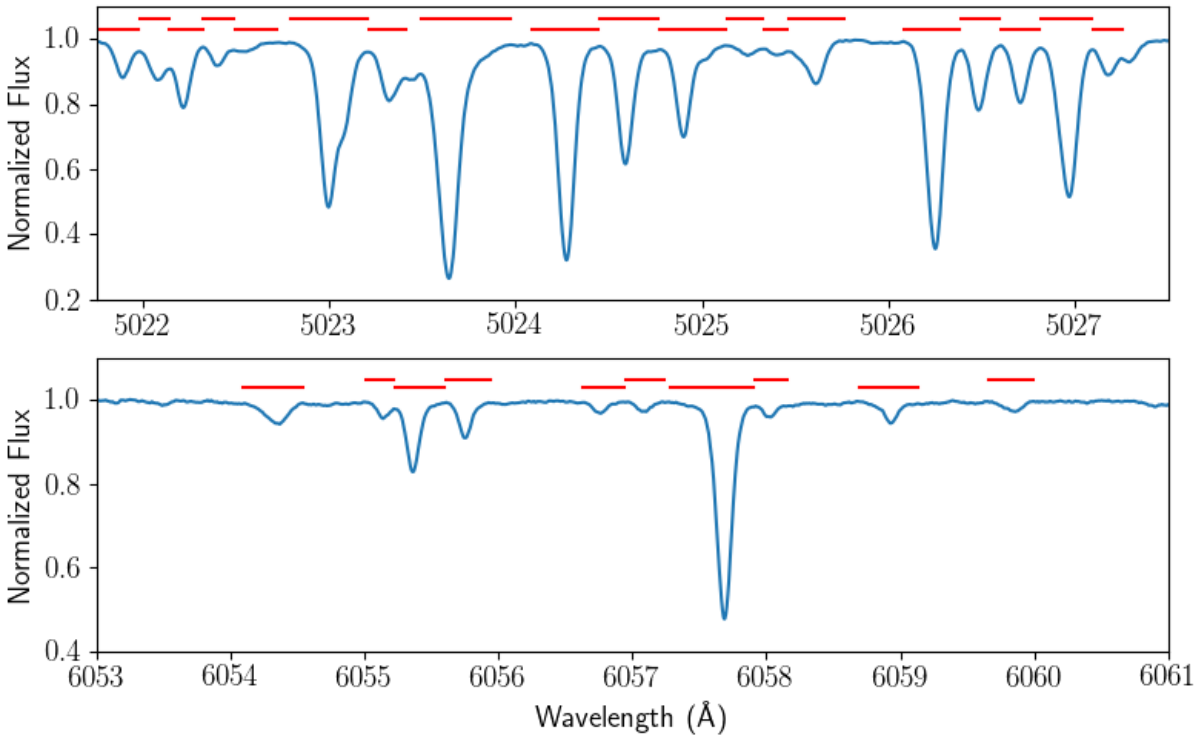

FIGURE 3. *Results of using Algorithm 1 on the NSO Spectrum. Red horizontal lines show the wavelength windows found to correspond to individual absorption features.*

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 Algorithm 1. Applying Algorithm 1 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 1 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 $0.015 \times \frac{500}{\text{S/N}}$. For details on this recommendation see Appendix A.

In addition we applied Algorithm 1 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 Appendix A.

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 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 outwards 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 Algorithm 1, 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 between a Gaussian and a multiplicative shift of it. 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 Appendix B). 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

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

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 (Wright and Eastman, 2014; Blackman et al., 2017; 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

$$\xi = 1 + \frac{v_r}{c}. \tag{2}$$

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

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

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

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

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 4.

According to Johnston (2014),

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

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

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

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,

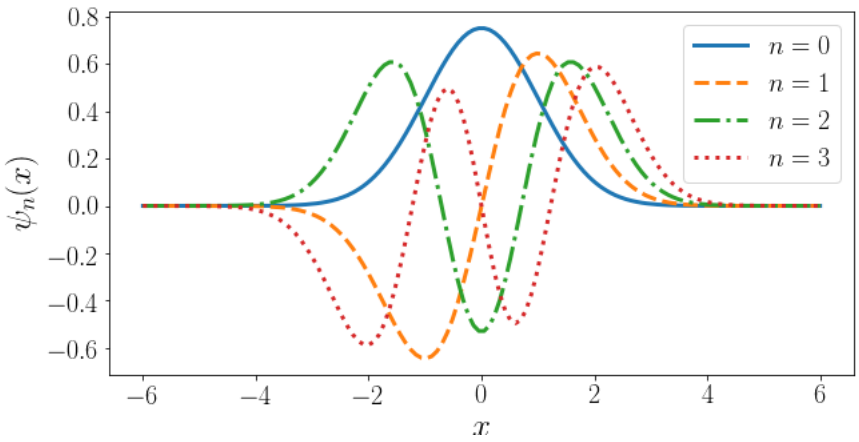

FIGURE 4. *The first* 4 *Hermite-Gaussian functions given by Equation* (3).

2014). One can also generalize the definition of the Hermite-Gaussian functions to have a general location, μ, and scale, σ:

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

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

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

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* $g(x;\xi) = e^{-\frac{(x-\mu)^2}{2\sigma^2}} - e^{-\frac{(\xi x-\mu)^2}{2\sigma^2}}$ *decomposed in the Hermite-Gaussian basis as* $g(x;\xi) = \sum_{n=0}^{\infty} c_n(\xi)\psi_n(x;\mu,\sigma)$,

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

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_x, is significantly larger than the width, σ_x, 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 Algorithm 1 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.

Some of the RV detection algorithms, such as the template matching method described in Bouchy et al. (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, multiplication of 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 multiplicative 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* $I_k(a,b,c) := \int_{-\infty}^{\infty} u^k e^{-\left(au^2+bu+c\right)} du$ *where* $a > 0$, *we have that*

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

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

$$\text{and for all } k \geq 2,\ I_k(a,b,c) = -\frac{b}{2a} I_{k-1}(a,b,c) + \frac{k-1}{2a} I_{k-2}(a,b,c). \tag{12}$$

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

LEMMA 2. *For* $g(x;\xi) = e^{-\frac{(x-\mu)^2}{2\sigma^2}} - e^{-\frac{(\xi x-\mu)^2}{2\sigma^2}}$ *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* $\varepsilon = \xi - 1$

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

and for all $k \geq 1$,

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

Using Lemmas 1 and 2 we numerically calculate the first seven coefficients as a function of RV and illustrate the results in Figure 5. 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

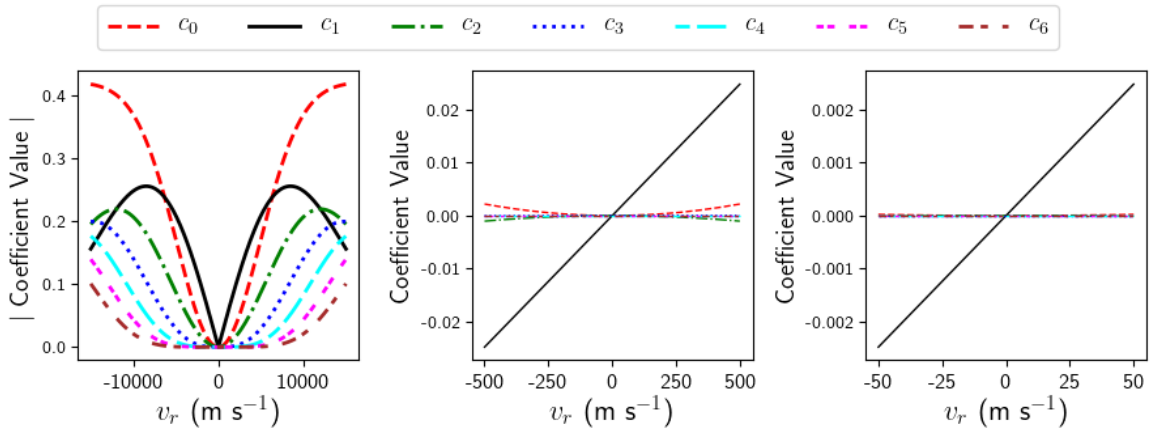

FIGURE 5. *The coefficient solutions that result from modeling a Doppler-shifted Gaussian with the Hermite-Gaussian basis are plotted here as a function of v_r. The left panel has the absolute value of the coefficients on the vertical axis and illustrates that at low values of v_r, c_1 is the dominating 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.*

zero-function. More importantly, though, Figure 5 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 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 $D(\phi||\varphi)$ as

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

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

LEMMA 3. *For* $g(x;\xi) = e^{-\frac{(x-\mu)^2}{2\sigma^2}} - e^{-\frac{(\xi x-\mu)^2}{2\sigma^2}}$ *decomposed as* $g(x;\xi) = \sum_{n=0}^{\infty} c_n(\xi)\psi_n(x;\mu,\sigma)$, *we have that*

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

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

Combining Lemmas 3 and 4 completes the proof of Theorem 1. (See Appendix B for a more detailed proof of each.)[2]

Theorem 1 does not explicitly give a rate at which the standardized approximation error approaches its limit. But by using Lemma 3 and Equation (52) from the proof of Lemma 4 in Appendix B, we illustrate the rate with Figure 6. Note that the standardized approximation error shown here is bounded between 0 and 1, and that the limit is actually non-zero. Figure 6 illustrates that as $\xi \to 1$, $D\left(g(x;\xi)||c_1(\xi)\psi_n(x;\mu,\sigma)\right)$ approaches its limit quadratically and that when $v_r < 50$ 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

[1] Since $g(x;\xi)$ approaches the zero function as $\xi \to 1$, and for any $k \geq 0$ $c_k(\xi) \to 0$ as $\xi \to 1$, 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 $D\left(g(x;\xi)||c_1(\xi)\psi_1(x;\mu,\sigma)\right)$ adjusts for this by standardizing the quantity.

[2] It is worth noting that in the proof of Lemma 4, L'hopital's rule must be applied twice. And since $c_1^2(\xi)$ is essentially an integral, one would naturally suggest that the proof could be simplified by interchanging two derivatives and the limit with the integration in both the numerator and denominator. However, it can be shown that this results in the limit of Lemma 4 incorrectly being 1. Therefore, this is an instance in which this interchange is not mathematically valid and cannot be used to simplify the proof.

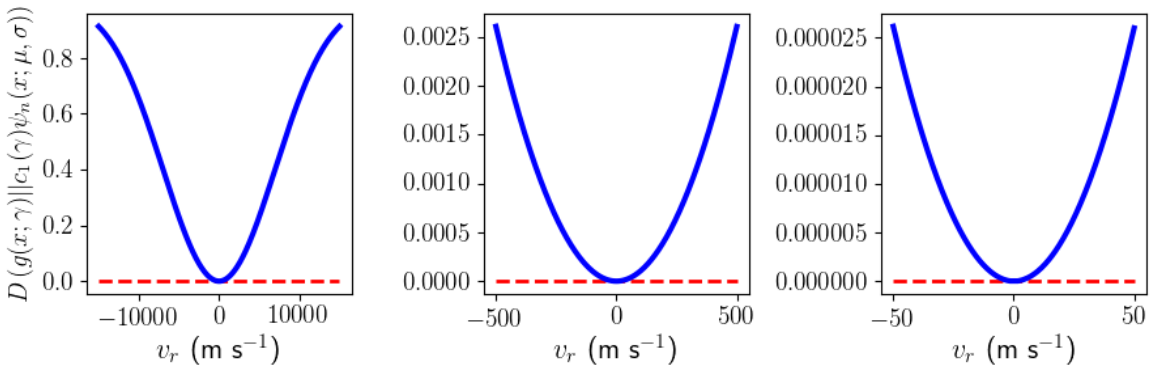

FIGURE 6. *The standardized approximation error* $D\left(g(x;\xi)||c_1(\xi)\psi_n(x;\mu,\sigma)\right)$ *in Theorem 1 as a function of* v_r *with parameters* $\mu = 5000$ *and* $\sigma = 0.1$ *is plotted in bold. The limit is also shown in the horizontal red dashed line.*

to the difference between a template spectrum and a Doppler-shifted spectrum and map the fitted coefficient to a RV. As illustrated in Figure 5, c_1 at low values of v_r is directly proportional to v_r.

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

Furthermore, using Equation (2) with $\varepsilon = \xi - 1$, we have that the mapping from ε to RV is $v_r(\varepsilon) = c\varepsilon$ and $\lim\limits_{\varepsilon\to 0}\dfrac{\partial}{\partial\varepsilon}v_r(\varepsilon) = c$. Hence, $\lim\limits_{\varepsilon\to 0}\dfrac{\partial}{\partial v_r}c_1\left(v_r(\varepsilon)\right) = \dfrac{\mu\sqrt{\sqrt{\pi}}}{c\sqrt{2\sigma}}$ which is the desired proportionality constant. So the proportionality that is valid at low values of RV, v_r, is

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

The strongest assumption made when applying the theorem is that the absorption features are Gaussian shaped. Because this may never be 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

(30) in the proof of Lemma 2, 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

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

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 $\varepsilon_i \sim N\left(0, \varrho_i^2\right)$ 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, we assume that the provided uncertainties, $\hat{\varrho}_i$, are accurate estimates of each ϱ_i, and estimate v_r in Equation (18) through weighted least squares with weights $w_i = 1/\hat{\varrho}_i^2$.

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 $v_r < 500$ 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 include interpolation[4], treats neighboring pixels similarly, accounts for the hetroskedastic noise, and easily allows for statistical inference.

3.4. *Model Misspecification.* The HGRV method assumes that the shape of absorption features is Gaussian, which often does not hold exactly. Various reasons are understood to contribute to this: a line following the Voigt profile may have a non-negligible Lorentzian component, the line may be deep enough to depart from the Voigt profile, or there may

[3] If 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.

[4] Interpolation 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.

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 7, 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 7.

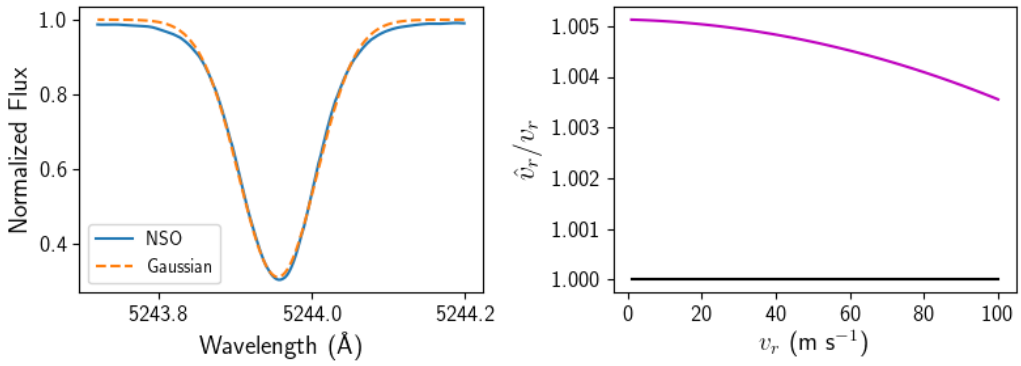

FIGURE 7. *Results for analyzing the effects of misspecifying the model of the absorption feature in the NSO spectrum between* 5243.7 *and* 5244.2 *Å as a Gaussian. The left panel shows the feature in solid blue and the best-fit Gaussian in dashed orange. The right panel shows the ratio of the RV estimated with Equation* (18) $\hat{v}_r$ *(with* $n = 1$*) and the true RV,* v_r*.*

Figure 7 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 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 Appendix C.

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 (Sneden et al., 2012). However,

in practice, these stellar parameters and the atomic line transition data are not well known enough to make this feasible. Therefore, we take a data-driven approach.

The method we propose for estimating the template is to stack all normalized, barycentric corrected, observed spectra across time epochs and fit a smooth curve to the combined spectrum 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 stacked 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 8 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 stacking 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 pre-specified 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.

[5]Using the stacking and smoothing template estimate 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

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

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) 2458641.952. 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 $t_1, ..., t_N$ where $t_k \sim$ iid Uniform$(0, 2\pi)$, (ii) calculate RV's $v_{r,1}, ..., v_{r,N}$ 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), (iv) apply the template estimation method described in Section 3.5 and calculate the resulting RMS$(\hat{\tau})$.

In our simulations, the number of spectra, N, ranges from 1 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 $(4965, 4970)$ which has a higher density of absorption features, as well as the window $(6381, 6386)$ 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 8 which shows the average RMS$(\hat{\tau})$ on the left panel, and the standard deviation of the RMS$(\hat{\tau})$ on the right, for each pair of S/N and number of spectra.

The left plot in Figure 8 illustrates that once the number of spectra reaches approximately 21, the average 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 RMS$(\hat{\tau})$ is more affected by the number of spectra than the S/N in this example.

The right plot in Figure 8 illustrates how the 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

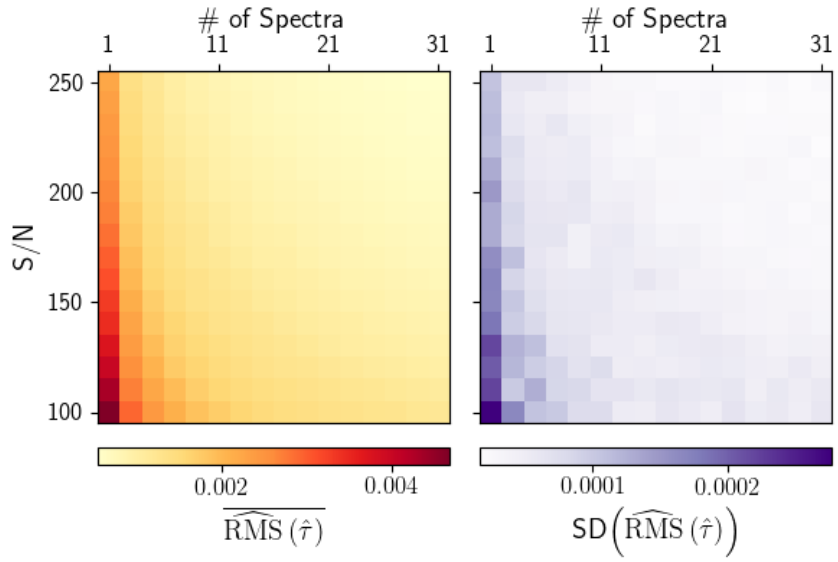

FIGURE 8. *Simulation study results for estimating the template spectrum between* 5240 *and* 5245 *Å. For each S/N and number of spectra, N, 50 simulations were carried out each with a different cadence. Each simulation involved estimating the template with local quadratic regression and calculating 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.*

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.

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

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

where n is the number of simulations at RV v_r. The square of $\widehat{\mathrm{RMS}}\left(\hat{v}_r\right)$ 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

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

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

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

We explore how the RMS($\hat{v}_r$), Bias($\hat{v}_r$), and SD($\hat{v}_r$) vary with S/N and v_r. Our simulation takes 5 equally spaced values of S/N $100, 150, ..., 300$ and 4 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 2458639.958 (see Section 5 for more details) and including Poisson noise of the specified S/N. The results for obtaining each $\hat{v}_r$ with the HGRV method are shown in Figure 9.

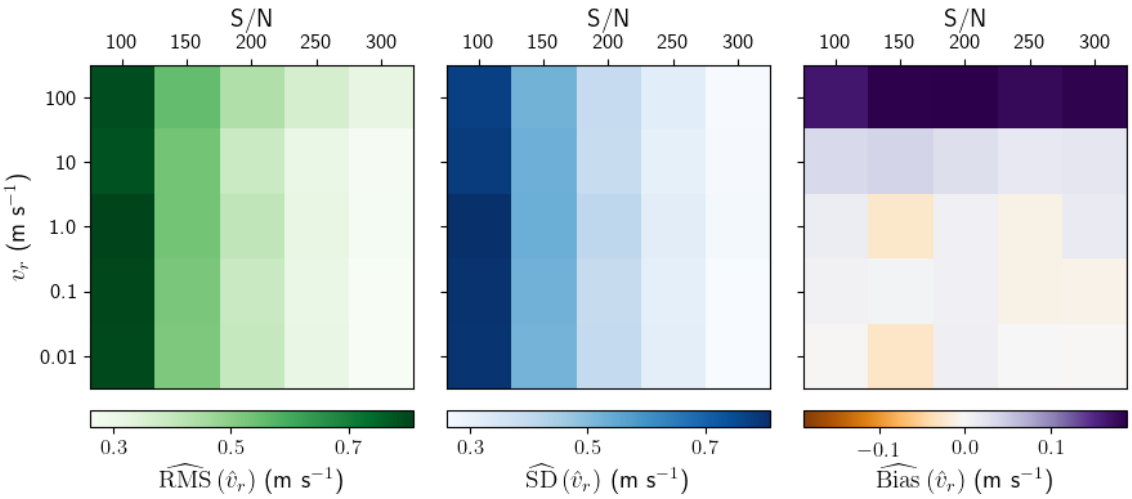

FIGURE 9. *The results for applying the HGRV method to spectra simulated from the estimated 51 Pegasi template spectrum. The left, middle, and right panels show the estimated RMS, SD, and bias of the estimated RV respectively. All three panels share the same vertical axis that represents the true RV each spectrum was simulated with. 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 left panel of Figure 9 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 9 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.[6]

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 HARPS G2 mask. 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 (-33168.5399 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 10 shows the difference in RMS between the HGRV and CCF methods. Since every pair of S/N and v_r in Figure 10 shows a negative RMS difference, this suggests that the HGRV method has higher RV-precision than the CCF approach in this regime. The EXPRES team finds that their FM code

likewise gives a

similarly improved fitting.

As a more detailed summary of the RMS improvement of the HGRV as demonstrated by Figure 10, the difference in the estimated SD and absolute bias (the sum of squares of which equal the squared RMS) is shown in Figure 11.

Figures 10 and 11 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.

To check the stability of this simulation, we used the wavelength solution for the 51 Pegasi spectrum from EXPRES observed on JD 2458804.588 instead of the wavelength solution from JD 2458639.958 used above. Running the HGRV and CCF approach each with 2000 independent simulations with $v_r = 1$ m s^{-1} and a S/N of 200 produced an RMS difference of -0.094 m s^{-1}. All estimated RVs from the CCF and HGRV methods for these simulations are provided 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 semi-amplitude of 55.57 ± 2.22 m s^{-1}and orbital period of 4.2292 ± 0.0003 days (Mayor and Queloz, 1995; Marcy et al., 1997; 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

[6]We 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 2.5×10^{-3} m s^{-1}. This demonstrates that the HGRV method has the capability of obtaining a RV precision less than 0.1 m s^{-1}.

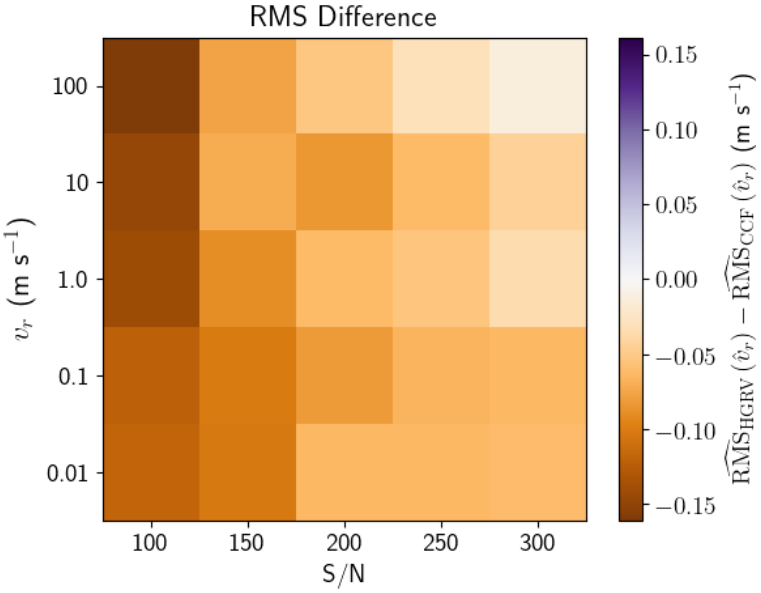

FIGURE 10. *The difference between the HGRV and CCF RMS for each pair of S/N and true v_r. Each pair consisted of* 2000 *independent simulations for each method. The difference is indicated on the right by the color bar which is centered at* 0.0 *m s^{-1}, and demonstrates the higher RV-precision of the HGRV method.*

observations of 51 Pegasi were unable to avoid, allowing for greater precision of derived RV. Our dataset consists of 56 observed spectra from JD 2458639 to 2458805 (June 5, 2019 to Nov. 18, 2019). The S/N of these spectra ranges from 89 to as high as 385, but most are close to 200 (see Table 2 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

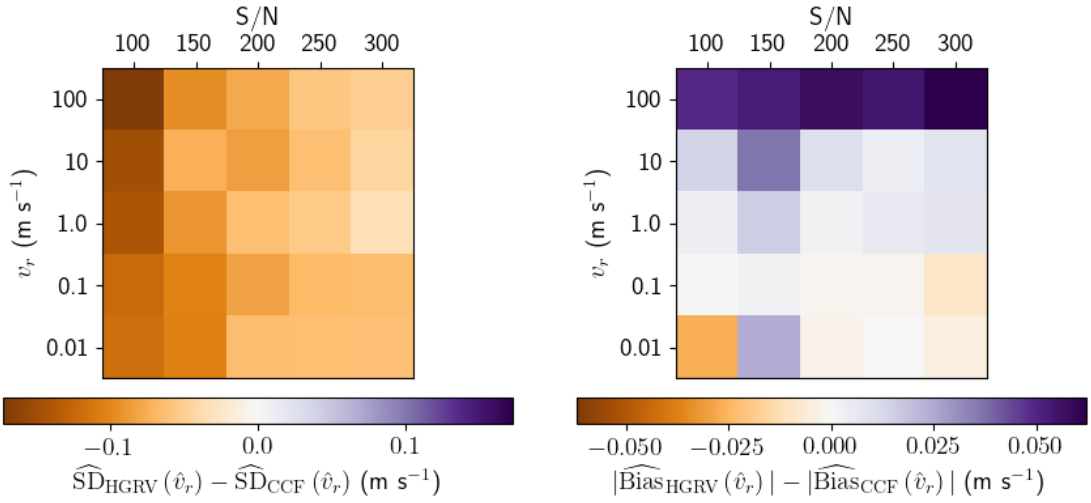

FIGURE 11. *The difference between the HGRV and CCF standard deviation and absolute bias for each pair of S/N and true* v_r. *Each pair consisted of* 2000 *independent simulations for each method. The differences are indicated below each panel by the color bars which are centered at* 0.0 $m\ s^{-1}$.

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 EXPRES with each spectrum that was created using the approach of Leet et al. (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.

Because a spectrum covers over 3000 Å 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 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 (point-wise) 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 12.

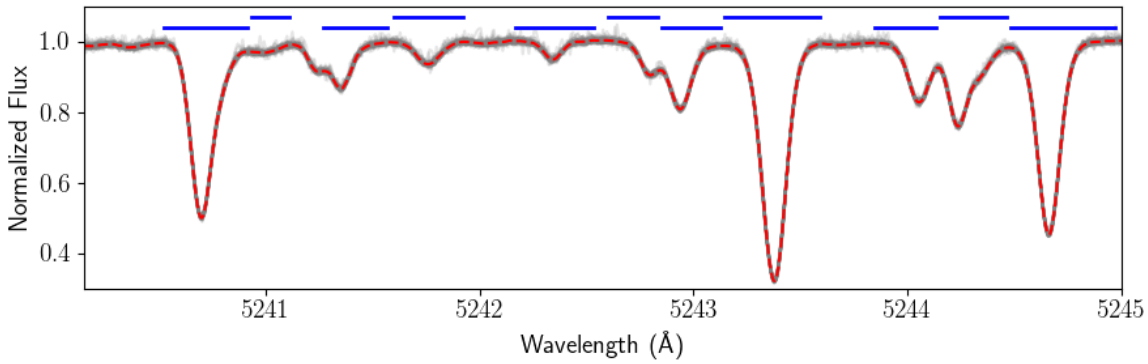

FIGURE 12. *A subset of the estimated template spectrum calculated from 51 Pegasi data is shown in the red dashed line on top of all observed spectra used in the calculation (shown in gray). The feature bounds that result from running Algorithm 1 on the estimated template spectrum are also shown in blue 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* 99*th percentile of* 1.1×10^{-3}.

Once we have the high S/N estimated template spectrum we can use it in Algorithm 1 to find absorption feature wavelength bounds. The tuning parameters of the algorithm that were found through the optimization process described in Appendix A 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 12 for the section of the spectrum displayed. Note that when neighboring features are strongly blended together, Algorithm 1 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 et al., 1999), which allows for initialization and bounds for each parameter to be fitted in non-linear 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 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[7], 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 Algorithm 1, 3868 were well-fitted with one or two Gaussians. An example of the fit model spectrum is shown in Figure 13. Most of the features that were eliminated at this stage were strongly blended with one or more neighboring features.

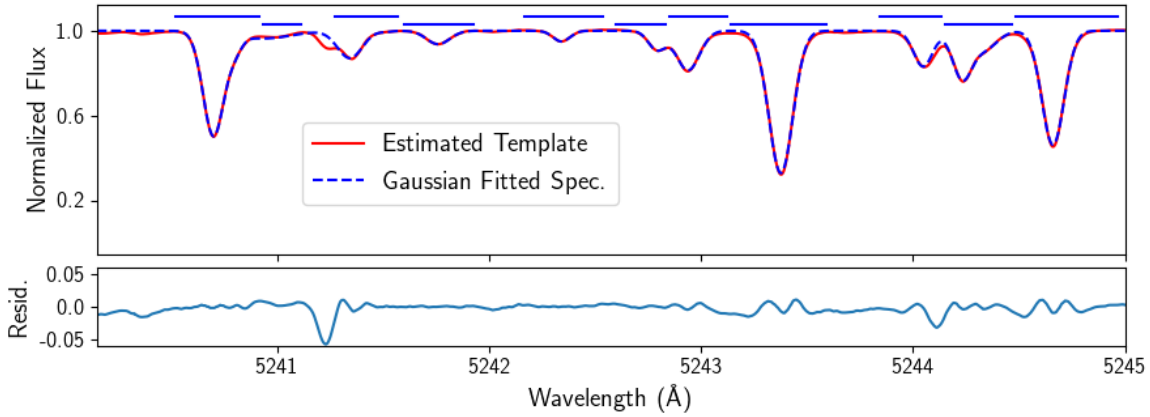

FIGURE 13. *The estimated template spectrum for 51 Pegasi is shown in solid red with the spectrum that approximates it as a sum of Gaussians shown in dashed blue. The full spectra go from* $4470-6800$ *Å, but for visualization only* $5240-5245$ *Å are displayed. All absorption features in this wavelength 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.*

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. Below 4850 Å the noise of

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

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. We furthermore 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 epoch[8] (Mészáros and Prieto, 2013), we calculate the difference spectrum between the two. We then transform each wavelength x_i using the sum, $\sum_{j=1}^{n} -\frac{\sqrt{\sqrt{\pi}}d_j\mu_j}{c\sqrt{2\sigma_j}}\psi_1\left(x_i;\mu_j,\sigma_j\right)$, from Equation (18). This transformation uses all fitted Gaussian parameters, after which we model the difference flux across the full stitched spectrum as a function of this new variable using weighted least-squares regression without an intercept to get the single RV estimate, $\hat{v}_r$.[9] 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 14.

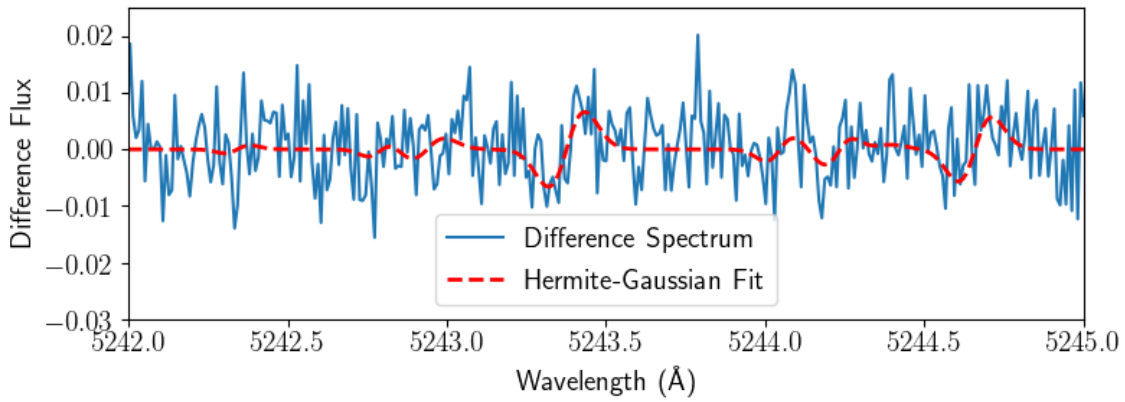

FIGURE 14. *The difference spectrum between the estimated template and the spectrum observed on June 7, 2019 (JD* 2458641.452*) by EXPRES is shown in solid blue. The curve fitted according to Equation* (18) *is shown in dashed red. For visualization, only* $5242 - 5245$ *Å is shown.*

Although we have a total of 56 spectra for 51 Pegasi that we derive individual RV's for with the HGRV method, for comparison we only use the 47 observations that were analyzed in Petersburg et al. (2020) to estimate orbital parameters. The reasons for why these 9 other spectra were not included are given in Table 2. All estimated RV's, together with an

[8] This is the only time in the proposed method that interpolation takes place.

[9] The 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.

indicator of whether or not they were among these 47, are given in Table 2 of Appendix D. Using these observations, and the RV's estimated from the HGRV method, we compare the orbital parameters and the overall fit to that 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 (Marcy et al., 1997; Wang and Ford, 2011; Bedell et al., 2019; 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

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

The semi-amplitude (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[10], and the results of this fitting are shown in Figure 15. 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, $\hat{P}$, which was -0.813. All other pairs had a correlation magnitudes less than 0.25.

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

TABLE 1
Fit parameters of Equation (23) for 51 Pegasi.

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

[10]The 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 pre-specified mask.

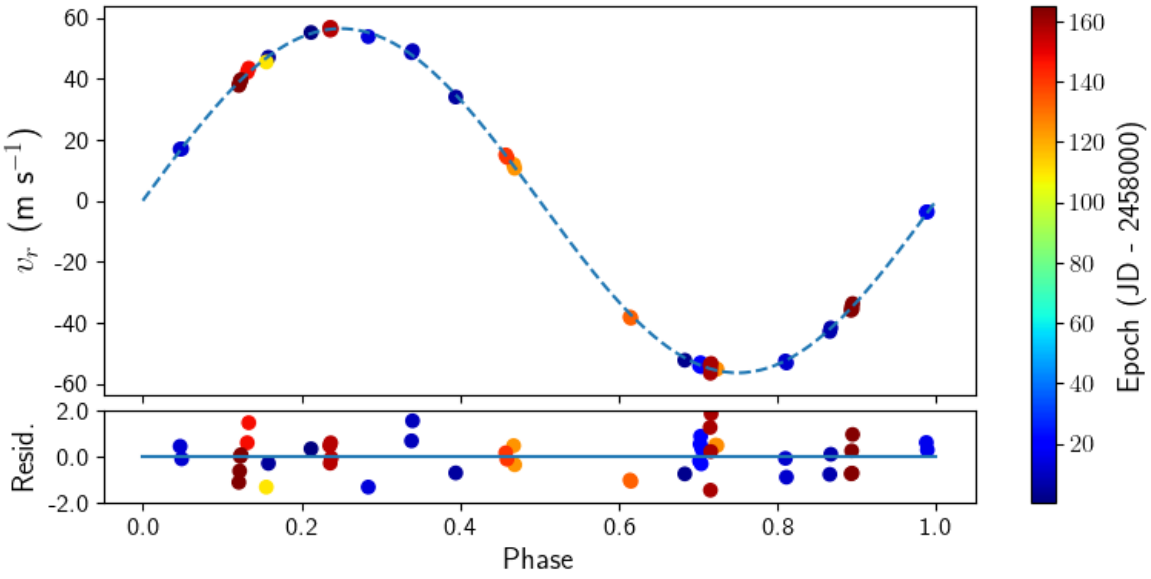

FIGURE 15. *The RV's derived for 51 Pegasi by the HGRV method, plotted as a function of orbital phase with solid points whose color indicates the epoch according to the colorbar on the right. All error bars are smaller than the size of the points. The fitted sine curve from Equation (23) is also shown 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.*

Including all 56 spectra gives an estimated $\hat{K} = 56.38 \pm 0.16$ m s^{-1}, $\hat{P} = 4.2308 \pm 0.0001$ days, $\hat{\phi} = -1.327 \pm 0.005$, and an RMS= 0.858 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 14), there is still an important Doppler signal present. By assuming that absorption features are approximately Gaussian and that $v_r < 500$ 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 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 straight-forward 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 2458639.958 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 photospheric 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 (Saar and Donahue, 1997; Queloz et al., 2001; Desort et al., 2007; Meunier et al., 2010). Stellar activity can impose a false RV of approximate magnitude 1 m s^{-1} for quiet stars (Hatzes, 2002; Lagrange et al., 2010; Isaacson and Fischer, 2010) to hundreds of m s^{-1} for the most active (Saar and Donahue, 1997; Paulson et al., 2004). While efforts have been made to model this activity (e.g., Tuomi et al. 2013; Rajpaul et al. 2015; 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 et al. 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) uses principal components analysis to show

that, at least according to simplified models of 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 et al. 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 non-zero (see Figure 5), therefore, the HGRV method would not work well. Fortunately, very few 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 orthogonalize the effects of stellar activity and a Doppler-shift. 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 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-prevision RMS that is on average approximately 10 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.

Acknowledgements. 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, non-profit 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.

References.

Anglada-Escudé, 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. *The Astrophysical Journal Supplement Series*, 200(2):15.

Astudillo-Defru, N., Bonfils, X., Delfosse, X., Ségransan, D., Forveille, T., Bouchy, F., Gillon, M., Lovis, C., Mayor, M., Neves, V., 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. *Astronomy & Astrophysics*, 575:A119.

Baranne, A., Queloz, D., Mayor, M., Adrianzyk, G., Knispel, G., Kohler, D., Lacroix, D., Meunier, J.-P., Rimbaud, G., and Vin, A. (1996). Elodie: A spectrograph for accurate radial velocity measurements. *Astronomy and Astrophysics Supplement Series*, 119(2):373–390.

Barker, P. (1984). Ripple correction of high-dispersion iue spectra-blazing echelles. *The Astronomical Journal*, 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. *arXiv preprint arXiv:1901.00503*.

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., Zhao, L. L., Leet, C., Buchhave, L. A., Tronsgaard,

R., Llama, J., Sawyer, T., Davis, A. B., Cabot, S. H. C., Shao, M., Trahan, R., Nemati, B., Genoni, M., Pariani, G., Riva, M., Probst, R. A., Holzwarth, R., Steinmetz, T., Fournier, P., and Pawluczyk, R. (2020). Performance Verification of the EXtreme PREcision Spectrograph. *arXiv e-prints*, page arXiv:2003.08852.

Blackman, R. T., Szymkowiak, A. E., Fischer, D. A., and Jurgenson, C. A. (2017). Accounting for chromatic atmospheric effects on barycentric corrections. *The Astrophysical Journal*, 837(1):18.

Bouchy, F., Pepe, F., and Queloz, D. (2001). Fundamental photon noise limit to radial velocity measurements. *Astronomy & Astrophysics*, 374(2):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 Journal on Scientific Computing*, 21(1):1–23.

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. *Publications of the Astronomical Society of the Pacific*, 108(724):500.

Ciuryło, R. (1998). Shapes of pressure-and doppler-broadened spectral lines in the core and near wings. *Physical Review A*, 58(2):1029.

Cretignier, M., Dumusque, X., Allart, R., Pepe, F., and Lovis, C. (2020). Measuring precise radial velocities on individual spectral lines-ii. dependance of stellar activity signal on line depth. *Astronomy & Astrophysics*, 633:A76.

Dai, C.-Q., Wang, Y., and Liu, J. (2016). Spatiotemporal hermite–gaussian solitons of a (3+ 1)-dimensional partially nonlocal nonlinear schrödinger equation. *Nonlinear Dynamics*, 84(3):1157–1161.

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. *The Astrophysical Journal*, 846(1):59.

Delisle, J.-B., Ségransan, D., Dumusque, X., Diaz, R., Bouchy, F., Lovis, C., Pepe, F., Udry, S., Alonso, R., Benz, W., et al. (2018). The harps search for southern extra-solar planets-xliii. a compact system of four super-earth planets orbiting hd 215152. *Astronomy & Astrophysics*, 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. *Astronomy & Astrophysics*, 473(3):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. *Astronomy & Astrophysics*, 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. *The Astrophysical Journal*, 796(2):132.

Dumusque, X., Borsa, F., Damasso, M., Díaz, R. F., Gregory, P., Hara, N., Hatzes, A., Rajpaul, V., Tuomi, M., Aigrain, S., et al. (2017). Radial-velocity fitting challenge-ii. first results of the analysis of the data set. *Astronomy & Astrophysics*, 598:A133.

Einstein, A. et al. (1905). On the electrodynamics of moving bodies. *Annalen der Physik*, 17(891):50.

Fischer, D. A., Anglada-Escude, G., Arriagada, P., Baluev, R. V., Bean, J. L., Bouchy, F., Buchhave, L. A., Carroll, T., Chakraborty, A., Crepp, J. R., et al. (2016). State of the field: extreme precision radial velocities. *Publications of the Astronomical Society of the Pacific*, 128(964):066001.

Frank, S. (2008). *OVI absorbers in SDSS spectra*. PhD thesis, The Ohio State University.

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. *The Astrophysical Journal*, 824(2):150.

Gray, D. F. (2005). *The observation and analysis of stellar photospheres*. Cambridge University Press.

Halverson, S., Terrien, R., Mahadevan, S., Roy, A., Bender, C., Stefánsson, G. K., Monson, A., Levi, E., Hearty, F., Blake, C., et al. (2016). A comprehensive radial velocity error budget for next generation doppler spectrometers. In *Ground-based and Airborne Instrumentation for Astronomy VI*, volume 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. , 126(943):827.

Hatzes, A. P. (2002). Starspots and exoplanets. *Astronomische Nachrichten*, 323(3-4):392–394.

Isaacson, H. and Fischer, D. (2010). Chromospheric activity and jitter measurements for 2630 stars on the california planet search. *The Astrophysical Journal*, 725(1):875.

Johnston, W. (2014). The weighted hermite polynomials form a basis for l 2 (). *The American Mathematical Monthly*, 121(3):249–253.

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*, volume 9908, page 99086T. International Society for Optics and Photonics.

Labutin, T. A., Zaytsev, S. M., and Popov, A. M. (2013). Automatic identification of emission lines in laser-induced plasma by correlation of model and experimental spectra. *Analytical chemistry*, 85(4):1985–1990.

Lagrange, A.-M., Desort, M., and Meunier, N. (2010). Using the sun to estimate earth-like planets detection capabilities-i. impact of cold spots. *Astronomy & Astrophysics*, 512:A38.

Lanczos, C. (1938). Trigonometric interpolation of empirical and analytical functions. *Journal of Mathematics and Physics*, 17(1-4):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. *arXiv preprint arXiv:1903.08350*.

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. *The Astrophysical Journal*, 481(2):926.

Marhic, M. (1978). Oscillating hermite-gaussian wave functions of the harmonic oscillator. *Lett. Nuovo Cim*, 22(9):376–378.

Mayor, M. and Queloz, D. (1995). A jupiter-mass companion to a solar-type star. *Nature*, 378(6555):355.

Mészáros, S. and Prieto, C. A. (2013). On the interpolation of model atmospheres and high-resolution synthetic stellar spectra. *Monthly Notices of the Royal Astronomical Society*, 430(4):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. *Astronomy & Astrophysics*, 512:A39.

Moré, J. J. (1978). The levenberg-marquardt algorithm: implementation and theory. In *Numerical analysis*, pages 105–116. Springer.

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. *The Astronomical Journal*, 127(6):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. *Astronomy & Astrophysics*, 388(2):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., Leet, C., McCracken, T. M., Sawyer, D., Sharov, M., Tronsgaard, R., Szymkowiak, A. E., and Fischer, D. A. (2020). An Extreme Precision Radial Velocity Pipeline: First Radial Velocities from EXPRES. *arXiv e-prints*, page arXiv:2003.08851.

Planck, M. (1901). On the law of distribution of energy in the normal spectrum. *Annalen der physik*, 4(553):1.

Queloz, D., Henry, G., Sivan, J., Baliunas, S., Beuzit, J., Donahue, R., Mayor, M., Naef, D., Perrier, C., and Udry, S. (2001). No planet for hd 166435. *Astronomy & Astrophysics*, 379(1):279–287.

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. *Monthly Notices of the Royal Astronomical Society*, 452(3):2269–2291.

Rajpaul, V. M., Aigrain, S., and Buchhave, L. A. (2020). A robust, template-free approach to precise radial velocity extraction. *Monthly Notices of the Royal Astronomical Society*, 492(3):3960–3983.

Riffel, R. A. (2010). profit: a new alternative for emission-line pro file fit ting. *Astrophysics and Space Science*, 327(2):239–244.

Rimmele, T. R. and Radick, R. R. (1998). Solar adaptive optics at the national solar observatory. In *Adaptive Optical System Technologies*, volume 3353, pages 72–82. International Society for Optics and Photonics.

Saar, S. H. and Donahue, R. A. (1997). Activity-related radial velocity variation in cool stars. *The Astrophysical Journal*, 485(1):319.

Sharpee, B., Williams, R., Baldwin, J. A., and van Hoof, P. A. (2003). Introducing emili: Computer-aided emission line identification. *The Astrophysical Journal Supplement Series*, 149(1):157.

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. *Astronomy & Astrophysics*, 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*.

Tonegawa, M., Totani, T., Iwamuro, F., Akiyama, M., Dalton, G., Glazebrook, K., Ohta, K., Okada, H., and Yabe, K. (2015). Field: Automated emission line detection software for subaru/fmos near-infrared spectroscopy. *Publications of the Astronomical Society of Japan*, 67(3).

Tuomi, M., Anglada-Escudé, 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. *Astronomy & Astrophysics*, 549:A48.

Wang, J. and Ford, E. B. (2011). On the eccentricity distribution of short-period single-planet systems. *Monthly Notices of the Royal Astronomical Society*, 418(3):1822–1833.

Wright, J. and Eastman, J. (2014). Barycentric corrections at 1 cm s-1 for precise doppler velocities. *Publications of the Astronomical Society of the Pacific*, 126(943):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. *The Astronomical Journal*, 157(6):243.

Zhao, Y., Ge, J., Yuan, X., Zhao, T., Wang, C., and Li, X. (2019). Identifying mg ii narrow absorption lines with deep learning. *Monthly Notices of the Royal Astronomical Society*, 487(1):801–811.

APPENDIX A: DETAILS OF ABSORPTION FEATURE FINDER ALGORITHM

Various algorithms already exist for detecting spectral features, particularly for emission lines in spectra of galaxies. However, they contain some limitations that make them unsuitable for the proposed methodology. For example, some were developed for absorption features of specific elemental species or line types[11] (Frank, 2008; Zhao et al., 2019), require experimental supervision (Labutin et al., 2013), partially consist of extensive human intervention and physical insight (Sharpee et al., 2003), or assume the features are sparse and well-separated (Tonegawa et al., 2015).

More importantly, these algorithms lack an important component needed for our analysis: estimating not just the central wavelength at which the feature occurs, but also the wavelength bounds that contain the feature. Dumusque (2018) approaches this by taking a fixed number of pixels around each feature center, but acknowledges that these windows could be further optimized. The reason for this is because a fixed pixel count for each wavelength window does not take into account different sizes of absorption features nor blends between neighboring features. Cretignier et al. (2020) improves upon this by allowing the number

[11]The central wavelength of each spectral line corresponds to a particular electron state transition of atoms responsible for absorbing photons in the stars photosphere. These central wavelengths depend on the species of the absorbing atom and its ionization state.

of pixels to vary for each feature but, by restricting the windows to be symmetric about the minimum, does not account for effects of line blends. Our proposed algorithm improves upon this by using an approach that accounts for these blends.

Our proposed Algorithm 1 works as follows. For a given pixel index i, let $\Lambda_{l,i}$ and $\Lambda_{r,i}$ be the wavelength regions of size m pixels to the left and right of the wavelength for pixel i, x_i, respectively. Also, let $Y_{l,i}$ and $Y_{r,i}$ be the corresponding flux regions. Algorithm 1 uses least-squares regression on each region to estimate coefficients $\beta_{0,l}$ and $\beta_{1,l}$ for the left region in addition to $\beta_{0,r}$ and $\beta_{1,r}$ for the right region (see Algorithm 1 for the model). If $\beta_{1,l}$ is found to be negative and $\beta_{1,r}$ positive with statistical significance, then x_i is considered a statistically significant minimum. At this point we apply a Bonferroni correction by using the significance level $\alpha/2$ for each slope. Algorithm 1 then proceeds outwards in wavelength until the estimates are no longer statistically significant, at which point the central wavelength of the window is taken as a feature bound. To further avoid the drawbacks of multiple testing, we eliminate any detected absorption features that do not have a depth above a certain threshold. We note, however, that multiple testing is not a concern since our goal is to find absorption features, and we do not use the statistical significance beyond the detection of the features.

It was found that when m is too small, many false absorption features are detected. When m is too large, many small features are missed. Even though similar effects come from α and η being too large or small, the effects appeared more sensitive to m. For fixed values of α and η, we adjusted m until the number of detected features was maximized. At this point we increased or decreased α if many small features were missed or many false features were detected. If many blended features were detected as single features or the wavelength bounds did not encompass full absorption features, we decreased or increased η, respectively, and repeated the full process.

When applying Algorithm 1 to the NSO spectrum, we get the results shown in Figures 16 and 17. Figure 16 displays the portion of the spectrum that was not contained in any detected absorption features and compares it to the full spectrum. Figure 17 displays some examples of absorption features that were missed by the algorithm. These figures illustrate that 97.7% of the squared deviation from 1.0 in the normalized flux is accounted for by the 64.4% of the spectrum contained in the wavelength bounds given by the algorithm.

It is also noticeable that some absorption features are missed by the algorithm, some of which are deep. Most of these were missed because, as illustrated in Figure 17, the features are strongly blended in a way that makes the slope in either direction at the core statistically insignificant. There are likely ways to improve upon this aspect of the algorithm, and we leave this to future work.

To analyze how the minimum line depth parameter depends on the S/N of the spectrum, we extend the false positive rate simulation done with a S/N of 500 described in Section 2.

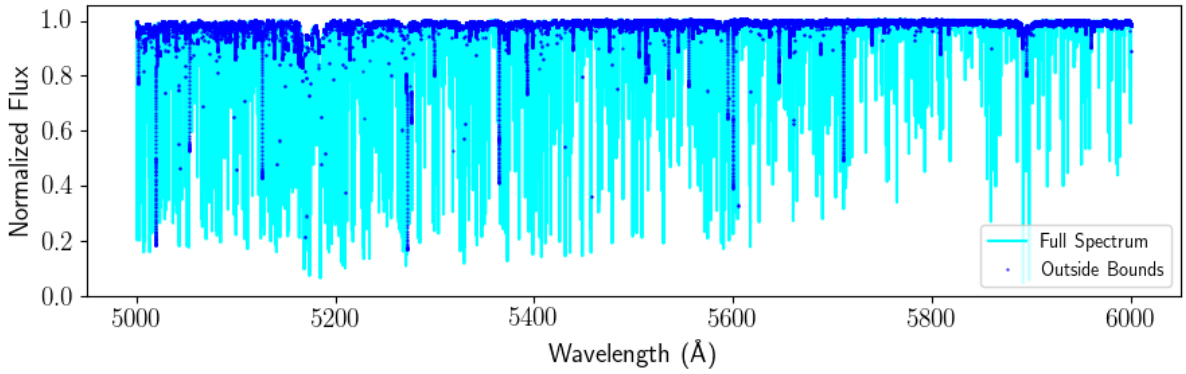

FIGURE 16. *The full NSO spectrum used in testing Algorithm 1. Normalized flux is plotted against the wavelength. The full spectrum is plotted in light blue. The thick dark blue points indicate the portions of the spectrum that are not contained in any of the wavelength bounds given by the algorithm.*

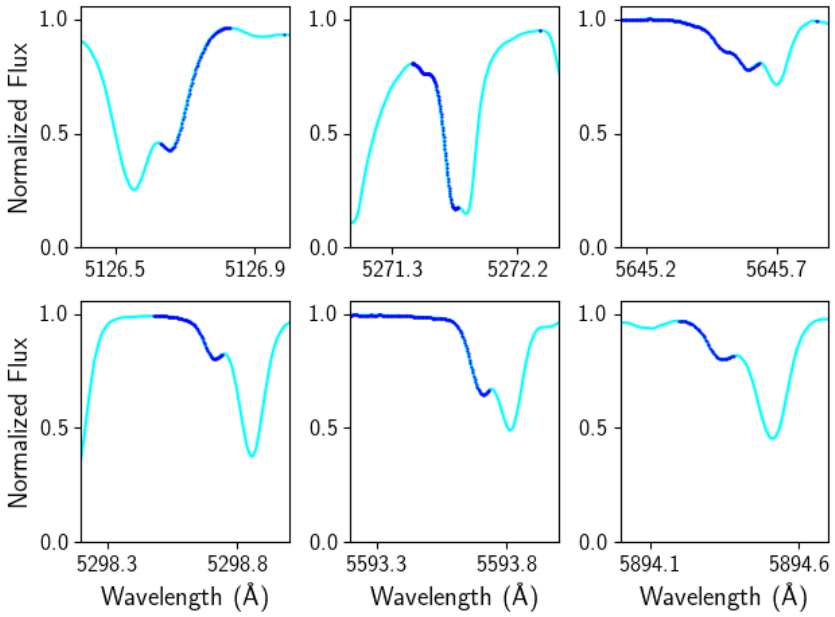

FIGURE 17. *Six of the absorption features in the NSO that were missed by Algorithm 1. Normalized flux is plotted against wavelength. The full spectrum is shown in light blue, and portions not included in any of the wavelength bounds given by the algorithm is shown in dark blue.*

For each S/N from 250 to 1500 in equal steps of 250, we take the NSO spectrum between 5000 and 6000 Å and replace the flux axis with noise 20 independent times. We then apply

Algorithm 1 to each of the 20 resulting spectra with parameters $m = 25$, $\alpha = 0.01$, and $\eta = 0.05$. We then collect all detected absorption features from the 20 spectra.

The total count of false absorption features detected ranged from 51 to 56 and showed no association with the S/N level. Furthermore, the depth of these false features is illustrated in Figure 18. The recommended minimum line depth parameter, $0.015 \times \frac{500}{\mathrm{S/N}}$, is also shown.

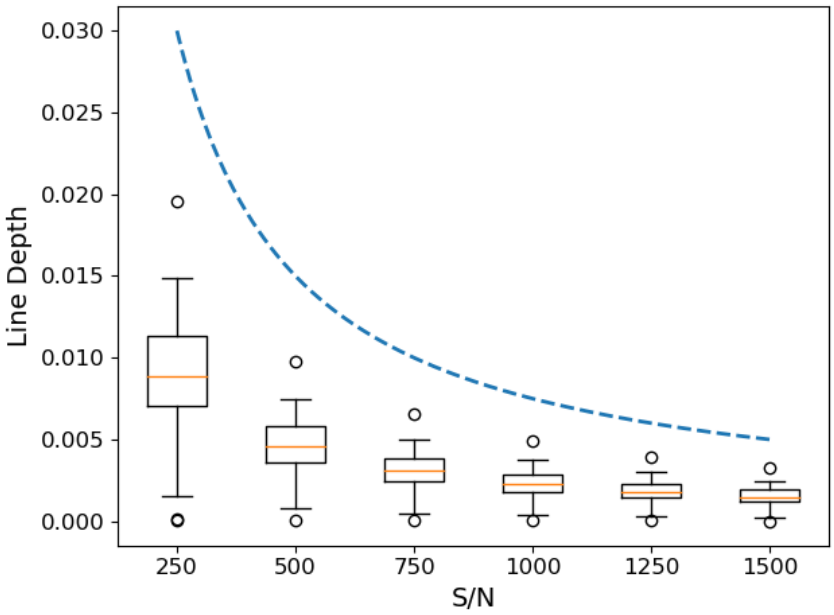

FIGURE 18. *Results from our simulation of the false positive rate of Algorithm 1 at various S/N, shown on the horizontal axis. The distribution of line depths for these false positives is represented by box plots according to the vertical axis. The count of false positives remained approximately constant at* 1 *absorption feature per* 363 *Å for each S/N. The dashed line represents our recommended value for the minimum line depth parameter in the algorithm given by the expression* $0.015 \times \frac{500}{\mathrm{S/N}}$.

APPENDIX B: PROOFS OF LEMMAS 1-4

PROOF. (of Lemma 1)
Choose constants $a \in \mathbb{R}^+$, $b, c \in \mathbb{R}$. Then, using integration by parts, we have that

$$I_1(a,b,c) = e^{\left(\frac{b^2}{4a} - c\right)} \left[\int_{-\infty}^{\infty} u e^{-au^2} du - \frac{b}{2a} \int_{-\infty}^{\infty} e^{-au^2} du \right] = -\frac{\sqrt{\pi} b}{2a^{3/2}} e^{\left(\frac{b^2}{4a} - c\right)} \quad . \tag{24}$$

$$I_0(a,b,c) = \int_{-\infty}^{\infty} e^{-\left(ax^2+bx+c\right)} dx = \frac{2a}{b} I_1(a,b,c) = \sqrt{\frac{\pi}{a}} e^{\left(\frac{b^2}{4a} - c\right)} \quad . \tag{25}$$

Now choose any $k \in \{n \in \mathbb{N} : n \geq 2\}$.

$$I_{k-1}(a,b,c) = \int_{-\infty}^{\infty} x^{k-1} e^{-ax^2} e^{-(bx+c)} dx \tag{26}$$

$$= \lim_{z \to \infty} \left[-\frac{1}{b} x^{k-1} e^{-\left(ax^2+bx+c\right)} |_{-z}^{z} \right] + \frac{1}{b} \int_{-\infty}^{\infty} \left((k-1)x^{k-2} - 2ax^k \right) e^{-\left(ax^2+bx+c\right)} dx \tag{27}$$

$$= \frac{k-1}{b} I_{k-2}(a,b,c) - \frac{2a}{b} I_k(a,b,c) \quad . \tag{28}$$

So we have that

$$I_k(a,b,c) = -\frac{b}{2a} I_{k-1}(a,b,c) + \frac{k-1}{2a} I_{k-2}(a,b,c). \tag{29}$$

□

PROOF. (of Lemma 2)
Since $g(x;\gamma) = \sum_{n=0}^{\infty} c_n(\gamma) \psi_n(x;\mu,\sigma)$ and $\psi_n(x;\mu,\sigma)$ are orthonormal, we have that

$$c_k(\gamma) = \int_{-\infty}^{\infty} \psi_k(x;\mu,\sigma) g(x;\gamma) dx \ . \tag{30}$$

Choose any $k \in \{n \in \mathbb{N} : n \geq 1\}$. By using Equation (4) for the k'th Hermite polynomial,

we have that for $\varepsilon = \gamma - 1$,

$$c_k(\varepsilon) = \int_{-\infty}^{\infty} g(x;\varepsilon)\psi_k(x;\mu,\sigma)dx \tag{31}$$

$$= \sqrt{\sigma\sqrt{\pi}} \int_{-\infty}^{\infty} \psi_0(x;\mu,\sigma)\psi_k(x;\mu,\sigma)dx - \int_{-\infty}^{\infty} e^{-\frac{1}{2\sigma^2}(x-\mu+\varepsilon x)^2} \psi_k(x;\mu,\sigma)dx \tag{32}$$

$$= 0 - \int_{-\infty}^{\infty} \frac{1}{\sqrt{\sigma 2^k k!\sqrt{\pi}}} H_k\left(\frac{x-\mu}{\sigma}\right) e^{-\frac{1}{2\sigma^2}\left[(x-\mu+\varepsilon x)^2+(x-\mu)^2\right]} dx \tag{33}$$

$$= -\frac{\sqrt{\sigma}}{\sqrt{2^k k!\sqrt{\pi}}} \int_{-\infty}^{\infty} H_k(u) e^{-\frac{1}{2}\left[\left(u+\varepsilon\left(u+\frac{\mu}{\sigma}\right)\right)^2+u^2\right]} du \tag{34}$$

$$= -\frac{\sqrt{\sigma}}{\sqrt{2^k k!\sqrt{\pi}}} \int_{-\infty}^{\infty} k! \sum_{m=0}^{\lfloor k/2 \rfloor} \frac{(-1)^m}{m!(k-2m)!} (2u)^{k-2m} e^{-\frac{1}{2}\left[\left(2+2\varepsilon+\varepsilon^2\right)u^2+2\varepsilon\frac{\mu}{\sigma}(1+\varepsilon)u+\varepsilon^2\frac{\mu^2}{\sigma^2}\right]} du \tag{35}$$

$$= -\frac{\sqrt{\sigma k! 2^k}}{\sqrt{\sqrt{\pi}}} \sum_{m=0}^{\lfloor k/2 \rfloor} \frac{(-1)^m}{m!(k-2m)!} \frac{1}{4^m} \int_{-\infty}^{\infty} u^{k-2m} e^{-\frac{1}{2}\left[\left(2+2\varepsilon+\varepsilon^2\right)u^2+2\varepsilon\frac{\mu}{\sigma}(1+\varepsilon)u+\varepsilon^2\frac{\mu^2}{\sigma^2}\right]} du \tag{36}$$

(37)

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

For $k = 0$, the only difference is that the first integral in Equation (32) becomes 1 instead of vanishing. Therefore,

$$c_0(\varepsilon) = \sqrt{\sigma\sqrt{\pi}} - \frac{1}{\sqrt{\sigma\sqrt{\pi}}} I_0\left(\frac{1+\varepsilon+\frac{\varepsilon^2}{2}}{\sigma^2}, -\frac{2\mu+\varepsilon\mu}{\sigma^2}, \left(\frac{\mu}{\sigma}\right)^2\right) \tag{38}$$

□

Proof. (of Lemma 3)
Decompose as $g(x;\gamma) = \sum_{n=0}^{\infty} c_n(\gamma)\psi_n(x;\mu,\sigma)$.

Then

$$\int_{-\infty}^{\infty} \left(g(x;\gamma) - c_1(\gamma)\psi_1(x;\mu,\sigma)\right)^2 dx \tag{39}$$

$$= \int_{-\infty}^{\infty} (g(x;\gamma))^2 dx - 2c_1(\gamma)\int_{-\infty}^{\infty} g(x;\gamma)\psi_1(x;\mu,\sigma)dx + c_1^2(\gamma)\int_{-\infty}^{\infty} \psi_1^2(x;\mu,\sigma)dx \tag{40}$$

$$= \int_{-\infty}^{\infty} (g(x;\gamma))^2 dx - 2c_1^2(\gamma) + c_1^2(\gamma) = \int_{-\infty}^{\infty} (g(x;\gamma))^2 dx - c_1^2(\gamma) \tag{41}$$

□

Proof. (of Lemma 4)
From Lemmas 1 and 2 we have that, with $\varepsilon = \gamma - 1$,

$$c_1^2(\varepsilon) = \varepsilon^2(1+\varepsilon)^2 h(\varepsilon) \tag{42}$$

where

$$h(\varepsilon) := \frac{4\sqrt{\pi}\mu^2}{\sigma} \frac{1}{\left(2+2\varepsilon+\varepsilon^2\right)^3} e^{-\left(\frac{\mu}{\sigma}\right)^2 \frac{\varepsilon^2}{2+2\varepsilon+\varepsilon^2}} . \tag{43}$$

We also have that

$$\frac{\partial}{\partial\varepsilon} c_1^2(\varepsilon) = \left(4\varepsilon^3 + 6\varepsilon^2 + 2\varepsilon\right) h(\varepsilon) + \varepsilon^2(1+\varepsilon)^2 \frac{\partial h(\varepsilon)}{\partial\varepsilon} \tag{44}$$

and

$$\frac{\partial^2}{\partial\varepsilon^2} c_1^2(\varepsilon) = \left(12\varepsilon^2 + 12\varepsilon + 2\right) h(\varepsilon) + 2\left(4\varepsilon^3 + 6\varepsilon^2 + 2\varepsilon\right) \frac{\partial h}{\partial\varepsilon} + \varepsilon^2(1+\varepsilon)^2 \frac{\partial^2 h}{\partial\varepsilon^2} . \tag{45}$$

Since $h(\varepsilon)$, $\frac{\partial h}{\partial\varepsilon}$, and $\frac{\partial^2 h}{\partial\varepsilon^2}$ are all continuous at 0, we have that

$$\lim_{\varepsilon\to 0} c_1^2(\varepsilon) = 0, \tag{46}$$

$$\lim_{\varepsilon\to 0} \frac{\partial}{\partial\varepsilon} c_1^2(\varepsilon) = 0, \tag{47}$$

$$\text{and } \lim_{\varepsilon\to 0} \frac{\partial^2}{\partial\varepsilon^2} c_1^2(\varepsilon) = 2\lim_{\varepsilon\to 0} h(\varepsilon) = \frac{\sqrt{\pi}\mu^2}{\sigma}. \tag{48}$$

With $g(x;\mu,\sigma)$ as in Lemma 2, we have that

(49)
$$\int_{-\infty}^{\infty} g^2(x;\gamma)dx = \int_{-\infty}^{\infty} e^{-\frac{(x-\mu)^2}{\sigma^2}}dx + \int_{-\infty}^{\infty} e^{-\frac{(\gamma x-\mu)^2}{\sigma^2}}dx - 2\int_{-\infty}^{\infty} e^{-\frac{1}{2\sigma^2}\left((1+\gamma^2)x^2-2\mu(1+\gamma)x+2\mu^2\right)}dx$$

(50)
$$= \sigma\sqrt{\pi} + \frac{\sigma}{\gamma}\sqrt{\pi} - 2e^{-\frac{\mu^2}{\sigma^2}\left(1-\frac{(1+\gamma)^2}{2(1+\gamma^2)}\right)}\int_{-\infty}^{\infty} e^{-\frac{1+\gamma^2}{2\sigma^2}\left(x-\frac{\mu(1+\gamma)}{1+\gamma^2}\right)^2}dx$$

(51)
$$= \sigma\sqrt{\pi}\left(1+\frac{1}{\gamma}-\frac{2^{3/2}}{\sqrt{1+\gamma^2}}e^{-\frac{\mu^2}{\sigma^2}\left(1-\frac{(1+\gamma)^2}{2(1+\gamma^2)}\right)}\right)$$

(52)
$$= \sigma\sqrt{\pi}\left(1+\frac{1}{1+\varepsilon}-\frac{2^{3/2}}{\sqrt{2+2\varepsilon+\varepsilon^2}}e^{-\frac{\mu^2}{2\sigma^2}\frac{\varepsilon^2}{2+2\varepsilon+\varepsilon^2}}\right)$$

Therefore, $\lim_{\varepsilon\to 0}\int_{-\infty}^{\infty} g^2(x;\varepsilon)dx = 0$. Furthermore, we have that

(53)
$$\frac{\partial}{\partial\varepsilon}\int_{-\infty}^{\infty} g^2(x;\varepsilon)dx = \sigma\sqrt{\pi}\left[-\frac{1}{(1+\varepsilon)^2}+2^{3/2}\left(\left(2+2\varepsilon+\varepsilon^2\right)^{-3/2}(1+\varepsilon)\right.\right.$$
$$\left.\left.+\frac{\mu^2}{\sigma^2}\left(2+2\varepsilon+\varepsilon^2\right)^{-5/2}\left(2\varepsilon+\varepsilon^2\right)\right)e^{-\frac{\mu^2}{2\sigma^2}\frac{\varepsilon^2}{2+2\varepsilon+\varepsilon^2}}\right].$$

Hence,

(54)
$$\lim_{\varepsilon\to 0}\frac{\partial}{\partial\varepsilon}\int_{-\infty}^{\infty} g^2(x;\varepsilon)dx = 0.$$

Defining

(55)
$$h(\varepsilon) := 2^{3/2}\left(\left(2+2\varepsilon+\varepsilon^2\right)^{-3/2}(1+\varepsilon)+\frac{\mu^2}{\sigma^2}\left(2+2\varepsilon+\varepsilon^2\right)^{-5/2}\left(2\varepsilon+\varepsilon^2\right)\right),$$

we have that $h(\varepsilon)$ is continuous and differentiable at 0.

Therefore, since

$$\lim_{\varepsilon\to 0} e^{-\frac{\mu^2}{2\sigma^2}\frac{\varepsilon^2}{2+2\varepsilon+\varepsilon^2}} = 1 \tag{56}$$

and

$$\lim_{\varepsilon\to 0} \frac{\partial}{\partial\varepsilon} e^{-\frac{\mu^2}{2\sigma^2}\frac{\varepsilon^2}{2+2\varepsilon+\varepsilon^2}} = 0, \tag{57}$$

we have that

$$\lim_{\varepsilon\to 0} \frac{\partial}{\partial\varepsilon} \left(h(\varepsilon) e^{-\frac{\mu^2}{2\sigma^2}\frac{\varepsilon^2}{2+2\varepsilon+\varepsilon^2}} \right) = \lim_{\varepsilon\to 0} \frac{\partial h(\varepsilon)}{\partial\varepsilon}. \tag{58}$$

Since

$$\lim_{\varepsilon\to 0} \frac{\partial}{\partial\varepsilon} \left(\left(2+2\varepsilon+\varepsilon^2\right)^{-3/2} (1+\varepsilon) \right) = -3\cdot 2^{-5/2} + 2^{-3/2} \tag{59}$$

and

$$\lim_{\varepsilon\to 0} \frac{\partial}{\partial\varepsilon} \left(\left(2+2\varepsilon+\varepsilon^2\right)^{-5/2} \left(2\varepsilon+\varepsilon^2\right) \right) = 2^{-3/2}, \tag{60}$$

we have that

$$\lim_{\varepsilon\to 0} \frac{\partial h(\varepsilon)}{\partial\varepsilon} = \frac{\mu^2}{\sigma^2} - \frac{1}{2}. \tag{61}$$

And since

$$\lim_{\varepsilon\to 0} \frac{\partial}{\partial\varepsilon} \left(\frac{-1}{(1+\varepsilon)^2} \right) = 2, \tag{62}$$

we have from Equation (53) that

$$\lim_{\varepsilon\to 0} \frac{\partial^2}{\partial\varepsilon^2} \int_{-\infty}^{\infty} g^2(x;\varepsilon)dx = \frac{3\sigma\sqrt{\pi}}{2} + \frac{\sqrt{\pi}\mu^2}{\sigma}\ . \tag{63}$$

So

$$\lim_{\varepsilon\to 0} \frac{c_1^2(\varepsilon)}{\int_{-\infty}^{\infty} g^2(x;\varepsilon)dx} = \frac{\frac{\sqrt{\pi}\mu^2}{\sigma}}{\frac{\sqrt{\pi}\mu^2}{\sigma} + \frac{3\sigma\sqrt{\pi}}{2}} = \frac{1}{1+\frac{3\sigma^2}{2\mu^2}}. \tag{64}$$

□

APPENDIX C: MODEL MISSPECIFICATION DETAILS

Following the same procedure as in Section 3.4, we considered 100 additional absorption features to analyze the effect of misspecifying their profile as Gaussian, five of which are displayed below in Figure 19. A Gaussian density shape is fit to each absorption feature, which is then Doppler-shifted by 50 equally spaced values of RV from 1 to 100 m s^{-1}. The RV is then estimated using the HGRV method. Most, but not all, of the additional features we analyzed lead to a slight overestimate of the RV. But for all 100 of these additional features, the difference for a 1 m s^{-1} RV is less than 1 cm s^{-1} away from the truth. Furthermore, the simulations in Section 4.2 indicate that when combining the lines in the HGRV method, the overall bias is not greater than with individual lines.

APPENDIX D: 51 PEGASI RADIAL VELOCITIES

Here we give the RVs derived using the HGRV method on the 56 observed spectra from EXPRES.

MJD (days)	**RV (m s^{-1})**	**S/N**	**In Petersburg et al. (2020)**
58639.458442	54.708 ± 0.404	385	Yes
58641.451749	−52.850 ± 0.516	179	Yes
58641.457773	−53.662 ± 0.710	140	No ∗
58643.462180	46.574 ± 0.521	225	Yes
58644.460959	33.536 ± 0.512	233	Yes
58646.455970	−43.411 ± 0.444	203	Yes
58646.461286	−42.241 ± 0.438	204	Yes
58648.456163	48.082 ± 0.505	244	Yes
58648.461529	48.711 ± 0.498	256	Yes
58650.450235	−53.092 ± 0.474	199	Yes
58650.455542	−53.741 ± 0.486	193	Yes
58651.443961	14.317 ± 1.130	99	No ∗
58651.452932	16.403 ± 0.431	284	Yes
58651.461117	16.515 ± 0.519	202	Yes
58652.456394	53.336 ± 0.653	172	Yes
58652.461797	52.696 ± 0.667		No
58655.432426	−4.268 ± 0.459	220	Yes
58655.437704	−4.148 ± 0.453	222	Yes
58657.456051	11.484 ± 0.706	142	No ∗
58657.461248	11.792 ± 0.623	157	No ∗
58658.453711	−54.614 ± 0.361	230	Yes
58658.456675	−53.959 ± 0.343	247	Yes
58658.459600	−53.690 ± 0.341	243	Yes
58658.462634	−54.958 ± 0.330	257	Yes
58658.465250	−54.357 ± 0.350	236	Yes
58664.447934	36.335 ± 1.171	94	No ∗
58664.458268	35.866 ± 1.270	89	No ∗
58665.461782	43.919 ± 0.538	214	No †
58749.221866	47.001 ± 0.684	163	Yes
58749.227235	48.181 ± 0.688		No
58763.233618	13.399 ± 0.466	230	Yes
58763.239194	12.112 ± 0.455	238	Yes
58764.311548	−53.651 ± 0.401	244	Yes

Continued on next page

Table 2 – continued from previous page

MJD (days)	RV (m s^{-1})	S/N	In Petersburg et al. (2020)
58764.318051	−53.765 ± 0.372	273	Yes
58772.315903	−36.680 ± 0.414	234	Yes
58772.321086	−37.061 ± 0.420	231	Yes
58780.114819	16.529 ± 0.464	237	Yes
58780.121270	15.728 ± 0.462	238	Yes
58787.198050	43.659 ± 0.557	194	Yes
58787.206110	44.981 ± 0.504	226	Yes
58796.099263	58.195 ± 0.546	235	Yes
58796.102083	57.456 ± 0.544	236	Yes
58796.104824	58.356 ± 0.546	235	Yes
58796.107532	57.717 ± 0.543	235	Yes
58798.128178	−52.396 ± 0.412	234	Yes
58798.129893	−55.148 ± 0.411	233	Yes
58798.131622	−53.502 ± 0.411	231	Yes
58798.133471	−51.899 ± 0.409	232	Yes
58803.110815	−34.492 ± 0.418	233	Yes
58803.114000	−33.286 ± 0.418	233	Yes
58803.116558	−34.086 ± 0.416	233	Yes
58803.118928	−32.252 ± 0.413	234	Yes
58804.076698	39.312 ± 0.503	239	Yes
58804.080907	40.058 ± 0.502	239	Yes
58804.084687	40.916 ± 0.504	240	Yes
58804.088298	41.212 ± 0.503	239	Yes

Table 2: Radial velocities derived from the HGRV method for 51 Pegasi. The first column gives the Modified Julian Day (MJD) which can be converted to JD by adding 2400000.5 days. The second column gives the estimated RV with its standard error. The third column identifies the S/N, and the fourth states whether it was included in the comparison of orbital parameters with Petersburg et al. (2020). ∗ indicates that it was not included in Petersburg et al. (2020) due to a S/N below 160. † indicates that the laser-frequency comb of the EXPRES spectrograph failed. A machine-readable version of this table is available on the online repository for this paper.

The EXPRES spectra used to obtain these estimated RV's with the HGRV method came with the barycentric corrected wavelength solutions provided which we used. All 56 were used in estimating the template spectrum for 51 Pegasi, and used the same set of identified absorption features and Gaussian fits to this estimating template.

E-MAIL: parker.holzer@yale.edu; jessica.cisewski@yale.edu; debra.fischer@yale.edu; lily.zhao@yale.edu

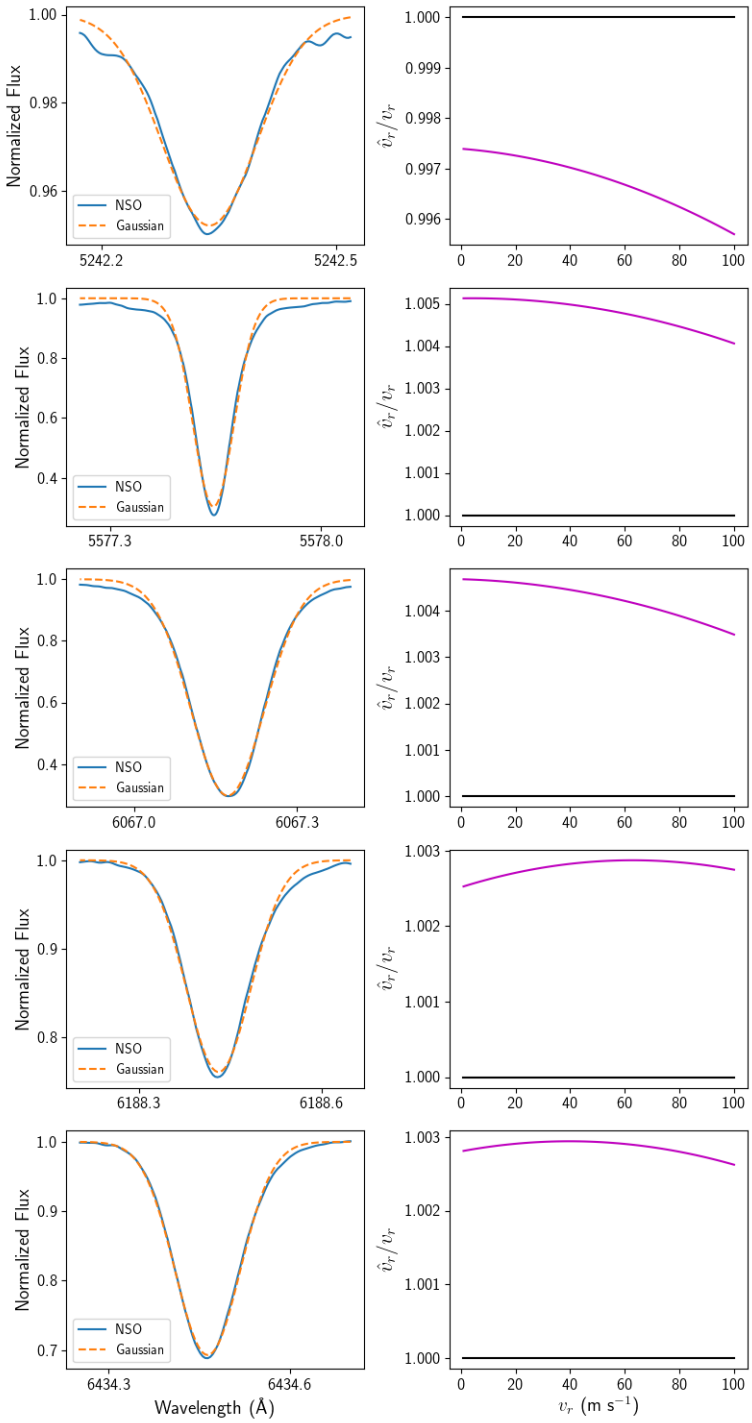

FIGURE 19. *Results for analyzing the effects of misspecifying the model of five different absorption features in the NSO spectrum as a Gaussian. The left panels show the feature in solid blue and the best fitted Gaussian in dashed orange. The right panels show the ratio of the RV estimated with Equation (18) $\hat{v}_r$ (with $n = 1$) and the true RV v_r.*