



# Radial Velocity Measurements of HR 8799 b and c with Medium Resolution Spectroscopy

Jean-Baptiste Ruffio<sup>1</sup> , Bruce Macintosh<sup>1</sup> , Quinn M. Konopacky<sup>2</sup> , Travis Barman<sup>3</sup> , Robert J. De Rosa<sup>1</sup> , Jason J. Wang<sup>4,8</sup> , Kielan K. Wilcomb<sup>2</sup> , Ian Czekala<sup>5,9</sup> , and Christian Marois<sup>6,7</sup>

<sup>1</sup> Kavli Institute for Particle Astrophysics and Cosmology, Stanford University, Stanford, CA 94305, USA; [jruffio@stanford.edu](mailto:jruffio@stanford.edu)

<sup>2</sup> Center for Astrophysics and Space Science, University of California, San Diego; La Jolla, CA 92093, USA

<sup>3</sup> Lunar and Planetary Laboratory, University of Arizona, Tucson, AZ 85721, USA

<sup>4</sup> Department of Astronomy, California Institute of Technology, Pasadena, CA 91125, USA

<sup>5</sup> Astronomy Department, University of California, Berkeley; Berkeley, CA 94720, USA

<sup>6</sup> National Research Council of Canada Herzberg, 5071 West Saanich Road, Victoria, BC, V9E 2E7, Canada

<sup>7</sup> University of Victoria, 3800 Finnerty Road, Victoria, BC, V8P 5C2, Canada

Received 2019 August 23; revised 2019 September 14; accepted 2019 September 16; published 2019 October 24

## Abstract

High-contrast medium resolution spectroscopy has been used to detect molecules such as water and carbon monoxide in the atmospheres of gas giant exoplanets. In this work, we show how it can be used to derive radial velocity (RV) measurements of directly imaged exoplanets. Improving upon the traditional cross-correlation technique, we develop a new likelihood based on joint forward modeling of the planetary signal and the starlight background (i.e., speckles). After marginalizing over the starlight model, we infer the barycentric RV of HR 8799 b and c in 2010 yielding  $-9.2 \pm 0.5 \text{ km s}^{-1}$  and  $-11.6 \pm 0.5 \text{ km s}^{-1}$ , respectively. These RV measurements help to constrain the 3D orientation of the orbit of the planet by resolving the degeneracy in the longitude of an ascending node. Assuming coplanar orbits for HR 8799 b and c, but not including d and e, we estimate  $\Omega = 89^{\circ}_{-17}^{+27}$  and  $i = 20^{\circ}_{-3.7}^{+4.5}$ .

*Unified Astronomy Thesaurus concepts:* Astrometry (80); Radial velocity (1332); Bayesian statistics (1900); Extrasolar gas giants (509); Direct imaging (387); High resolution spectroscopy (2096)

## 1. Introduction

The four planets orbiting the star HR 8799 have been poster children of the direct detection of exoplanets since their discovery (Marois et al. 2008, 2010). They orbit around their star at semimajor axes between 15 and 70 au for a system  $\approx 41$  pc (Gaia Collaboration 2018) away from the Sun. Their masses lie between 5 and  $10 M_{\text{Jup}}$ , which are derived from their luminosity, the estimated age of the system ( $\approx 41$  Myr; Zuckerman et al. 2011; Bell et al. 2015), and evolutionary models (Baraffe et al. 2003). The planets are surrounded by warm ( $< 10$  au) and cold ( $> 100$  au) dust belts, similar to the asteroid and the Kuiper belt in our own solar system (Reidemeister et al. 2009; Su et al. 2009; Hughes et al. 2011; Matthews et al. 2014; Booth et al. 2016).

The orbital parameters of the four planets have been constrained by a decade's worth of monitoring (Fabrycky & Murray-Clay 2010; Soummer et al. 2011; Pueyo et al. 2015; Konopacky et al. 2016; Zurlo et al. 2016; Wertz et al. 2017; Wang et al. 2018b; O'Neil et al. 2019). The planets are thought to only slightly deviate from coplanarity, with a small inclination around  $27^{\circ}$ , and close to 1:2:4:8 resonance (Wang et al. 2018). In a recent milestone, Gravity Collaboration et al. (2019) used optical interferometry to measure  $100 \mu\text{as}$  astrometry of the closest planet in the system (HR 8799 e). Unfortunately, direct imaging data does not distinguish between in-the-plane or out-the-plane of the sky motion of the planets, leading in particular to a  $180^{\circ}$  degeneracy in the longitude of the ascending node. Measuring the radial velocity (RV) of the planets can provide important constraints to the 3D

orientation of the orbits, but this requires medium ( $R = \lambda/d\lambda \sim 4000$ ) to high ( $R > 25,000$ ) resolution spectra (HRS). HR 8799 b and c have already been studied at medium resolution with Keck/OSIRIS ( $R \approx 4000$ ; Larkin et al. 2006), for example providing unambiguous detection of water ( $\text{H}_2\text{O}$ ) and carbon monoxide (CO) (Barman et al. 2011; Konopacky et al. 2013; Barman et al. 2015; Petit dit de la Roche et al. 2018). However, this data was never used to estimate the RV of the planets.

High resolution spectroscopy of exoplanets has so far only been possible for two classes of exoplanets, hot Jupiters and widely separated super-Jupiters. The strong and rapidly varying RVs of hot Jupiters can be used to isolate the planetary signal from one of their host stars, with which they are blended, as well as distinguishing them from telluric lines. The Cryogenic High-Resolution Infrared Echelle Spectrograph (CRIRES;  $R = 100,000$ ) at the Very Large Telescope (VLT; Kaeuff et al. 2004) has been a prolific instrument in this field, characterizing both transiting (e.g., Snellen et al. 2010) and nontransiting hot Jupiters (e.g., Brogi et al. 2012). Such measurements have enabled the detection of atomic lines (Hoeijmakers et al. 2018; Salz et al. 2018) and molecular lines (e.g., water,  $\text{H}_2\text{O}$ , and carbon monoxide, CO; Snellen et al. 2010; Brogi et al. 2012, 2013, 2014, 2016; Rodler et al. 2012; Birkby et al. 2013, 2017; de Kok et al. 2013; Lockwood et al. 2014). HRS has also probed day side to night side winds, and the possibility of thermal inversion layers (Brogi et al. 2012, 2017; Schwarz et al. 2015; Nugroho et al. 2017).

Current high-resolution spectrographs are not designed for high-contrast observations and are therefore limited to widely separated and bright directly imaged exoplanets. The sensitivity of high-contrast imaging is limited by the diffracted starlight at

<sup>8</sup> 51 Pegasi b Fellow.

<sup>9</sup> NASA Hubble Fellowship Program Sagan Fellow.

the location of the planet, called speckles, which originate from atmospheric turbulence and optical aberrations within the instrument. The intensity of the speckles varies on large spectral scales (i.e., low spectral resolution) and can be described as a modulated stellar spectrum. The higher resolution noise is caused by photon shot-noise, detector read-noise, and imperfect modeling of the atmospheric transmission (i.e., telluric lines). The important feature of the high spectral resolution instrument is that speckles can be removed with a high-pass filter while preserving the molecular signatures of the planet spectrum almost intact.

Measuring the RV of directly imaged planets is challenging due to their flux ratio with respect to the star. The first RV measurement of an exoplanet with high-contrast imaging is  $\beta$  Pictoris b. Using CRIRES observations and the cross-correlation of a carbon monoxide (CO) molecular template, Snellen et al. (2014) measured the RV ( $-15.4 \pm 1.7 \text{ km s}^{-1}$  relative to the star) and the spin of the planet ( $25 \pm 3 \text{ km s}^{-1}$ ). Observations of HR 8799 c in the  $L$  band with Keck/NIRSPEC ( $R = 15,000$ ) in adaptive optics mode (NIRSPAO) provided detection of water in the atmosphere of the planet as well as a first estimate of its RV of  $-8.9 \pm 2.5 \text{ km s}^{-2}$  (Wang et al. 2018a). Recently,  $\text{H}\alpha$  was detected around two accreting exoplanets orbiting the star PDS 70, but the measured RVs of the emission line probe the accretion mechanism and not the orbital motion of the planet (Haffert et al. 2019). Additionally, the RV and spin of a handful of low-mass brown dwarf companions, and larger separation or lower contrast planetary mass companions, have also been made at high spectral resolution (Metchev et al. 2015; Schwarz et al. 2016; Bryan et al. 2018).

In this work, we develop a new likelihood for the analysis of high-resolution spectroscopic data. After marginalizing over the starlight modeling, we use it to measure the RV of HR 8799 b and c from Keck/OSIRIS observations. First, in Section 2, we discuss the current paradigm for the reduction of high-resolution spectroscopic data, which is based on cross-correlation. Then, we describe the observations and data reduction in Section 3. The planet detection in individual exposure is described in Section 4. The RV measurements are presented in Section 5. In Section 6, the new RV data is used to better constrain the orbits of the planets. We discuss the results and conclude in Sections 7 and 8 respectively.

A more detailed description of the calibrations and supplemental information for this article are provided in Appendices A–C. The mathematical background and derivations can be found in Appendix D.

## 2. Preamble: The Cross-correlation Function

RV measurement, detection of molecule, and abundance estimation from spectroscopic data are only possible with the use of atmospheric models and molecular templates. With HRS, the signal-to-noise ratio (S/N) of individual lines is too low to allow their independent detection and characterization. HRS therefore extensively relies on the concept of a cross-correlation function (CCF), which can be seen as way to stack the signal of the individual lines together. Briefly, the data processing steps generally include the division by a transmission spectrum of the atmosphere, a high-pass filter to remove the diffracted starlight, and the cross-correlation of a template in the spectral direction to harvest the signal of the planet. The peak value of the cross-correlation, expressed as a function of

the RV shift of the model, is a measure of the detection strength. The shape of the CCF has also been used to infer the rotational broadening of planets and brown dwarfs (Snellen et al. 2014; Schwarz et al. 2016; Bryan et al. 2018).

Assuming Gaussian and uncorrelated noise, the cross-correlation is closely related to maximum likelihood estimation. The definition of a likelihood is important for parameter estimation (e.g., RV, spin, molecular abundance, etc.) and the calculation of their uncertainty. An accurate likelihood is also paramount when combining data from different instruments with different resolutions. Indeed, the main drawback of HRS is the difficulty to constrain the local continuum used as a reference to measure the lines depths, because it is generally removed with a high-pass filter to mitigate the stellar light contamination. Low-resolution spectra or broadband photometry, on the other hand, can provide a good estimate of the continuum strength and shape. Therefore, it is desirable to jointly analyze data from different instruments, but it is seldom attempted because of the difficulty in defining the joint likelihood.

We define a data vector  $\mathbf{d}$  of size  $N_d$ , a centered Gaussian noise vector  $\mathbf{n}$  with variance  $\sigma^2$ , and a model template  $\mathbf{m}$ , such that  $\mathbf{d} = \epsilon \mathbf{m} + \mathbf{n}$ . We use the bold font convention for vectors and matrices. If  $\mathbf{m}_{\text{RV}}$  is defined as the Doppler-shifted model spectrum, the value of the cross-correlation corresponding to that RV shift is given by  $\text{CCF}(\text{RV}) = \mathbf{m}_{\text{RV}}^\top \mathbf{d}$ . In the following, the subscript in  $\mathbf{m}_{\text{RV}}$  will be omitted. Using matrix notations, the discrete CCF, expressed as a vector, can be written  $\text{CCF} = \mathbf{T} \mathbf{d}$ , with  $\mathbf{T}$  a rectangular Toeplitz matrix where each row is a shifted version of  $\mathbf{m}^\top$ , e.g.,

$$\mathbf{T} = [\mathbf{m}_{-1 \text{ km s}^{-1}}^\top, \dots, \mathbf{m}_{0 \text{ km s}^{-1}}^\top, \dots, \mathbf{m}_{+1 \text{ km s}^{-1}}^\top]. \quad (1)$$

In the context of HRS, the cross-correlation should not be understood as the simple mathematical operation, but as a Doppler shift for which the spectral shift is wavelength dependent. Brogi & Line (2019) propose a new CCF to a log-likelihood relationship and discuss past examples from the literature. All these mappings are statistically grounded, and only differ in their data modeling assumptions. More specifically, they differ in their choice of the free parameters, which are the amplitude of the signal,  $\epsilon$ , and the variance of the noise,  $\sigma^2$ . There are four possible cases: fitting for both, none, and one or the other parameters. First, jointly estimating  $\epsilon$  and  $\sigma^2$  yields the result from Zucker (2003)

$$\log(\mathcal{L}) \propto -\frac{1}{N_d} \log \left( 1 - \left( \frac{\mathbf{m}^\top \mathbf{d}}{N(\mathbf{m}^\top \mathbf{m})(\mathbf{d}^\top \mathbf{d})} \right)^2 \right). \quad (2)$$

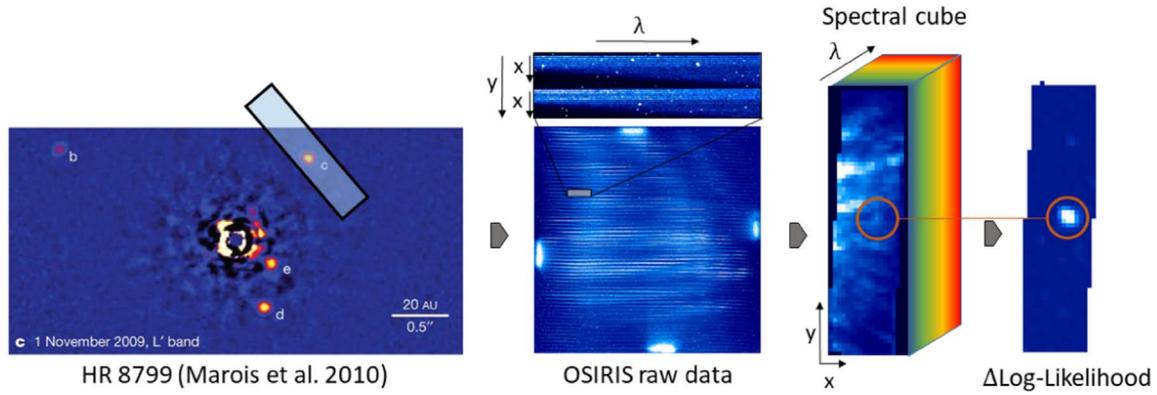
Conversely, if both parameters are fixed, there is no free parameter to estimate and the log-likelihood is given by Lockwood et al. (2014) as

$$\log(\mathcal{L}) \propto \mathbf{m}^\top \mathbf{d}, \quad (3)$$

Brogi & Line (2019) only estimate the variance of the noise and assume the planet amplitude fixed, resulting in the following log-likelihood:

$$\log(\mathcal{L}) \propto -\frac{1}{N_d} \log(\mathbf{m}^\top \mathbf{m} + \mathbf{m}^\top \mathbf{d} - 2\mathbf{m}^\top \mathbf{d}). \quad (4)$$

Finally, assuming the variance is fixed and fitting for the amplitude of the planet model yields the following likelihood



**Figure 1.** Schematics of the observations of the planets orbiting HR 8799 using the OSIRIS integral field spectrograph at the Keck observatory. From left to right: the HR 8799 four-planet system (Marois et al. 2010) over which is laid the OSIRIS field of view (2 mas platescale mode), a raw detector image featuring the individual spaxel spectra, a sky-subtracted and mean-combined image of the reconstructed data cube from the OSIRIS data reduction pipeline, and the final detection map of a single 10 minute exposure based on this work.

(Cantalloube et al. 2015; Ruffio et al. 2017):

$$\log(\mathcal{L}) \propto -\frac{(\mathbf{m}^T \mathbf{d} / \sigma^2)^2}{\mathbf{m}^T \mathbf{m} / \sigma^2}, \quad \text{S/N} = \frac{\mathbf{m}^T \mathbf{d} / \sigma^2}{\sqrt{\mathbf{m}^T \mathbf{m} / \sigma^2}}. \quad (5)$$

These four examples directly relate the CCF to a likelihood, but the model of the data remains simple. We argue for a more fundamental approach to the modeling of HRS data. In statistical analysis, it is considered better practice to forward model the signal and use minimal preprocessing to avoid modifying the distribution of the noise and decrease the validity of the likelihood. An example of heavily processed data used for inference in HRS is the estimation of planet RV and spin from the CCF. Assuming Gaussian noise in the original spectrum, the CCF is also expected to follow a multivariate Gaussian distribution as it is a linear function of the data,  $CCF = \mathbf{T}\mathbf{d}$ . However, its values are correlated, with a characteristic correlation length given by the autocorrelation of the planet model spectrum. Note that we are here discussing the correlation of another correlation—the CCF. Indeed, assuming homoskedastic, uncorrelated, and centered Gaussian noise for the data, the covariance matrix of the CCF values is given by  $\Sigma_{CCF} = \mathbf{T}\mathbf{T}^T / \sigma^2$ . The row vectors of  $\Sigma_{CCF}$  are then shifted copies of the autocorrelation function of the model  $\mathbf{m}$ . Therefore, a simple  $\chi^2$  analysis of the shape of the CCF will not enable robust parameter estimation. Another more common example of data preprocessing is the interpolation of a spectrum on a regular wavelength grid, which also introduces correlated noise. Generally, any operation on the data modifies the property of the noise, making it harder to model, and therefore limiting the accuracy of the likelihood.

### 3. Observations and Data Reduction

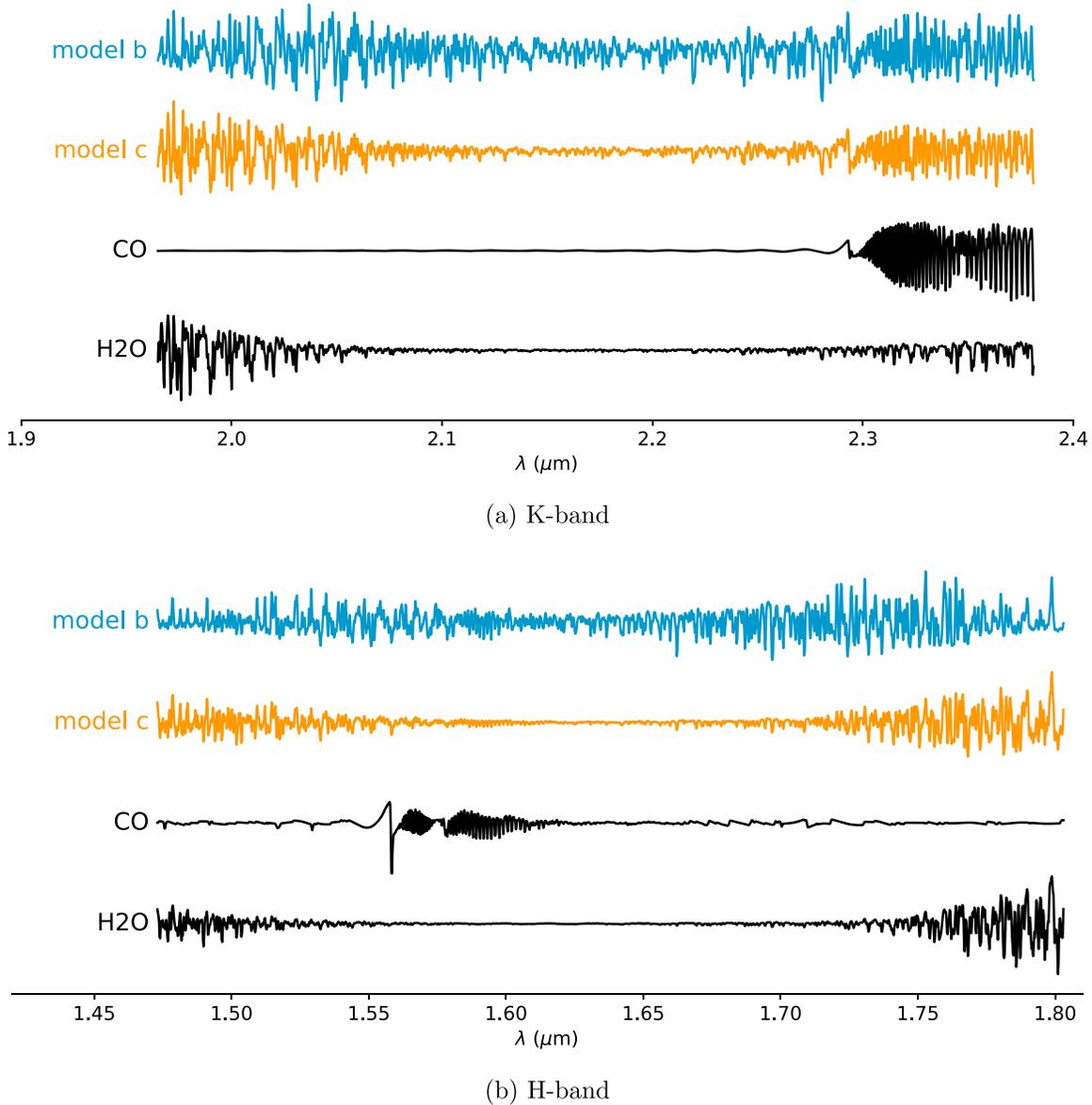
Over the past 10 years, a total of 38 hr of observations were carried out on HR 8799 b and c in the  $K$  and  $H$  bands with Keck/OSIRIS (Barman et al. 2011, 2015; Konopacky et al. 2013; Petit dit de la Roche et al. 2018). A description of the data and calibration acquisition strategy can be found in Barman et al. (2011). A summary table of the relevant observations is provided in Appendix A. Keck/OSIRIS includes a lenslet grid and a diffraction grating providing  $64 \times 16$  spectra at a resolution of  $\sim R = 4000$ . For all but one night, the smallest platescale of 20 mas/pix was used, providing a  $\sim 1.3 \times 0''.3$  field of view

(FOV), which only allows the observation of a single planet at once as shown in Figure 1. Because of mechanical uncertainties in offsetting the telescope pointing and because both the star and the planet cannot fit in the FOV, their locations are not precisely known. This requires the planet to be detected in each individual frame before the signal from different exposures can be combined. OSIRIS is fed by the Keck adaptive optics system, which therefore provides a diffraction limited point-spread function (PSF). The lack of a coronagraph hurts the sensitivity to faint companions close to the star. However, the limited raw contrast is balanced by the higher spectral resolution.

OSIRIS underwent repairs and upgrades over the years changing the data quality over time (Lockhart et al. 2019). A few notable events are relevant to this work. The cooling system experienced issues during most of 2009, which significantly increased the dark current (Barman et al. 2011). We have therefore excluded the 2009 epochs from this analysis. The dispersion grating was upgraded in 2012 December, which increased its efficiency by a factor  $\sim 2$  (Mieda et al. 2014). In 2016 January, the Hawaii-2 detector was replaced by a Hawaii-2RG (Boehle et al. 2016), greatly improving the data quality. The vast majority of the usable data has been taken prior to the upgrade.

The 1019 spectra of each spatial location (i.e., spaxel) are vertically offset from each other by only two pixels and horizontally dispersed on the detector (see Figure 1). The OSIRIS data reduction pipeline (DRP) is used to build a 3D spectral cube with an iterative algorithm for the deconvolution of the overlapping spectra (Krabbe et al. 2004; Lyke et al. 2017). The location of the spectra on the detector is stable and therefore stored in rectification matrices that are only recalibrated after hardware interventions. Rectification matrices are computed from a series of white light scans where columns of lenslets are illuminated one at a time. The wavelength solution is calculated separately using arc lamp calibrations and hard coded in the OSIRIS DRP, which chooses the suitable file based on the date of the observation. Finally, the DRP interpolates the spectral cubes on a regular wavelength grid.

Despite the stability of the instrument, the default wavelength solution suffers from biases. For example, in 2017, the mean error over the FOV was estimated to be  $\sim 13\%$  of a pixel ( $\sim 5 \text{ km s}^{-1}$  compared to a  $\sim 38 \text{ km s}^{-1}$  pixel) and spatial standard deviation up to 5% ( $\sim 2 \text{ km s}^{-1}$ ; Lockhart et al. 2019).



**Figure 2.** Atmospheric and molecular templates described in Barman et al. (2011, 2015). The spectra are convolved to the Keck/OSIRIS resolution ( $R = 4000$ ) in both spectral filters: *K* band (a) and *H* band (b). In both panels, the first two spectra are the best-fit atmospheric models resulting from Barman et al. (2015) for HR 8799 b and Konopacky et al. (2013) for HR 8799c. These spectra are used to detect and estimate the RV of the planets. As the most prominent components of the spectra, the spectral signatures of water and carbon monoxide are shown separately, but they are not used in this work. All spectra have been normalized to a unit maximum deviation from zero and then vertically offset from each other.

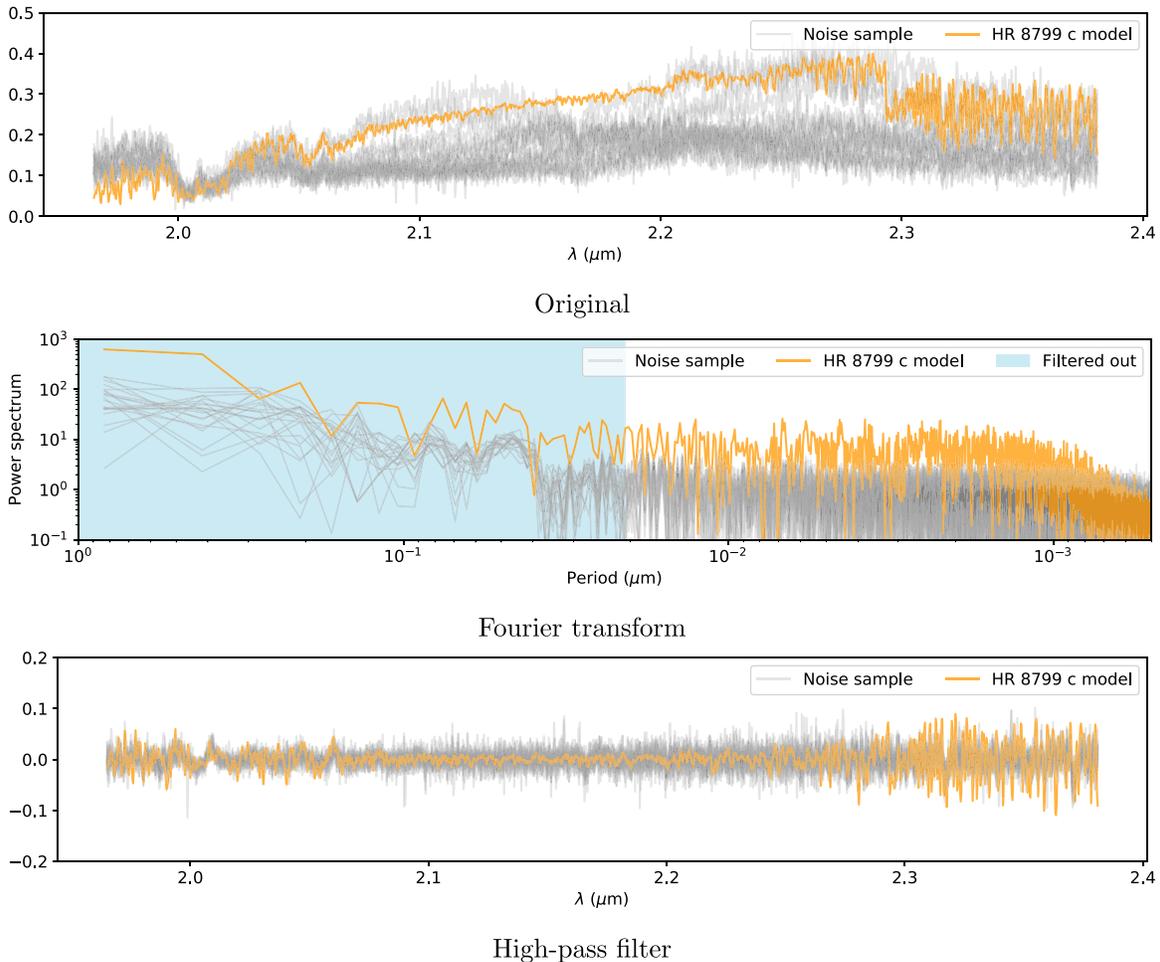
In this work, we aim at measuring RVs at a precision and accuracy of under a kilometer per second and therefore need to correct such offsets. As prescribed in the OSIRIS pipeline user manual,<sup>10</sup> we use the prominent  $\text{OH}^-$  radical emission lines from sky-background observations as a wavelength reference (Appendix B). To simplify its implementation, we do not derive a full third-order polynomial wavelength solution such as the one used in the instrument pipeline. We instead limit the correction to a single offset per pixel and per night of observation. As a consequence, the default wavelength solution and instrument pipeline are still used to build the spectral cubes. The wavelength offset is only used in the calculation of the Doppler-shifted planet spectrum model, but it is neither accounted for in the modeling of the transmission spectrum nor

the spurious starlight model. The calibration is described in more detail in Appendix B.

The transmission profile of the instrument and the atmosphere, the super-sampled PSF, and the flux are calibrated nightly using A0 reference star observations according to Appendix C.

A critical piece of the data analysis is the atmospheric models used for the planet detection, and RV measurement. We use the templates described in Barman et al. (2011, 2015), which are shown in Figure 2 for both the *H* and *K* band. The molecular templates are generated from the full atmospheric model, therefore including a realistic temperature–pressure profile, but only including a subset of opacity sources when computing the outgoing spectrum of the planet. The opacities that are accounted for are the pseudo-continuum (e.g.,  $\text{H}_2\text{--H}_2$  collision induced absorption) and the specific molecule that we

<sup>10</sup> [https://www2.keck.hawaii.edu/inst/osiris/OSIRIS\\_Manual\\_v2.3.pdf](https://www2.keck.hawaii.edu/inst/osiris/OSIRIS_Manual_v2.3.pdf)



**Figure 3.** Medium resolution spectrum ( $R \approx 4000$ ) of a planet model compared to speckle noise from Keck/OSIRIS at the  $K$  band. We use the best-fit atmospheric model of HR 8799c from Konopacky et al. (2013). In panel (a), the noise sample spectra correspond to individual spaxels in the Keck/OSIRIS field of view, which have been corrected for atmospheric and instrumental transmission using reference star observations. Panel (b) features the Fourier transform of each spectrum, as well as the high-pass filter cutoff used in this work to subtract the speckle noise. The resulting high-pass filter spectra are illustrated in panel (c).

are trying to detect. The spectra are convolved using a Gaussian with a wavelength dependent FWHM matching the Keck/OSIRIS resolution  $R = 4000$ .

#### 4. Data Model and Planet Detection

The stellar halo makes the detection of planets close to the star challenging. With a  $1''.7$  separation from the star, HR 8799 b can easily be seen in a mean-combined raw data cube, but spotting HR 8799c ( $0''.94$ ) is significantly more challenging. The dithering of the image between exposures—used to average wavelength calibration, detector noise, and sky-subtraction noise—adds another level of difficulty. Indeed, the offsets recorded in the headers are unfortunately not precise enough to allow blind stacking of the signal from different images. As a consequence, it is first necessary to detect and localize the planet in individual spectral cubes before any subsequent analysis. In this section, we will introduce the modeling of the data and describe the planet detection algorithm, which is based on a likelihood ratio test. The statistics background and derivations used in this work are detailed in Appendix D.

We chose a linear statistical model of the data  $\mathbf{d}$ , consisting of the model matrix  $\mathbf{M}_\psi$ , itself a function of the parameters  $\psi$

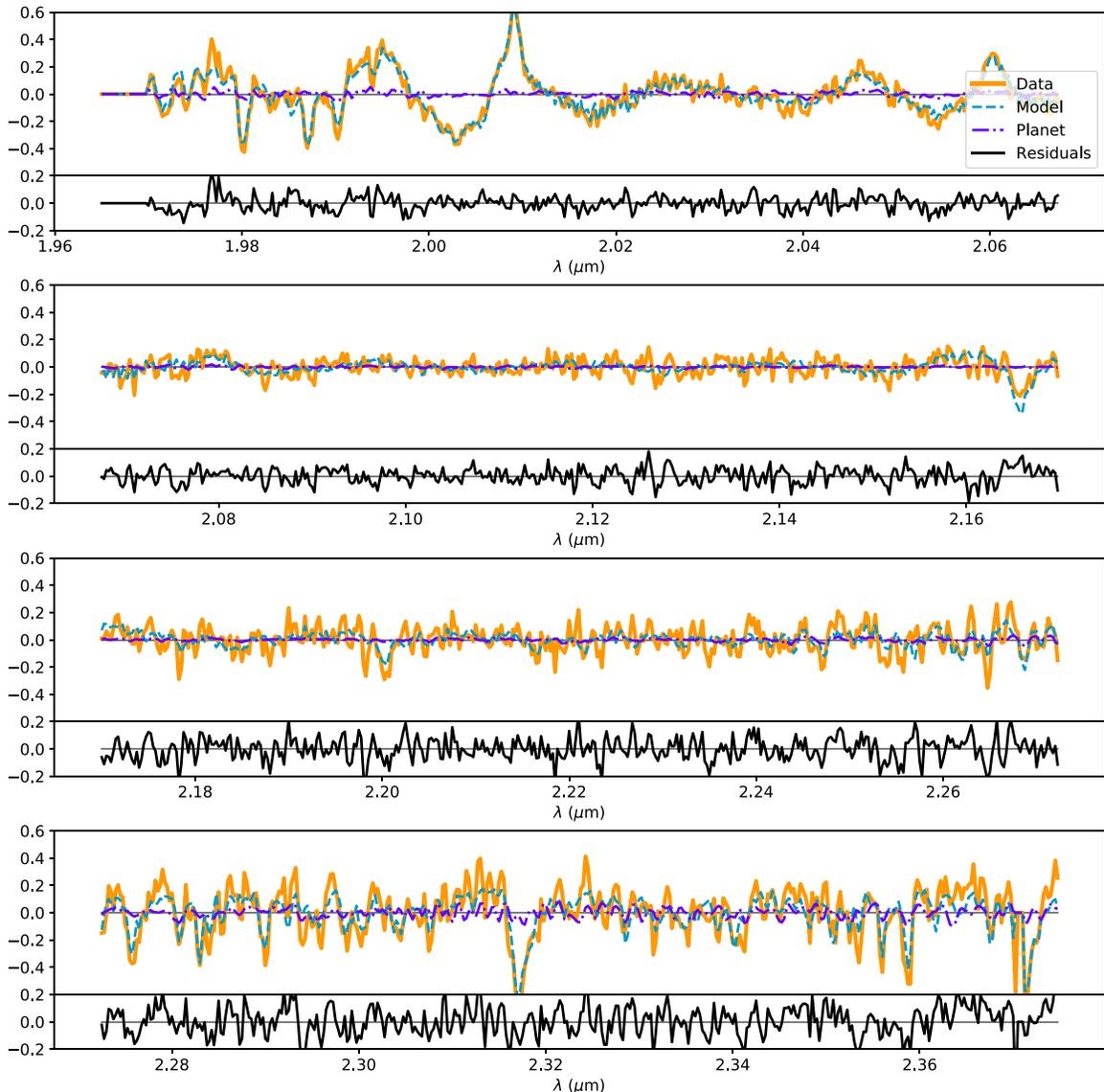
(e.g., RV, atmospheric model, etc.), linear parameters  $\phi$ , and centered uncorrelated Gaussian noise vector  $\mathbf{n}$ , such that

$$\mathbf{d} = \mathbf{M}_\psi \phi + \mathbf{n}. \quad (6)$$

Each term is described in more detailed in the following.

The data  $\mathbf{d}$  is a vector of  $N_d = 5 \times 5 \times N_\lambda$  elements corresponding to a  $5 \times 5$  spaxel subset of the spectral cube, including  $N_\lambda$  spectral channels with  $N_\lambda = 1665$  in the  $K$  band. It is a vectorized postage-stamp-sized data cube centered on the assumed position of the planet. The stellar halo, or speckles, can be seen as a modulation of the stellar spectrum varying with position and wavelength. Because it is a distortion of the continuum, the effect of the speckles is minimized at higher resolution, which is why the data is generally high-pass filtered.

Before high-pass filtering, we flag and mask bad pixels as followed. First, we retrieve the bad pixels identified by the OSIRIS DRP and saved as fits file extensions. Then, we flag additional bad pixels by removing the pseudo-continuum with a 100 pixel median filter, and identifying outliers with a threshold equal to seven times the median absolute deviation. Immediately neighboring pixels to any bad pixels are also flagged as bad, which also include the edge of the FOV. The



**Figure 4.** Residuals of a single model fit as a function of wavelength. The residuals were averaged over the 5 pixel wide area around the location of the planet. Here, the planet model is also plotted separately from the data model, which includes both the starlight and the planet. For this figure, we used a single spectral cube of HR 8799c.

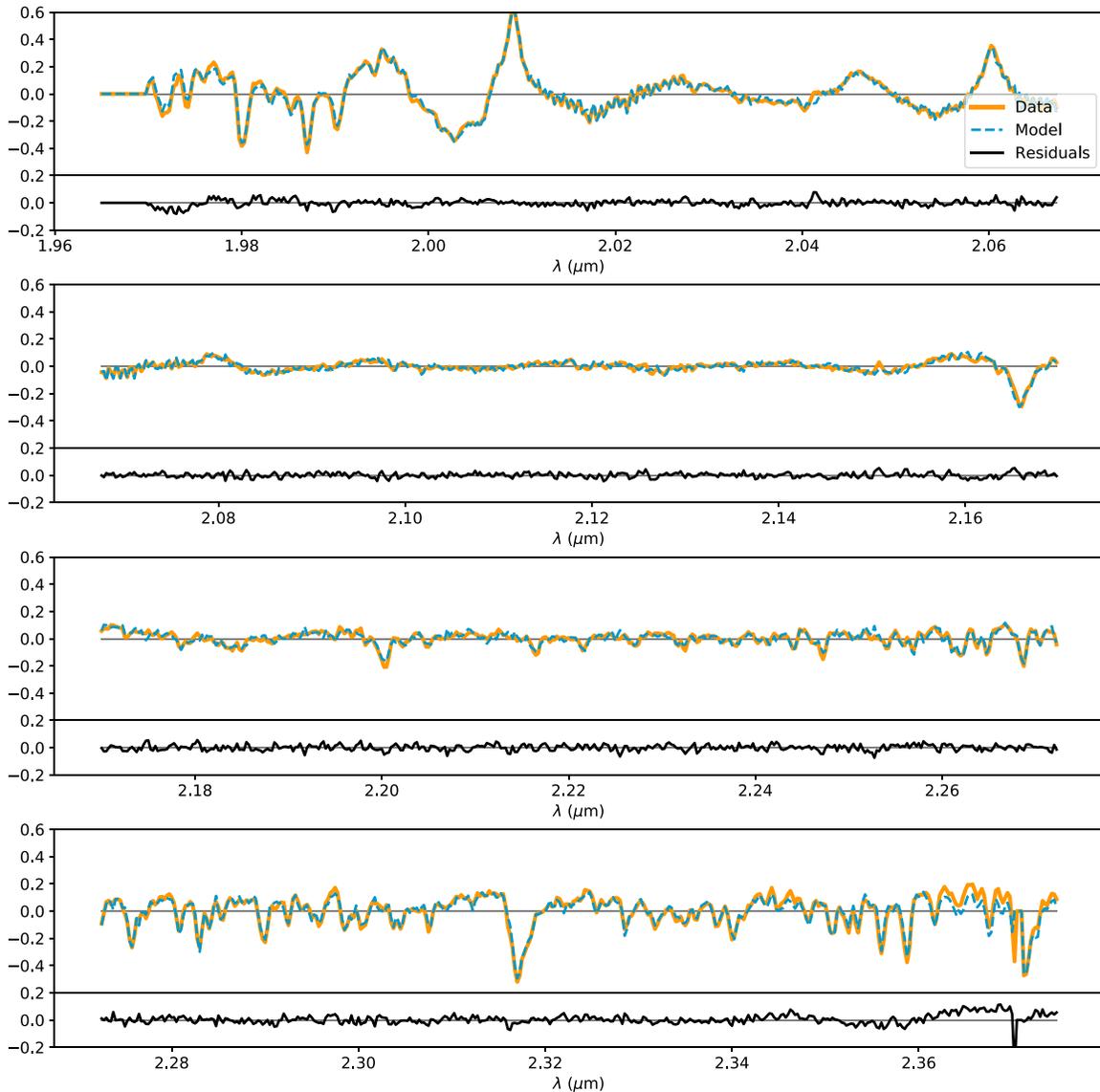
bad pixels are temporarily replaced by the estimated pseudo-continuum.

Figure 3 illustrates the bad-pixel corrected data and the high-pass filtering for a single spaxel. We chose a linear Fourier-based high-pass filter with a cutoff corresponding to  $1/20$  of the spectral band (i.e., periodicity cutoff of 21 nm compared the 416 nm width of the  $K$  band). In practice, we need to ensure continuity of the spectrum at the edge of the spectral band, which is why we first concatenate a mirrored copy of the spectrum before calculating the Fourier transform. The cutoff can be justified from Figure 3; indeed, the noise is assumed to be uncorrelated and the Fourier transform of white noise should have a flat power spectrum. While this is approximately true at higher spectral resolution, the speckle noise introduces a lot of power at lower spectral resolution. Therefore, the cutoff is chosen to match the location where the power spectrum flattens. The data can then be written in terms of its low- and high-pass filtered component so that  $\mathbf{d} = \mathbf{d}_L + \mathbf{d}_H$ . However,

to avoid cluttering the notation, we will omit the subscript and assume high-pass filtered quantities, unless specified otherwise.

The matrix  $\mathbf{M}_\psi$  includes the model of the planet as well as a model of the starlight from the host star at the location of the planet. It is therefore a function of the planet and the star spectra, the planet RV, the combined transmission of the instrument and the atmosphere, and the PSF of the instrument, which are all represented by the nonlinear parameters  $\psi$ . We write  $\mathbf{M}_\psi = [\mathbf{c}_{0,\text{planet}}, \mathbf{c}_1, \dots, \mathbf{c}_{25}]$ , where the  $\mathbf{c}_i$  are column vectors with the same size as the data vector  $\mathbf{d}$  and also function of  $\psi$ . The first column  $\mathbf{c}_{0,\text{planet}}$  is the high-pass filtered model of the planet, therefore equal to the vectorized PSF multiplied by the planet spectrum and the transmission, and normalized to the flux of the star. Each  $\mathbf{c}_{i>0}$  represents the diffracted starlight of one of the 25 spaxels under the planet PSF. They are defined as

$$\mathbf{c}_i = \mathbf{d}_{i,L} \frac{(TS_{\text{star}})_H}{(TS_{\text{star}})_L}, \quad (7)$$



**Figure 5.** Average residuals over the entire field of view of a single exposure shown as a function of wavelength. The residuals were averaged over both the location of the planet and the 5 pixel box of the data vector. For this figure, we used a single spectral cube of HR 8799c.

where  $\mathbf{d}_{i,L}$  is the low-pass filtered data vector of a spaxel  $i$ ,  $\mathcal{T}$  is the mean transmission profile defined in Appendix C, and  $\mathcal{S}_{\text{star}}$  is a model spectrum of HR 8799 from the Phoenix library (Husser et al. 2013). The multiplication by  $\mathbf{d}_{i,L}$  effectively modulates the spectral lines, such that their depths match the strength of the speckles at a given position and wavelength.

We then define the 26 linear parameters as  $\phi = [\epsilon, a_1, \dots, a_{K-1}]$ , with  $\epsilon$  being the planet to star flux ratio and  $a_i$  the amplitude of the starlight model  $c_i$ .

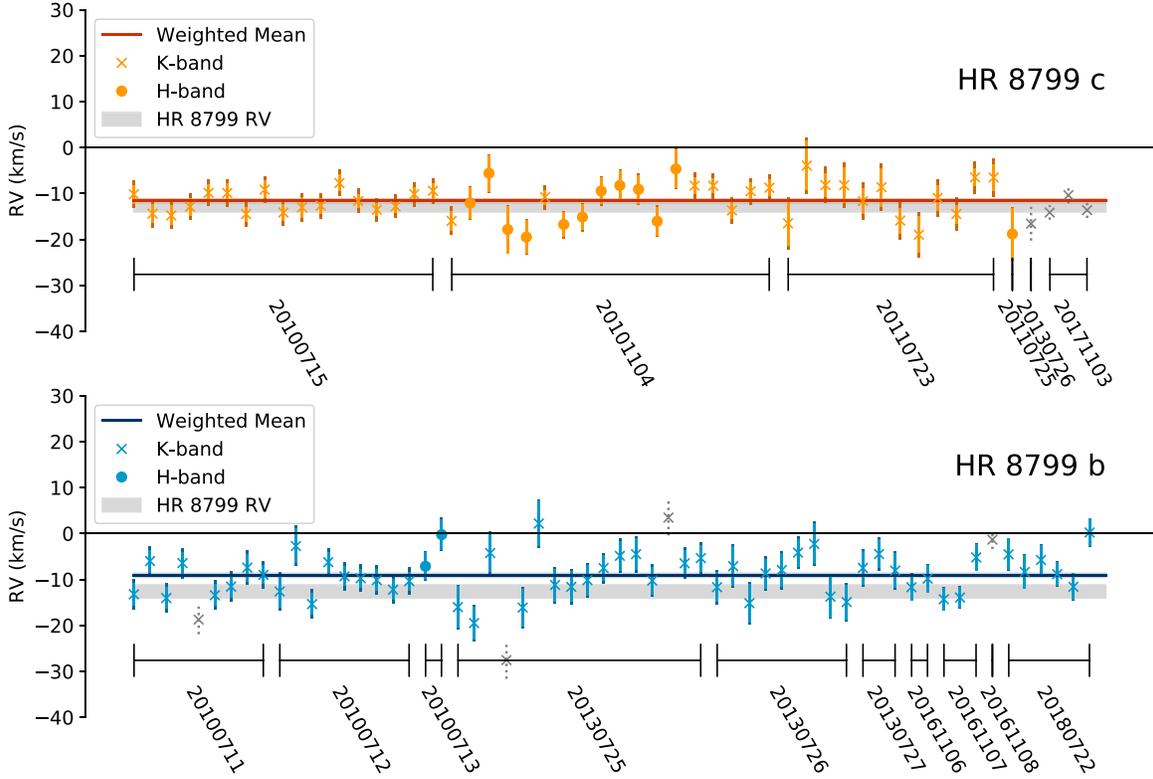
Finally,  $\mathbf{n}$  is a centered Gaussian random vector with covariance  $\Sigma = s^2 \Sigma_0$ , where  $\Sigma_0$  is a diagonal matrix with diagonal elements equal to  $\mathbf{d}_L$  and  $s^2$  is a variance scaling factor.

The likelihood is given in Appendix D.1 and the maximum likelihood estimate of the parameters is described in Appendix D.2. We define the maximum likelihood estimate of the linear parameters  $\tilde{\phi}$  and the maximum likelihood estimate of the covariance scaling factor  $\tilde{s}$  corresponding to Equations (16) and (18), respectively. The previously identified bad pixels are masked when fitting the model. In Figure 4, we

show the data, model, and residuals after the maximum likelihood fit, all averaged over the 5 pixel wide area at the location of the planet. It shows that the model is a fine fit to the data. The average residuals over the entire FOV is shown in Figure 5; it is calculated by fitting the planet at each location of the image.

While we assume a diagonal covariance of the noise in this work, a Cholesky decomposition could be used to reduce the problem to a diagonal covariance matrix. This statistical framework is flexible, and applicable to a wide variety of problems since there are no constraints on the definition of the linear model  $\mathbf{M}$ . The covariance of the estimated linear parameter and the S/N of the planet are discussed in Appendices D.3 and D.4, respectively.

The planet detection consists in comparing the maximized likelihood of the data ( $\log \mathcal{L}(\tilde{\phi}, \tilde{s}^2)$ ; (30)) for two different hypotheses, assuming there is a planet or not (Appendix D.5). The planet-free model is derived from the model described previously, but omitting the planet component, which means  $\mathbf{M}_{0\psi} = [c_1, \dots, c_{25}]$  and  $\phi_0 = [a_1, \dots, a_{K-1}]$ . We evaluate the



**Figure 6.** Radial velocity (RV) measurement of HR 8799c (upper panel) and b (lower panel) by individual exposures. The error bars are inflated to account for the error in the wavelength solution, which is added in quadrature to the statistical error (small darker extensions of the error bars). The *K* and *H* band observations are indicated by crosses and circles, respectively. The gray regions represent the current uncertainty for the RV of the star. The grayed data points have been ignored in the subsequent analysis.

logarithm of the likelihood ratio,

$$\log \mathcal{L}(\tilde{\phi}, \tilde{s}^2) - \log \mathcal{L}(\tilde{\phi}_0, \tilde{s}_0^2), \quad (8)$$

at each location in the image and with different assumed RVs for the planet. The likelihood ratio is maximized at the location and for the correct RV of the planet. The maps of log-likelihood differences at the RV of star for HR 8799 b and c are shown in Appendix D.5. Because the planets have wide almost face-on orbits, their RVs are expected to be within a few kilometers per second of that of the star ( $-12.6 \pm 1.4 \text{ km s}^{-1}$ , Gontcharov 2006; or  $-10.9 \pm 0.5 \text{ km s}^{-1}$ , Wang et al. 2018a). The barycentric correction was calculated with the Python module `barycorrpy`<sup>11</sup> (Kanodia & Wright 2018a, 2018b). We then selected the best spectral cubes, with strong detections, to be used in subsequent analysis and flagged exposures with strong artifacts as bad.

## 5. Radial Velocity

### 5.1. Inference

We propose a framework for the RV measurements of directly imaged planets and the derivation of their uncertainties. The challenge of such observations is the dominance of the stellar halo at the location of the planet. We will show how the statistical model introduced in Section 4 is well suited for such inference, and even more generally the estimation of any nonlinear parameters  $\psi$ . In particular, it allows the analytical

marginalization of the stellar halo subtraction. We will demonstrate that the method can provide the best constraints on the RV of HR 8799 b and c using medium resolution spectroscopy ( $R \approx 4000$ ). The *H* and *K* band observations prior to 2013, which represents most of the data used in this work, were already published (Barman et al. 2011, 2015; Konopacky et al. 2013; Petit dit de la Roche et al. 2018). However, these studies did not attempt to constrain the RV of the planets.

In Appendix D.6, we show that the posterior on  $\psi$  marginalized over the linear parameters  $\Phi$  and the covariance scaling parameter  $s$  is directly related to the minimized chi-squared,  $\chi_{\phi=\tilde{\phi},\psi}^2$ . We note that the RV of the planet is a component of  $\psi$ , and we will assume that the other components of  $\psi$  are known and fixed. Assuming an improper uniform prior for  $\mathcal{P}(\text{RV})$  and  $\mathcal{P}(\phi)$ , and an improper scaling parameter prior  $\mathcal{P}(s) \propto 1/s^2$ , we get Equation (40),

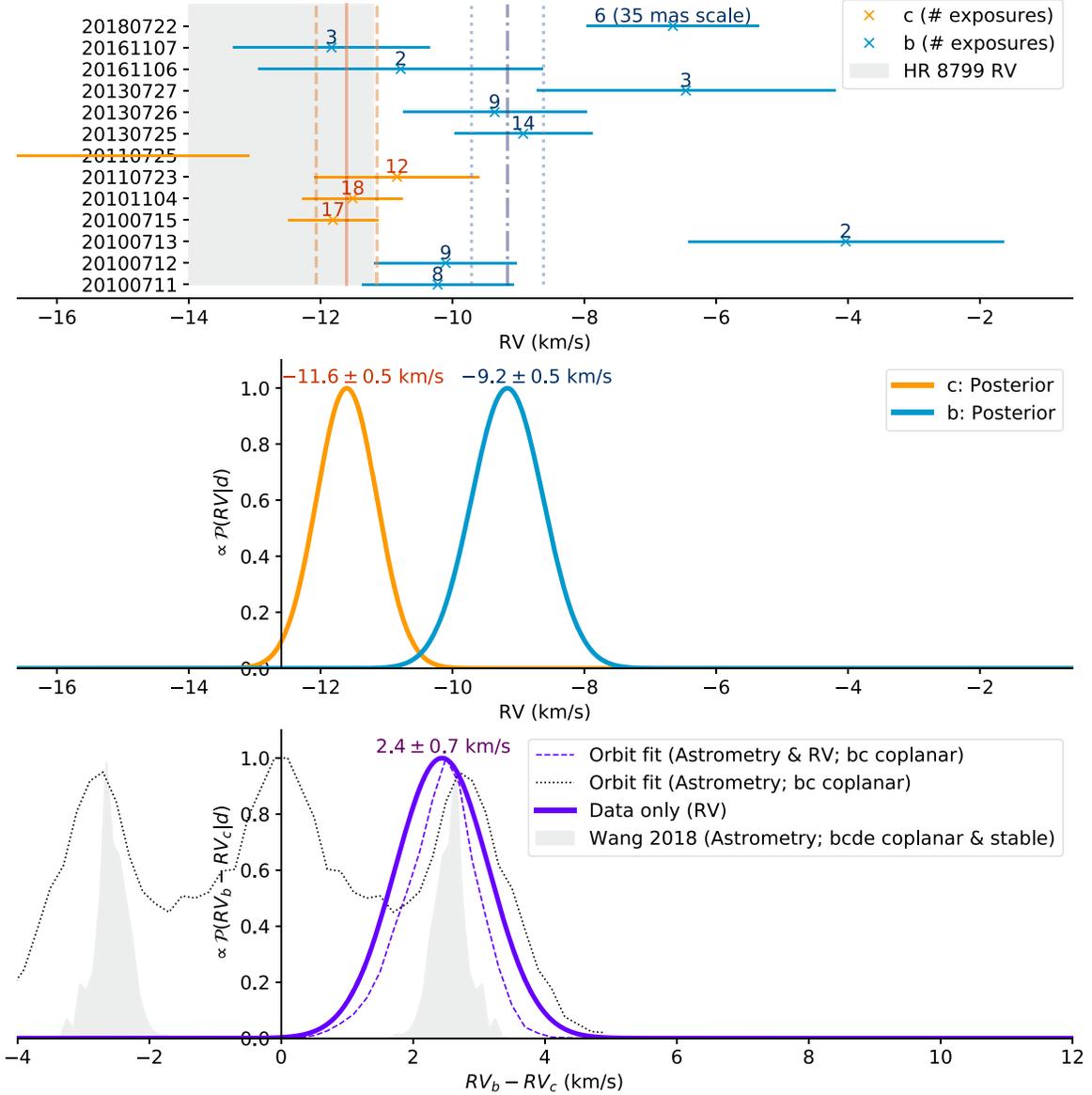
$$\mathcal{P}(\text{RV}|\mathbf{d}) \propto \frac{1}{\sqrt{|\Sigma_0| \times |\mathbf{M}_{\text{RV}}^\top \Sigma_0^{-1} \mathbf{M}_{\text{RV}}|}} \left( \frac{1}{\chi_{\phi=\tilde{\phi},\text{RV}}^2} \right)^{\frac{N_{\mathbf{d}} - N_{\phi} + 1}{2}}. \quad (9)$$

The  $\chi^2$  is defined as

$$\begin{aligned} \chi_{0,\phi=\tilde{\phi},\text{RV}}^2 &= (\mathbf{d} - \mathbf{M}\tilde{\phi})^\top \Sigma_0^{-1} (\mathbf{d} - \mathbf{M}\tilde{\phi}), \\ &= \mathbf{d}^\top \Sigma_0^{-1} \mathbf{d} - \mathbf{d}^\top \Sigma_0^{-1} \mathbf{M} (\mathbf{M}^\top \Sigma_0^{-1} \mathbf{M})^{-1} \mathbf{M}^\top \Sigma_0^{-1} \mathbf{d}. \end{aligned} \quad (10)$$

In Section 2, we discussed how the cross-correlation can be related to the log-likelihood and therefore to the  $\chi^2$ . Because the cross-correlation can be interpreted in the context of a

<sup>11</sup> <https://github.com/shbhuk/barycorrpy>



**Figure 7.** Combined analysis of the radial velocity (RV) measurements of HR 8799 b and c. The upper panel features the RV estimate for each individual night annotated with the number of exposures used. The middle panel shows the final posteriors when combining all data. For each planet, the tighter posterior results from multiplying individual posteriors assuming they are independent, but corrected to yield  $\chi_r^2 = 1$ . The lower panel compares the posteriors on the differential RV between HR 8799 b and c resulting from a set of orbital fits. The solid purple posterior is directly derived from the middle panel. The grayed area corresponds to the posterior derived from Wang et al. (2018b), which only uses astrometric data from direct imaging and assumes that the four planets, HR 8799 b, c, d, and e, are coplanar and stable. The purple dashed and black dotted posteriors are derived from the orbital fits presented in this work, with and without the RV data respectively. In both cases, the astrometric data is used and HR 8799 b and c are assumed to be coplanar, but HR 8799 e and d are not included.

single parameter linear model, this framework can be thought of its multidimensional generalization. Indeed, the term

$$\mathbf{d}^T \Sigma_0^{-1} \mathbf{M} (\mathbf{M}^T \Sigma_0^{-1} \mathbf{M})^{-1} \mathbf{M}^T \Sigma_0^{-1} \mathbf{d},$$

is conceptually similar to Equation (5),

$$\log(\mathcal{L}) \propto (\mathbf{d}^T \mathbf{m} / \sigma^2) (\mathbf{m}^T \mathbf{m} / \sigma^2)^{-1} (\mathbf{m}^T \mathbf{d} / \sigma^2).$$

In practice, the calculation of the minimized  $\chi^2$  of a linear model is straightforward and computationally efficient. As part of the planet detection step, the minimized  $\chi^2$  is already calculated as a function of the RV of the planet, which can be converted into the RV posterior using Equation (9).

## 5.2. Results

The RV estimates and associated uncertainties for the selected data are shown in Figure 6. Assuming coplanar and stable orbits for the HR 8799 planets, the RV of HR 8799c,  $RV_c$ , is expected to have changed by  $0.6 \text{ km s}^{-1}$  since 2010 (Wang et al. 2018b). We therefore only combine data spanning 2010–2011, during which the RV change is significantly below our precision. The RV of HR 8799 b,  $RV_b$ , is not expected to have changed in any measurable amount, which is why all available epochs are used.

A wavelength offset was calculated from  $\text{OH}^-$  emission lines in sky-background observations to account for spatial and night-to-night variations of the wavelength solution. As

described in Appendix B, the spatial variations and the mean offset over the FOV are calculated separately. The spatial variations of the wavelength offset relative to the mean of the FOV is assumed to be constant within a year, while the mean offset is assumed to vary from night to night. An error term for the spatially varying offset is added in quadrature to the RV statistical uncertainties. The values of the error used are listed in Table 3. Because each exposure is dithered with respect to each other, this additional error is assumed to be independent for each exposure. The inflated errors are shown in Figure 6 with a darker color, which suggests that the effect of the additional error is minimal compared to the statistical error. After removing  $3\sigma$  outliers, the  $\chi_r^2$  in Figure 6 for HR 8799 b and c are 1.0 and 1.4, respectively, showing fine agreement with the error bars. However, due to the larger value for HR 8799 b, we inflate the errors such as to normalize the  $\chi_r^2$  to unity. The combined posteriors for  $RV_c$  and  $RV_b$  shown in the middle panel of Figure 7 are calculated by multiplying the individual posteriors for each exposure.

Due to the limited number of samples, it is harder to estimate the error on the nightly mean offset. The error bars in the upper panel of Figure 7 seem to already be consistent with the weighted mean of the RV, and therefore are unlikely to require an additional error term for the wavelength solution. In order to verify this hypothesis, we compare the Bayesian evidence of two models. The first model uses the uncertainties from Figure 7 to infer the mean RV of HR 8799 b. The second model includes an additional error term, which is added in quadrature to the uncertainties on the nightly estimates. When using a uniform prior for the additional error term, the Bayes factor is 3.8 in favor of the latter model, which is not a significant difference according to Kass & Raftery (1995). Additionally, the mean RV posterior marginalized over the additional error term does not significantly differ from the original posterior. We conclude that an additional calibration error is not necessary to explain the data.

As a summary, the barycentric corrected final RVs for HR 8799 b and c, in 2010, are estimated to be  $-9.2 \pm 0.5 \text{ km s}^{-1}$  and  $-11.6 \pm 0.5 \text{ km s}^{-1}$ , respectively. The night-by-night estimates are also reported in Table 1.

Due to the uncertainty on the RV of the star, it is not possible to derive tight constraints on the relative velocity of the planets to the system center of mass. However, the relative velocity of the planets with respect to each other can be used to derive the 3D orientation of the orbits. The tightest constraints on the orbital motion of HR 8799 b and c comes from direct imaging astrometry combined with coplanarity and stability priors of the orbits (Wang et al. 2018b). Because direct imaging does not distinguish between in-the-plane or out-of-the-plane of the sky motion of the planets, there exist two families of orbits with equal probability. They can be described in terms of  $RV_b$  and  $RV_c$  such that

$$\begin{aligned} \text{1st family: } & RV_b \approx +2 \text{ km s}^{-1}, RV_{c,2010} \approx -0.8 \text{ km s}^{-1}, \\ & RV_b - RV_{c,2010} \approx +3 \text{ km s}^{-1}, \\ \text{2nd family: } & RV_b \approx -2 \text{ km s}^{-1}, RV_{c,2010} \approx +0.8 \text{ km s}^{-1}, \\ & \text{and } RV_b - RV_{c,2010} \approx -3 \text{ km s}^{-1}. \end{aligned}$$

In the bottom panel of Figure 7, we therefore compare the posterior of  $RV_b - RV_{c,2010}$  predicted from direct imaging with the posterior derived from our RV measurements. Our result unambiguously favors the first family of orbits.

Table 1

Barycentric Corrected RVs of HR 8799 b and c Measured with Keck/OSIRIS

Planet	Date	RV ( $\text{km s}^{-1}$ )	# Cubes
b	2010 Jul 11	$-10.2 \pm 1.2$	8
	2010 Jul 12	$-10.1 \pm 1.1$	9
	2010 Jul 13	$-4.0 \pm 2.4$	2
	2013 Jul 25	$-8.9 \pm 1.0$	14
	2013 Jul 26	$-9.4 \pm 1.4$	9
	2013 Jul 27	$-6.5 \pm 2.3$	3
	2016 Nov 6	$-10.8 \pm 2.2$	2
	2016 Nov 7	$-11.8 \pm 1.5$	3
	2018 Jul 22	$-6.7 \pm 1.3$	6
c	2010 Jul 15	$-11.8 \pm 0.7$	17
	2010 Nov 4	$-11.5 \pm 0.8$	18
	2011 Jul 23	$-10.8 \pm 1.3$	12
	2011 Jul 25	$-18.8 \pm 5.7$	1
b	Combined	$-9.2 \pm 0.5$	
c	Combined	$-11.6 \pm 0.5$	

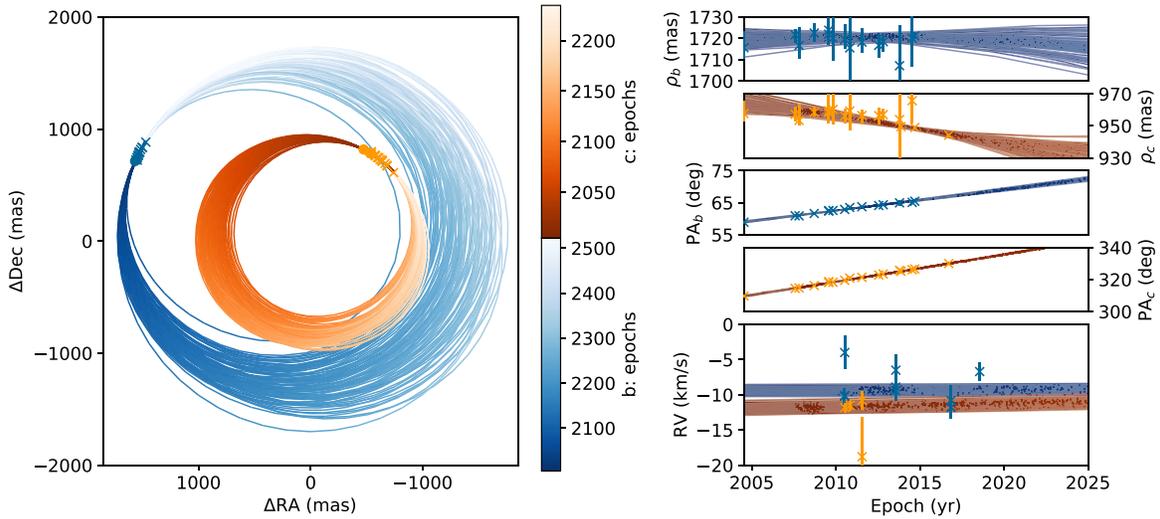
**Note.** For reference, the current estimate of the RV of the star is  $-12.6 \pm 1.4 \text{ km s}^{-1}$  (Gontcharov 2006) or  $-10.9 \pm 0.5 \text{ km s}^{-1}$  (Wang et al. 2018a), which has not been subtracted from the values in this table due to the large uncertainty.

## 6. Orbits Fit

In the previous section, the RV predictions resulting from direct imaging assumed that the four planets orbiting HR 8799 are coplanar and stable (Wang et al. 2018b). Here, we will relax the stability constraint and study the effect of the RV on the joint orbital fit of HR 8799 b and c. The orbital parameters are estimated for two cases: with and without RV measurements. In both cases, we assume that the two orbits are coplanar. The assumption is necessary because the RV of the star is not precisely known and needs to be fitted for. The relative RV of HR 8799c with respect to the star is expected to be of the order of the measurement uncertainty, which means that its sign is not well constrained by the data. However, the sign of the relative RV of HR 8799 b is positive with high confidence. The uncertainty on the sign of HR 8799c RV is creating two families of acceptable orbits, one of which with high mutual inclination between the two planets. We argue that high mutual inclinations are far less likely than near coplanar orbits for stability reasons; stable orbits with mutual inclination  $>8^\circ$  are difficult to find (Wang et al. 2018b). The two planets are forced to share the same longitude of ascending node, inclination, parallax, stellar mass, and, if applicable, stellar RV. We assume that the stellar RV and the center of mass RV are identical. We use the RV measurements from Table 1, the Keck/NIRC2 astrometric measurements from Konopacky et al. (2016, therein Table 2), and the Gemini/GPI data from Wang et al. (2018b, therein Table 2).

The posterior distribution of the parameters is calculated using the open source python package `orbitize!` (Blunt et al. 2019). Within `orbitize!`, we use the parallel tempered implementation of the affine-invariant ensemble sampler for Markov chain Monte Carlo (`ptemcee`; Foreman-Mackey et al. 2013; Vausden et al. 2016). Modifications were made to `Orbitize!` to allow the use of planetary RVs and enforce coplanarity.

An inverse prior was used for the semimajor axis ( $a \in [10^{-3}, 10^7]$ ). The inclination prior follows the geometric



**Figure 8.** Orbits of HR 8799 b (blue) and c (orange) randomly sampled from their posterior. The orbital fits include the radial velocity (RV) measurements of the planets. From top to bottom on the right, the panels show the separation of b then c, the position angle of b then c, and the RV of both planets. The error bars were converted from R.A. and decl. to separation and position angle using a Monte Carlo approach.

prior ( $\sin(i)$  with  $i \in [0, \pi]$ ). Uniform priors were used for the eccentricity ( $e \in [0, 1]$ ), the argument of periastron ( $\omega \in [0, 2\pi]$ ), the longitude of ascending node ( $\Omega \in [0, 2\pi]$  if using RV,  $\Omega \in [0, \pi]$  otherwise), and the epoch of periastron passage ( $\tau \in [0, 1]$ ) expressed as the fraction of orbital period past 1858 November 17 at midnight (i.e., MJD = 0). Gaussian priors were used on the parallax ( $24.2175 \pm 0.0881$  mas; Gaia Collaboration 2018), the stellar mass ( $1.52 \pm 0.15 M_{\odot}$ ; Baines et al. 2012), and stellar RV ( $-12.6 \pm 1.4$  km s<sup>-1</sup>; Gontcharov 2006).

The number of parameters in the fits can make the convergence of the chains difficult: 13 parameters when including the RV measurements, and 12 otherwise. Based on past work, we used 16 temperatures, 1024 walkers, and the chains were thinned by a factor 50 to remove the correlation between samples. The chains were initialized from the individual fits of the planets and run for 180,000 steps, including a 90,000 burn-in phase. Therefore, the results presented below include 1800 independent samples per walker after thinning the chains ( $\approx 2 \times 10^6$  samples in total).

For the fit including the RVs, Figure 8 shows 50 randomly sampled orbits from the final posterior and their comparison to the data. The main advantage of the RV is to constrain the 3D orientation of the orbital plane of the planets, longitude of ascending node ( $\Omega$ ), and inclination. However, the precision of the current measurements do not significantly constrain the shape of the orbit, semimajor axis, and eccentricity. The posteriors of these four orbital parameters are shown in Figure 9 in both cases: with and without RV data. When the RV is not considered, the  $\Omega$  posterior is only calculated in  $[0, \pi]$ , but equivalent solutions exist in  $[0, 2\pi]$ , which is why the posterior is bi-modal. The RV unambiguously breaks down the degeneracy on  $\Omega$  that exists when solely using direct imaging data. The posteriors on the differential RV between HR 8799 b and c resulting from these orbits fits is also shown in Figure 7.

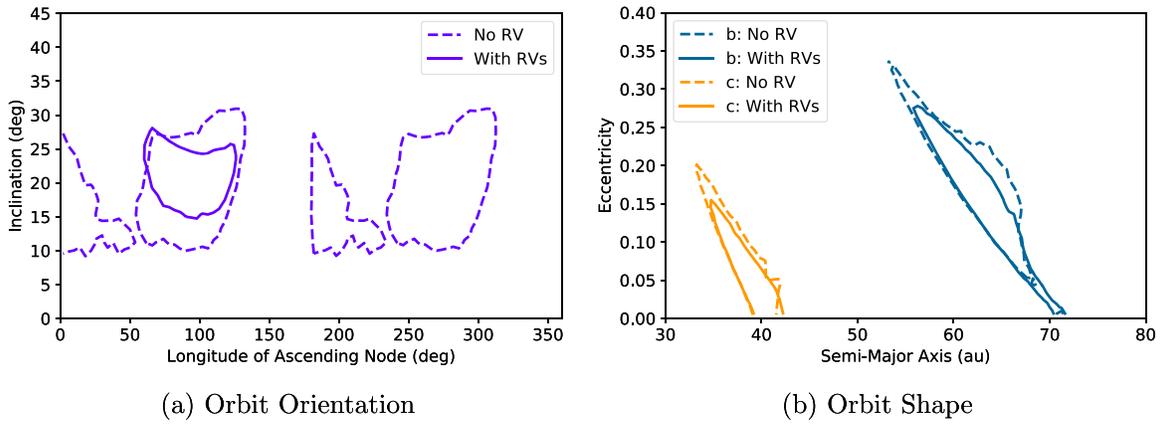
As a result, from the joint fit of HR 8799 b and c, assuming coplanar orbits and including the RV measurements, we estimate the longitude of ascending node and inclination to be  $\Omega = 89^{\circ+27}_{-17}$  and  $i = 20^{\circ+4.5}_{-3.7}$ . As a byproduct of the fit, we can also infer the

RV of HR 8799 to be  $RV_{\star} = -10.5^{+0.5}_{-0.6}$  km s<sup>-1</sup>. The derived orbital parameters are summarized in Table 2.

## 7. Discussion

The values of the inclination and the longitude of ascending node derived in this work remain consistent with Wang et al. (2018b):  $\Omega = 67^{\circ}9^{+5.9}_{-5.2}$  and  $i = 26^{\circ}8 \pm 2.3$  for the stable and coplanar orbits fit using the four planets in the system. Therefore, the discussion in Wang et al. (2018b) about the alignment of the planets with the debris disks remains valid; no deviations from coplanarity with mid-infrared observations with Herschel are detected (Mathews et al. 2014), but there is a slight discrepancy with the longitude of the ascending node of the debris disk at a millimeter wavelength with the Submillimeter Array (SMA) and the Atacama Large Millimeter/submillimeter Array (ALMA; Wilner et al. 2018). A more precise measurement of the RV of the star, combined with this work, would better constrain the longitude of the ascending node of the outer planets and test the assumption of coplanarity. The tightest constraints on  $\Omega$  are provided by HR 8799 d and e, which have covered longer portions of their orbital arcs. However, the two inner planets are not included in this work.

The prospects of such measurements are exciting. Indeed, the mass of directly imaged planets is currently estimated from atmospheric and evolutionary models, which lack absolute calibration. RV measurements of planets can contribute to their precise orbit characterization, which can enable the detection of deviations from Keplerian motion due to other planetary bodies in the system, and therefore provide independent mass measurements (Gravity Collaboration et al. 2019). Another application of RV measurements of exoplanets is the unambiguous detection of exomoons, in the very same way as the RV of stars has been a powerful exoplanet detection method. To date, the work of Teachey & Kipping (2018) is the only tentative detection of an exomoon around Kepler-1625b, which used transit timing and flux decrement of the star. Photodynamical modeling of the system shows that it is consistent with a Neptune-sized moon orbiting a several Jupiter-mass planet. Although the nature of the system, binary planet, or exomoon is up for debate, this discovery suggests



**Figure 9.** Contours of a 68% confidence interval for (a) the orientation of the orbital plane and (b) the shape of the orbit for HR 8799 b and c. The dashed contours only include the astrometric data from direct imaging, while the solid contours also include the planetary radial velocity. HR 8799 b and c are assumed to be coplanar and therefore share the same inclination and longitude of ascending node.

**Table 2**

Estimated Orbital Parameters from a Joint Fit of HR 8799 b and c using Planetary RVs and Assuming Coplanar Orbits

Parameters	Values
$a_b$	$62.3^{+4.8}_{-3.6}$ au
$a_c$	$39.7^{+1.5}_{-3.1}$ au
$e_b$	$<0.2$
$e_c$	$<0.09$
$\omega_b$	$116.^{\circ}3^{+42}_{-16}$
$\omega_c$	$61.^{\circ}{}^{+11}_{-48}$
$\tau_b$	$0.43^{+0.04}_{-0.06}$
$\tau_c$	$0.42^{+0.10}_{-0.07}$
Parallax	$24.2^{+0.1}_{-0.1}$ mas
$M_{\text{tot}}$	$1.53^{+0.11}_{-0.07} M_{\odot}$
$RV_{\star}$	$-10.5^{+0.5}_{-0.6}$ km s $^{-1}$
$\Omega$	$89.^{\circ}{}^{+27}_{-17}$
$i$	$20.^{\circ}8^{+4.5}_{-3.7}$

**Note.** The error bars represent the 68% confidence interval. The last five parameters are shared by the two planets.

that precise Doppler measurement of planets with precision under  $0.1 \text{ km s}^{-1}$  could already provide detections of exomoons (Vanderburg et al. 2018).

This science case will be supported by ongoing and future instrument development. For example, upgrades to NIRSPEC (Martin et al. 2014), CRIRES+ (Follert et al. 2014), and the Keck Planet Imager and Characterizer (KPIC; Mawet et al. 2017) will keep opening new science opportunities. KPIC is specifically designed for high-contrast exoplanets. These new capabilities were specifically designed for high-contrast imaging and studies of planetary RV or spin. This work demonstrates that such measurements can also be made at medium spectral resolution, opening new possibilities for the *James Webb Space Telescope (JWST)*. Indeed, two instruments, NIRSPEC ( $1\text{--}5 \mu\text{m}$ ,  $R = 1000\text{--}2700$ ) and MIRI ( $5\text{--}30 \mu\text{m}$ ), onboard the *JWST* are mid-infrared medium resolution spectrographs very similar to Keck/OSIRIS.

## 8. Conclusion

Using data from Keck/OSIRIS, we measured the first RVs of HR 8799 b and c. Keck/OSIRIS is an integral field

spectrograph providing 1000 near-infrared spectra at  $R \approx 4000$  resolution over of small  $\sim 1.3 \times 0.3$  FOV. Improving upon the traditional cross-correlation analysis, we developed a new forward model based statistical framework for the analysis of medium-resolution spectroscopic data of directly imaged planets. After analytically marginalizing over the starlight subtraction, we inferred the RV of HR 8799 b and c:  $RV_b = -9.2 \pm 0.5 \text{ km s}^{-1}$  and  $RV_c = -11.6 \pm 0.5 \text{ km s}^{-1}$ . Despite the lower spectral resolution, these are the most precise RV measurements of directly imaged exoplanets at high contrast. We were able to break the degeneracy in the longitude of ascending node resulting from direct imaging astrometry using the relative RV between the planets. Assuming coplanarity, we jointly estimated the orbital parameters of the two planets and constrained the spatial orientation of the orbital plane ( $\Omega = 89.^{\circ}{}^{+27}_{-17}$  and  $i = 20.^{\circ}8^{+4.5}_{-3.7}$ ) as well the RV of HR 8799 ( $RV_{\star} = -10.5^{+0.5}_{-0.6} \text{ km s}^{-1}$ ). This work demonstrates that planetary RV is possible with medium resolution spectroscopy, providing exciting prospects for the *JWST*.

We would like to thank Tuan Do and Adam Mantz for their help with OSIRIS data calibration, and statistics related issues respectively. The research was supported by grants from NSF, including AST-1411868 (J.-B.R., B.M.), 1614492, and 1614492 (T.S.B.). Support was provided by grants from the National Aeronautics and Space Administration (NASA) NNX15AD95G (J.-B.R., R.J.D.R.). This work benefited from NASA’s Nexus for Exoplanet System Science (NExSS) research coordination network sponsored by NASA’s Science Mission Directorate. Q.M.K., T.S.B., and K.K.W. were supported by NASA under grant no. NNX17AB63G issued through the Astrophysics Division of the Science Mission Directorate. Any opinions, findings, and conclusions or recommendations expressed in this work are those of the author(s) and do not necessarily reflect the views of the National Aeronautics and Space Administration. This work made use of the sky-background models made available by the Gemini observatory.<sup>12</sup> The W. M. Keck Observatory is operated as a scientific partnership among the California Institute of Technology, the University of California, and NASA. The Keck Observatory was made possible by the

<sup>12</sup> <http://www.gemini.edu/sciops/telescopes-and-sites/observing-condition-constraints/ir-background-spectra>

generous financial support of the W. M. Keck Foundation. We also wish to recognize the very important cultural role and reverence that the summit of Maunakea has always had within the indigenous Hawaiian community. We are most fortunate to have the opportunity to conduct observations from this mountain.

*Facility:* Keck I (OSIRIS).

*Software:* `astropy`<sup>13</sup> (Astropy Collaboration et al. 2013), `Matplotlib`<sup>14</sup> (Hunter 2007), `orbitize!`<sup>15</sup> (Blunt et al. 2019), `ptemcee`<sup>16</sup> (Vousden et al. 2016; Foreman-Mackey et al. 2013).

## Appendix A Observations Summary Table

A summary table of the Keck/OSIRIS observations used in the work is provided in Table 3.

## Appendix B Wavelength Calibration

In this section, we describe the additional wavelength solution calibration used in Section 5. We use the OH<sup>-</sup> radical emission lines featured in sky-background observations as a wavelength reference. While there are fewer lines in the *K* band (Figure 10) than in the *H* band (Figure 11), there are enough for calibration purposes. To simplify its implementation, we do not derive a full third-order polynomial wavelength solution such as the one used in the instrument pipeline. We instead limit the correction to a single offset, which varies as a function of pixel position and date of the observation. The spatial variations and the mean offset over the FOV are calculated separately. The spatial variations of the wavelength offset relative to the mean of the FOV is assumed to be constant within a year, while the mean offset is assumed to vary from night to night.

A few skies are typically acquired during each observing night and they are reduced as follow. First, each sky is dark subtracted and shaped into a 3D spectral cube using the OSIRIS DRP standard recipes. We identify bad pixels in

individual slices of the cube by spatially high-pass filtering them with a  $5 \times 5$  box median filter, and using a threshold equal to seven times the median absolute deviation (7MAD). For each spaxel, the wavelength solution offset is calculated from a maximum likelihood fit between a sky emission model spectrum and the data after they have been high-pass filtered in the spectral direction (Figures 10–11). The model of the sky was downloaded from the Gemini observatory website<sup>17</sup> (Lord 1992), convolved with a Gaussian matching the resolution of the instrument (FWHM corresponding to  $R = 4000$ ), shifted by a wavelength offset, evaluated on the OSIRIS wavelength grid, and high-pass filtered before being fitted to the data. Due to the sharpness of the lines, we use a 40 pixel wide median high-pass filter instead of the Fourier-based high-pass filtered used elsewhere in this work.

We then use a matched filter to find the optimal wavelength offset for each spaxel in the FOV. The matched filter is here identical to a cross-correlation since we use a solid wavelength offset instead of a Doppler shift, and also because we assume a covariance matrix equal to the identity. As a result of this process, a spatial map of wavelength offsets is obtained for each sky observation in our data set. For each year, we then derive a master map of the spatial variations by averaging the mean-subtracted maps. This spatial calibration maps for each year are shown in Figure 12. The error on the spatial offset calibration is calculated for each sky from the residual map after subtracting the mean-corrected master map. We report the smallest standard deviation for each year in Table 3. This error is used to inflate the RV measurement error for each individual exposure. The mean offset is computed by taking the average of all the wavelength offsets maps taken during a single night. The resulting mean offset for each night is given in Table 3, which show deviations from the instrument calibration up to  $13.4 \text{ km s}^{-1}$  in 2017.

The master calibration for each night is calculated by adding the corresponding mean offset, calculated for each night, to the map of spatial variation, calculated for each year.

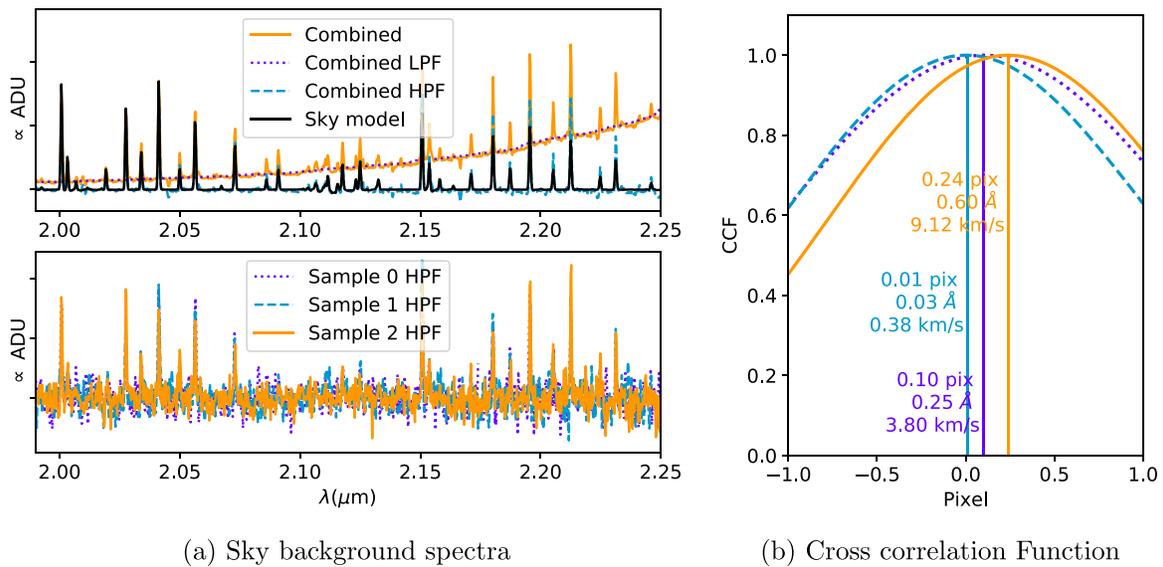
<sup>13</sup> <http://www.astropy.org>

<sup>14</sup> <https://matplotlib.org>

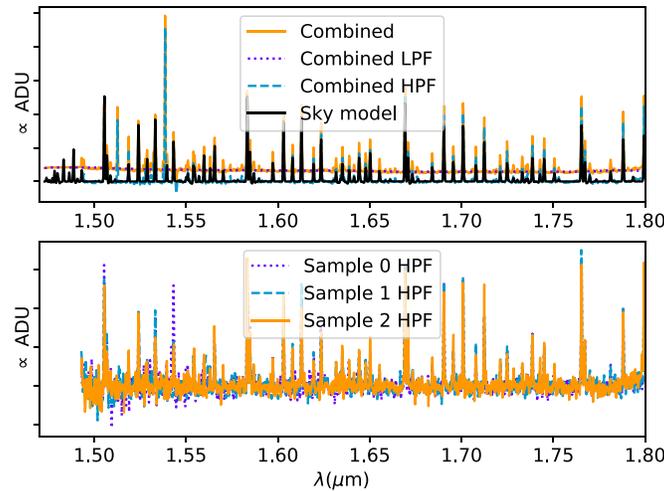
<sup>15</sup> <https://github.com/sblunt/orbitize>

<sup>16</sup> <https://github.com/willvouden/ptemcee>

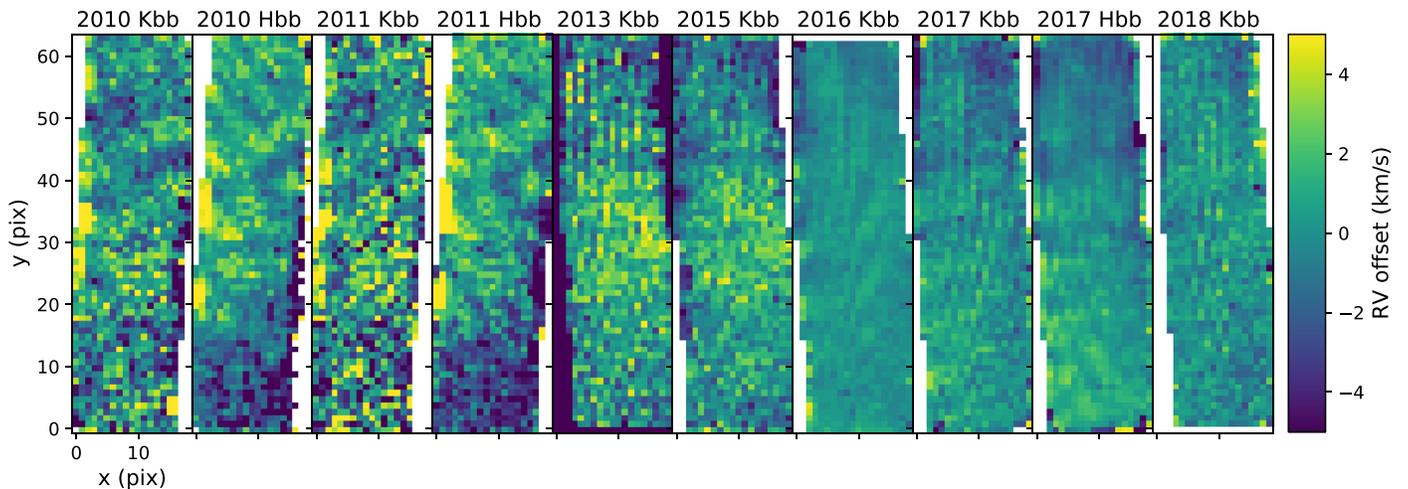
<sup>17</sup> <http://www.gemini.edu/sciops/telescopes-and-sites/observing-condition-constraints/ir-background-spectra>



**Figure 10.** Calibration of the wavelength solution from fitting of the OH emission lines in sky-background observations. The upper panel of (a) features the spatially averaged spectrum of a sky observation taken in the *K* band. It includes the original spectrum as well as its low-pass filtered (LPF) and high-pass filtered (HPF) components. The latter is compared to the high-pass filtered Earth atmosphere model from the Gemini observatory website. The lower panel of (a) shows three sample spectra of individual spaxels. The right panel (b) includes the cross-correlation function of the same three spectra with the model as well as the derived offsets.



**Figure 11.** Similar to Figure 10(a), but for the *H* band.



**Figure 12.** Spatial variation of the wavelength solution offset calculated for each year and spectral band.

**Table 3**  
Night-by-night Summary of the HR 8799 Observations with OSIRIS/Keck

Planet	Date	Band	Cubes	Time (hr)	Skies	RV Mean Offset (km s <sup>-1</sup> )	RV Residual Error (km s <sup>-1</sup> )	Notes
b	2009 Jul 22	Kbb	12	2.8	1	3.1		
	2009 Jul 23	Hbb	6	1.5	1	3.3		
	2009 Jul 30	Hbb	8	2.0				Cooling issue
	2009 Sep 3	Hbb	12	2.8	1	-5.3		
	2010 Jul 11	Kbb	9	2.2	2	6.5	2.2	
	2010 Jul 12	Kbb	9	2.2	1	5.3	2.2	
	2010 Jul 13	Hbb	2	0.5	1	2.9	1.3	
	2013 Jul 25	Kbb	16	2.7	2	2.7	1.9	
	2013 Jul 26	Kbb	9	1.5	3	3.4	1.9	
	2013 Jul 27	Kbb	7	1.2	2	3.1	1.9	
	2016 Nov 6	Kbb	2	0.3	2	6.9	0.8	
	2016 Nov 7	Kbb	3	0.5	2	6.9	0.8	
	2016 Nov 8	Kbb	1	0.2	1	6.9	0.8	
	2018 Jul 22	Kbb	6	0.5	1	-0.8	1.0	35 mas platescale
c	2010 Jul 15	Kbb	17	2.8	3	5.7	2.2	
	2010 Oct 28	Hbb	5	0.8	1	2.0	1.3	
	2010 Nov 04	Hbb	12	2.0	2	2.9	1.3	
		Kbb	8	1.3	2	5.7	2.2	
	2011 Jul 23	Kbb	12	2.0	3	3.4	3.0	
	2011 Jul 24	Hbb	12	2.0	3	0.8	1.5	
		Kbb	3	0.5	1	3.4	3.0	
		Hbb	15	2.5	3	0.8	1.5	
		Kbb	2	0.3	2	3.4	3.0	
		Hbb	1	0.2	3	3.4	1.9	
	2013 Jul 26	Kbb	1	0.2	3	3.4	1.9	
	2017 Nov 03	Hbb	14	2.3	2	-11.4	1.0	
		Kbb	3	0.5	2	-13.4	1.0	

**Note.** The RV mean offset corresponds to the spatially averaged wavelength calibration offset calculated from the OH<sup>-</sup> emission lines within a given night. The RV residual error is an estimated error of the spatially dependent wavelength offset relative to the spatial average (see Appendix B).

### Appendix C Transmission and PSF Calibration

In this section, we describe the calibration of the transmission profile of the atmosphere, the flux calibration, and the derivation of the PSF of the instrument, which are all based on reference star observations.

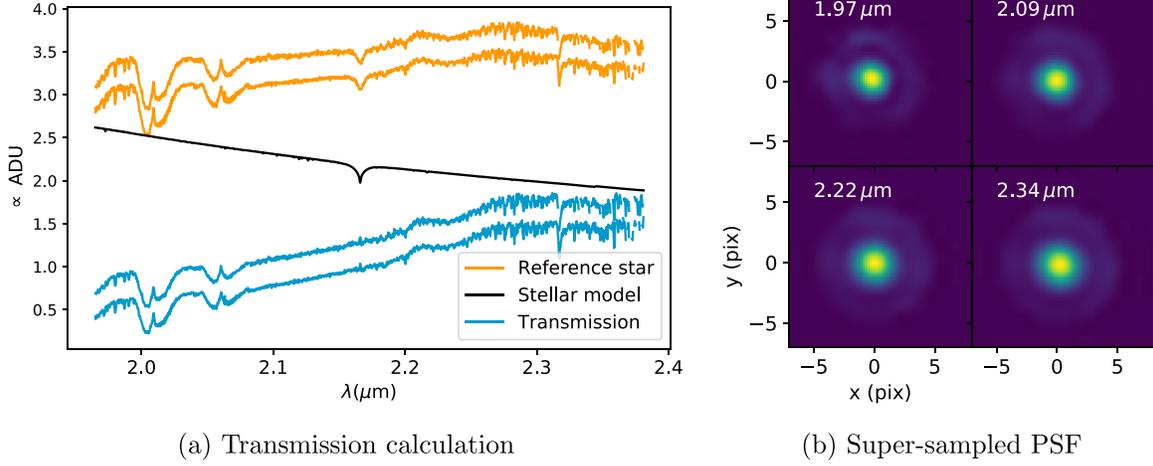
To avoid saturating the detector, the reference stars were sometimes acquired in open loop with no adaptive optics (AO) correction. Only the AO-corrected observations were used to derive the PSF and the flux calibration, while all were used to derive the transmission profile. Instead of a sky subtraction, we subtract each reference star observation in a pair thanks to the dithering pattern. Bad pixels are identified in the spectral direction using a 100 pixel box median high-pass filter and a 7MAD threshold. The neighboring pixels of such identified bad pixels are also marked as bad and subsequently masked. The centroid of the stellar PSF is calculated in each slice from a 2D Gaussian fit. Despite the atmospheric dispersion correction from the instrument, the centroid of the star varies by half a

pixel in the *H* band, which needs to be accounted for when modeling the planet signal. The reference center is chosen to be the median centroid of the star across the spectral band. At each wavelength, we build a spline interpolated model of the super-sampled PSF by combining the reference star observations for each spectral band on a nightly basis. An example of a super-sampled PSF is given in Figure 13. For the AO-corrected exposures, the star spectrum is then simply calculated by summing the flux in a five-pixel radius aperture. When the AO was turned off, we mask out the lower tenth percentile of pixels, replace the bad pixels by the median value in the given slice, and integrate the resulting cube in the spatial direction to derive the spectrum.

The transmission spectrum of the atmosphere and of the instrument is calculated on a nightly basis by dividing the reference star spectrum with a Phoenix stellar model (Figure 13; Husser et al. 2013).<sup>18</sup> The stellar models are broadened using the Python module PyAstronomy<sup>19</sup> to account for the spin of the star, convolved to OSIRIS resolution, and evaluated onto the data wavelength grid.

<sup>18</sup> <http://phoenix.astro.physik.uni-goettingen.de>

<sup>19</sup> <https://www.hs.uni-hamburg.de/DE/Ins/Per/Czesla/PyA/PyA/index.html>



**Figure 13.** Calibration of the transmission profile and super-sampled point-spread function (PSF) from reference star observations. (a) The left panel shows the spectra extracted from two different reference star observations in the  $K$  band, the Phoenix stellar model, and the resulting transmission spectrum. (b) The right panel shows the super-sampled PSF calculated for one night at four different wavelengths in the  $K$  band.

## Appendix D Statistics of Multivariate Linear Model

### D.1. Data Model

We define  $\mathbf{d}$  as the data vector of size  $N$  representing pixel values and  $\mathbf{n}$  as a Gaussian random vector with zero mean and covariance matrix  $\Sigma = s^2 \Sigma_0$ ;  $s^2$  being a scaling parameter. We define a linear model of the data with a matrix  $\mathbf{M}$  of size  $N \times N_\phi$ . The  $N_\phi$  linear parameters are denoted by the vector  $\phi$ . Additionally, we can assume that the model is a function of nonlinear parameters  $\psi$ , such that the data model can be written as

$$\mathbf{d} = \mathbf{M}_\psi \phi + \mathbf{n}. \quad (11)$$

More specifically, the linear parameters  $\phi$  represent the amplitude of the planet and the amplitude of the stellar light at any given pixel, while  $\psi$  includes parameters defining the atmospheric model of the planet, its RV, and its spin.

The corresponding likelihood is given by

$$\begin{aligned} \mathcal{L}(\psi, \phi, s^2) &= \mathcal{P}(\mathbf{d} | \psi, \phi, s), \\ &= \frac{1}{\sqrt{(2\pi)^N |\Sigma|}} \exp \left\{ -\frac{1}{2} (\mathbf{d} - \mathbf{M}_\psi \phi)^\top \Sigma^{-1} (\mathbf{d} - \mathbf{M}_\psi \phi) \right\}, \\ &= \frac{1}{\sqrt{(2\pi)^N |\Sigma_0| s^{2N}}} \\ &\quad \times \exp \left\{ -\frac{1}{2s^2} (\mathbf{d} - \mathbf{M}_\psi \phi)^\top \Sigma_0^{-1} (\mathbf{d} - \mathbf{M}_\psi \phi) \right\}. \end{aligned} \quad (12)$$

### D.2. Maximum Likelihood Estimate

The most likely value of  $\phi$ , noted  $\tilde{\phi}$ , for given values of  $s^2$  and  $\psi$ , is calculated by minimizing the negative log-likelihood. The subscript in  $\mathbf{M}_\psi$  will be dropped in this section, because  $\psi$  is assumed to be fixed.

$$\begin{aligned} \tilde{\phi} &= \text{Argmax}_\phi \mathcal{L}(\psi, \phi, s^2), \\ &= \text{Argmin}_\phi -\log \mathcal{L}(\psi, \phi, s^2). \end{aligned} \quad (13)$$

The negative log-likelihood is given by

$$\begin{aligned} -\log \mathcal{L} &= \frac{N}{2} \log(2\pi) + \frac{1}{2} \log(|\Sigma_0|) + \frac{N}{2} \log s^2 \\ &\quad + \frac{1}{2s^2} (\mathbf{d} - \mathbf{M}\phi)^\top \Sigma_0^{-1} (\mathbf{d} - \mathbf{M}\phi), \\ &= \frac{N}{2} \log(2\pi) + \frac{1}{2} \log(|\Sigma_0|) + \frac{N}{2} \log s^2 + \frac{1}{2s^2} \chi_0^2, \end{aligned} \quad (14)$$

where we defined  $\chi_0^2 = s^2 \chi^2$  as

$$\begin{aligned} \chi_0^2 &= (\mathbf{d} - \mathbf{M}\phi)^\top \Sigma_0^{-1} (\mathbf{d} - \mathbf{M}\phi), \\ &= \mathbf{d}^\top \Sigma^{-1} \mathbf{d} + \phi^\top \mathbf{M}^\top \Sigma^{-1} \mathbf{M} \phi - 2\phi^\top \mathbf{M}^\top \Sigma^{-1} \mathbf{d}. \end{aligned} \quad (15)$$

The solution to a linear  $\chi^2$  minimization problem is the pseudo-inverse.

$$\tilde{\phi} = (\mathbf{M}^\top \Sigma_0^{-1} \mathbf{M})^{-1} \mathbf{M}^\top \Sigma_0^{-1} \mathbf{d} \quad (16)$$

Indeed, let us solve for  $\nabla_\phi \chi_0^2 = 0$ .

$$\begin{aligned} \nabla_\phi \chi_0^2 = 0 &\Leftrightarrow 2\mathbf{M}^\top \Sigma_0^{-1} \mathbf{M} \tilde{\phi} - 2\mathbf{M}^\top \Sigma_0^{-1} \mathbf{d} = 0, \\ &\Leftrightarrow \mathbf{M}^\top \Sigma_0^{-1} \mathbf{M} \tilde{\phi} = \mathbf{M}^\top \Sigma_0^{-1} \mathbf{d}. \end{aligned} \quad (17)$$

Multiplying by the inverse of  $\mathbf{M}^\top \Sigma_0^{-1} \mathbf{M}$  returns Equation (16). In practice, that inversion is never performed because it is algorithmically much faster and more stable to solve the system of linear equations directly rather than inverting the matrix.

We can also jointly optimize for the covariance scaling factor  $s^2$ , which corresponds to the variance of the noise if  $\Sigma_0$  is the identity matrix. The optimization remains convex, so there is still a unique solution. The optimal  $s^2$  can be derived from minimizing the profile negative log-likelihood  $\mathcal{L}(\tilde{\phi}, s^2)$  as

$$\begin{aligned} \frac{d(-\log \mathcal{L})}{ds^2} \Big|_{\phi=\tilde{\phi}} = 0 &\Leftrightarrow \frac{N}{2s^2} - \frac{1}{2s^4} \chi_{0,\phi=\tilde{\phi}}^2 = 0, \\ &\Leftrightarrow \tilde{s}^2 = \frac{1}{N} \chi_{0,\phi=\tilde{\phi}}^2. \end{aligned} \quad (18)$$

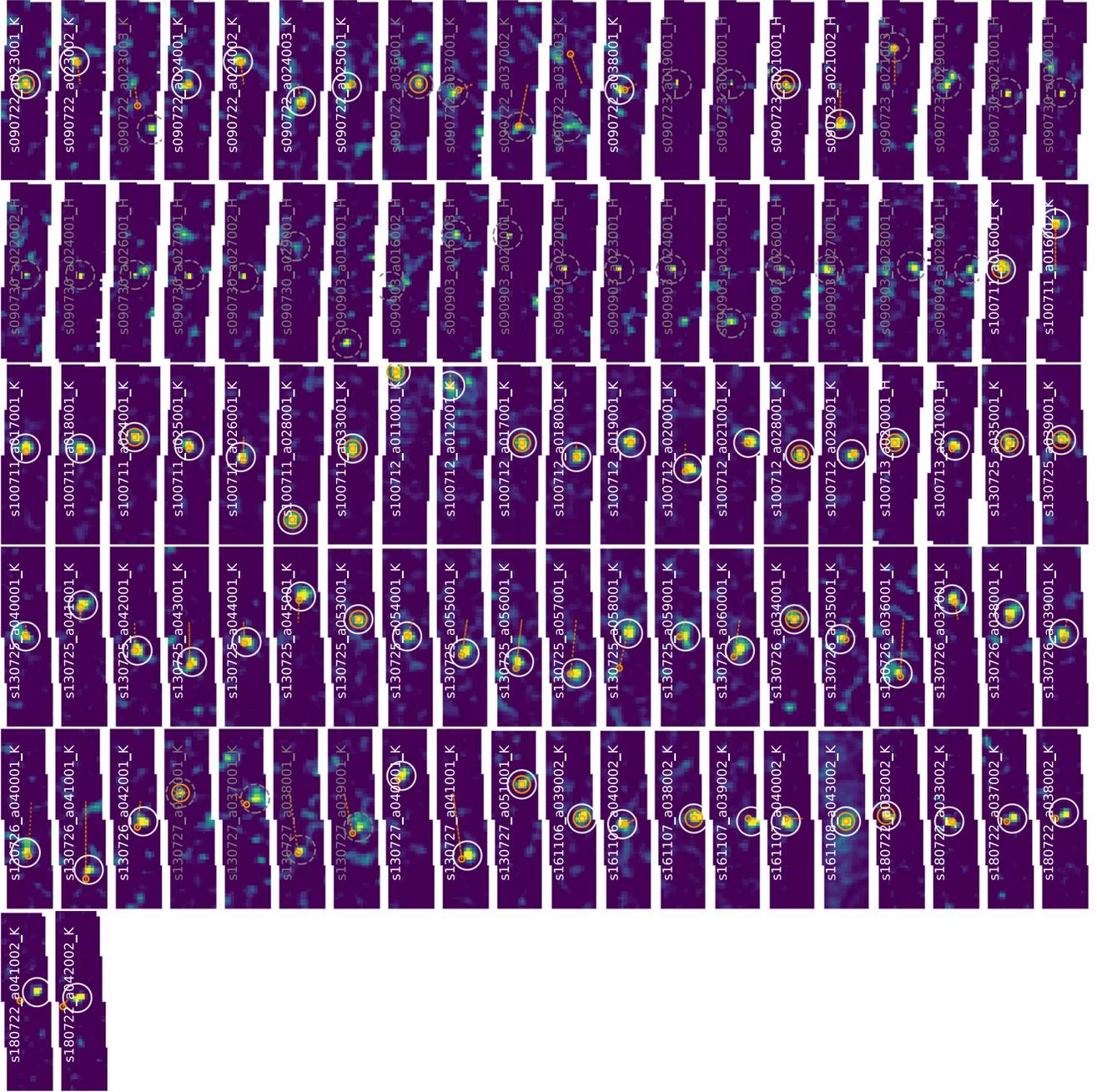


Figure 14. Individual exposure detection maps for HR 8799 b in the *Hor K* band.

### D.3. Covariance of the Linear Parameters

In this section, we will calculate the covariance of the estimated linear parameters  $\tilde{\phi}$ , which is given in Equation (16)

$$\begin{aligned}\tilde{\phi} &= (\mathbf{M}^\top \Sigma_0^{-1} \mathbf{M})^{-1} \mathbf{M}^\top \Sigma_0^{-1} \mathbf{d}, \\ &= (\mathbf{M}^\top \Sigma_0^{-1} \mathbf{M})^{-1} \mathbf{M}^\top \Sigma_0^{-1} (\mathbf{M} \phi_{\text{true}} + \mathbf{n}).\end{aligned}\quad (19)$$

In the second line, we replaced the data vector by its signal and noise components.

The linear transformation of a Gaussian random vector remains Gaussian, so the posterior of  $\tilde{\phi}$  must be Gaussian too. Indeed, if  $\mathbf{X} \sim \mathcal{N}(\boldsymbol{\mu}_x, \Sigma_x)$ , then  $\mathbf{Y} = \mathbf{A}\mathbf{X}$  also follows a normal distribution with vector mean

$$\boldsymbol{\mu}_y = \mathbf{A}\boldsymbol{\mu}_x \quad (20)$$

and covariance matrix

$$\Sigma_y = \mathbf{A}\Sigma_x\mathbf{A}^\top. \quad (21)$$

The proof is as follows:

$$\begin{aligned}\Sigma_y &= E((\mathbf{y} - \boldsymbol{\mu}_y)(\mathbf{y} - \boldsymbol{\mu}_y)^\top), \\ &= E(\mathbf{y}\mathbf{y}^\top - \boldsymbol{\mu}_y\mathbf{y}^\top - \mathbf{y}\boldsymbol{\mu}_y^\top + \boldsymbol{\mu}_y\boldsymbol{\mu}_y^\top), \\ &= E(\mathbf{y}\mathbf{y}^\top) - \boldsymbol{\mu}_y\boldsymbol{\mu}_y^\top, \\ &= E((\mathbf{A}\mathbf{x})(\mathbf{A}\mathbf{x})^\top) - (\mathbf{A}\boldsymbol{\mu}_x)(\mathbf{A}\boldsymbol{\mu}_x)^\top, \\ &= \mathbf{A}E(\mathbf{x}^\top\mathbf{x})\mathbf{A}^\top - \mathbf{A}\boldsymbol{\mu}_x^\top\boldsymbol{\mu}_x\mathbf{A}^\top, \\ &= \mathbf{A}\Sigma_x\mathbf{A}^\top.\end{aligned}\quad (22)$$

As a result, the covariance matrix of  $\tilde{\phi}$  is given by

$$\begin{aligned} \text{cov}(\tilde{\phi}) &= [(\mathbf{M}^\top \Sigma_0^{-1} \mathbf{M})^{-1} \mathbf{M}^\top \Sigma_0^{-1}] s^2 \Sigma_0 \\ &\quad \times [(\mathbf{M}^\top \Sigma_0^{-1} \mathbf{M})^{-1} \mathbf{M}^\top \Sigma_0^{-1}]^\top, \\ &= s^2 (\mathbf{M}^\top \Sigma_0^{-1} \mathbf{M})^{-1} \mathbf{M}^\top \Sigma_0^{-1} \Sigma_0 \Sigma_0^{-1} \mathbf{M} (\mathbf{M}^\top \Sigma_0^{-1} \mathbf{M})^{-1}, \\ &= s^2 (\mathbf{M}^\top \Sigma_0^{-1} \mathbf{M})^{-1} \mathbf{M}^\top \Sigma_0^{-1} \mathbf{M} (\mathbf{M}^\top \Sigma_0^{-1} \mathbf{M})^{-1}, \\ &= s^2 (\mathbf{M}^\top \Sigma_0^{-1} \mathbf{M})^{-1}. \end{aligned} \quad (23)$$

Another way to show this result is to consider the likelihood, and write  $\phi = \tilde{\phi} + \Delta\phi$  as

$$\begin{aligned} L &= \frac{1}{\sqrt{(2\pi)^N |\Sigma_0| s^{2N}}} \\ &\quad \times \exp \left\{ -\frac{1}{2s^2} (\mathbf{d} - \mathbf{M}\phi)^\top \Sigma_0^{-1} (\mathbf{d} - \mathbf{M}\phi) \right\}, \\ &= \frac{1}{\sqrt{(2\pi)^N |\Sigma_0| s^{2N}}} \exp \left\{ -\frac{1}{2s^2} (\mathbf{d} - \mathbf{M}\tilde{\phi} - \mathbf{M}\Delta\phi)^\top \right. \\ &\quad \left. \times \Sigma_0^{-1} (\mathbf{d} - \mathbf{M}\tilde{\phi} - \mathbf{M}\Delta\phi) \right\}, \\ &= \frac{1}{\sqrt{(2\pi)^N |\Sigma_0| s^{2N}}} \\ &\quad \times \exp \left\{ -\frac{1}{2s^2} (\mathbf{d} - \mathbf{M}\tilde{\phi})^\top \Sigma_0^{-1} (\mathbf{d} - \mathbf{M}\tilde{\phi}) \right\} \\ &\quad \times \exp \left\{ -\frac{1}{2s^2} (\mathbf{M}\Delta\phi)^\top \Sigma_0^{-1} (\mathbf{M}\Delta\phi) \right\} \\ &\quad \times \exp \left\{ +\frac{1}{s^2} (\mathbf{d} - \mathbf{M}\tilde{\phi})^\top \Sigma_0^{-1} (\mathbf{M}\Delta\phi) \right\}, \\ &= \frac{1}{\sqrt{(2\pi)^N |\Sigma_0| s^{2N}}} \exp \left\{ -\frac{1}{2s^2} \chi_{0,\phi=\tilde{\phi}}^2 \right\} \\ &\quad \times \exp \left\{ -\frac{1}{2s^2} \Delta\phi^\top (\mathbf{M}^\top \Sigma_0^{-1} \mathbf{M}) \Delta\phi \right\}. \end{aligned} \quad (24)$$

In the third equality, we used the fact that the argument of last exponential term vanishes because

$$\begin{aligned} \nabla_\phi \chi_0^2 = 0 &\Leftrightarrow 2\mathbf{M}^\top \Sigma_0^{-1} \mathbf{M}\tilde{\phi} - 2\mathbf{M}^\top \Sigma_0^{-1} \mathbf{d} = 0, \\ &\Leftrightarrow \mathbf{M}^\top \Sigma_0^{-1} (\mathbf{d} - \mathbf{M}\tilde{\phi}) = 0, \\ &\Leftrightarrow (\mathbf{d} - \mathbf{M}\tilde{\phi})^\top \Sigma_0^{-1} \mathbf{M} = 0. \end{aligned} \quad (25)$$

Therefore,

$$\mathcal{L} \propto \exp \left\{ -\frac{1}{2s^2} \Delta\phi^\top (\mathbf{M}^\top \Sigma_0^{-1} \mathbf{M}) \Delta\phi \right\}, \quad (26)$$

which shows that the posterior of  $\phi$  with uniform priors is Gaussian with covariance

$$\text{cov}(\tilde{\phi}) = s^2 (\mathbf{M}^\top \Sigma_0^{-1} \mathbf{M})^{-1}. \quad (27)$$

#### D.4. Signal-to-noise Ratio

If the planet signal corresponds to the first column of the model  $\phi$ , which is assumed to be case, then the standard deviation is the square root of the first diagonal element of

$\text{cov}(\tilde{\phi})$ . So the theoretical S/N of the planet can be written as

$$\mathcal{S} = \frac{\tilde{\phi}[0]}{\sqrt{\text{cov}(\tilde{\phi})[0, 0]}}. \quad (28)$$

This expression includes a marginalization over the all the other linear parameters of  $\phi$ .

#### D.5. Planet Detection

We first define two hypotheses,  $\mathcal{H}_0$  and  $\mathcal{H}_1$ , which correspond to the null hypothesis and the planet hypothesis, respectively,

$$\mathbf{d} = \mathbf{M}_{\psi_0} \phi_0 + \mathbf{n} \quad \text{and} \quad \mathbf{d} = \mathbf{M}_{\psi_1} \phi_1 + \mathbf{n}. \quad (29)$$

Assuming the null hypothesis, the only signal in the data is the diffracted starlight from the star (i.e., speckles). The planet case is simply modeled from the null hypothesis by adding one extra parameter to model the planet signal. As a result,  $\mathbf{M}_{\psi_1}$  has one extra column, which includes the spectrum and the PSF of the planet, compared to  $\mathbf{M}_{\psi_0}$ .

When jointly fitting for  $\phi$  and  $s^2$ , the minimized negative log-likelihood is given by

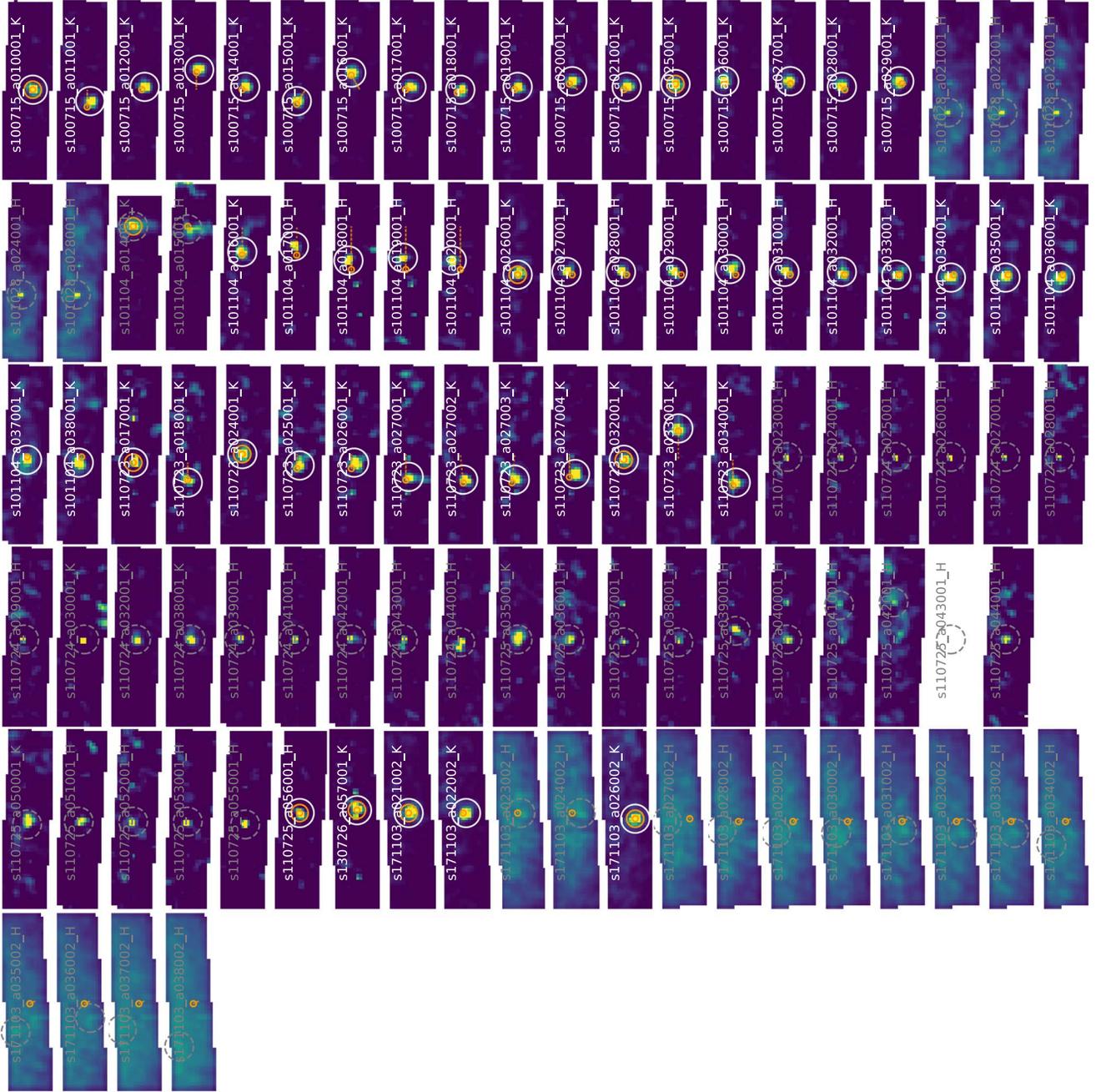
$$\begin{aligned} -\log \mathcal{L}(\tilde{\phi}, \tilde{s}^2) &= \frac{N}{2} \log(2\pi) + \frac{1}{2} \log(|\Sigma_0|) + \frac{N}{2} \log \tilde{s}^2 \\ &\quad + \frac{1}{2\tilde{s}^2} \chi_{0,\phi=\tilde{\phi}}^2, \\ &= \frac{N}{2} \log(2\pi) + \frac{1}{2} \log(|\Sigma_0|) + \frac{N}{2} \log \tilde{s}^2 \\ &\quad + \frac{1}{2\tilde{s}^2} \chi_{0,\phi=\tilde{\phi}}^2, \\ -2 \log \mathcal{L}(\tilde{\phi}, \tilde{s}^2) &= N \log(2\pi) + \log|\Sigma_0| + N \log \chi_{0,\phi=\tilde{\phi}}^2 \\ &\quad - N \log N + N. \end{aligned} \quad (30)$$

There are different methods to estimate the likelihood of one model compared to another. A few common examples are: likelihood ratio, the Akaike information criterion (AIC), or the Bayesian information criterion (BIC). AIC and BIC include a penalty for extra model parameters. This can be useful if one wants to model the planet signal as a linear combination of molecular templates for example. However the penalty for extra parameters is a constant, which means that for our simple planet detection scheme, a likelihood ratio is satisfactory.

Figures 14 and 15 show the individual exposure detection maps for HR 8799 b and c, respectively. The corresponding filename is annotated on each frame. The white solid circles indicate the selected frames which have been used in the analyses of Section 5. The grayed dashed circles represent nondetections or detections suffering from an image artifact. The pointing offsets calculated from the fits file header keywords are drawn with orange lines. The reference pointing, marked as a double orange circle, is the chosen to be the first detection in the given sequence of observations.

For the sake of completeness, we here remind the definition of both the AIC and BIC, which is based on the maximized likelihood

$$\text{AIC} = 2N_\phi - 2 \log \mathcal{L}(\tilde{\phi}, \tilde{s}^2), \quad (31)$$



**Figure 15.** Individual exposure detection maps for HR 8799c in the H or K band.

where  $N_\phi$  the number of parameters in  $\phi$ . The smaller the AIC, the more likely the model is.

$$\text{BIC} = \log(N_d)N_\phi - 2 \log \mathcal{L}(\tilde{\phi}, \tilde{\sigma}^2), \quad (32)$$

The model comparison consists in calculating the difference in the criterion for two models. In the case of AIC, we have,

$$\Delta \text{AIC} = \text{AIC}_0 - \text{AIC}_1 \quad (33)$$

with the index 0 or 1 corresponding to  $\mathcal{H}_0$  and  $\mathcal{H}_1$ , respectively. The probability of one hypothesis compared to another is proportional to  $\exp(\Delta \text{AIC}/2)$ . For example, if  $\Delta \text{AIC} = 4.6$ ,

then  $\mathcal{H}_1$  is 10 times more likely than  $\mathcal{H}_0$ . In practice, the mismatch between the models and the data can make the interpretation more challenging.

#### D.6. Marginalizing Over Linear Parameters

We will show how to derive the posterior of the nonlinear parameters  $\psi$  marginalized over  $\phi$ .  $\phi$  is here seen as a background parameter. For example, it can be used to calculate the posterior of the RV estimate of the planet, while marginalizing over the planet flux and the starlight subtraction.

The marginalized posterior of  $\psi$  is given by

$$\begin{aligned} \mathcal{P}(\psi|\mathbf{d}) &= \int_s \int_\phi \mathcal{P}(\psi, \phi, s|\mathbf{d}), \\ &= \frac{1}{\mathcal{P}(\mathbf{d})} \int_s \int_\phi \mathcal{P}(\mathbf{d}|\psi, \phi, s) \mathcal{P}(\psi, \phi, s), \\ &= \frac{1}{\mathcal{P}(\mathbf{d})} \int_s \int_\phi \mathcal{P}(\mathbf{d}|\psi, \phi, s) \mathcal{P}(\psi) \mathcal{P}(\phi) \mathcal{P}(s), \\ &= \frac{\mathcal{P}(\psi)}{\mathcal{P}(\mathbf{d})} \int_s \mathcal{P}(s) \int_\phi \mathcal{P}(\mathbf{d}|\psi, \phi, s) \mathcal{P}(\phi). \end{aligned} \quad (34)$$

In Equation (24), we showed that the likelihood can be written as

$$\begin{aligned} P(\mathbf{d}|\psi, \phi, s) &= \frac{1}{\sqrt{(2\pi)^{N_d} |\Sigma_0| s^{2N_d}}} \exp \left\{ -\frac{1}{2s^2} \chi_{0, \phi = \bar{\phi}, \psi}^2 \right\} \\ &\quad \times \exp \left\{ -\frac{1}{2s^2} \Delta \phi^\top (\mathbf{M}_\psi^\top \Sigma_0^{-1} \mathbf{M}_\psi) \Delta \phi \right\}. \end{aligned} \quad (35)$$

Assuming unbound uniform improper priors on the elements of  $\phi$ , we have

$$\begin{aligned} &\int_\phi P(\mathbf{d}|\psi, \phi, s^2) P(\phi) \\ &= \frac{1}{\sqrt{(2\pi)^{N_d} |\Sigma_0| s^{2N_d}}} \exp \left\{ -\frac{1}{2s^2} \chi_{0, \phi = \bar{\phi}, \psi}^2 \right\} \\ &\quad \times \int_{\Delta \phi} \exp \left\{ -\frac{1}{2s^2} \Delta \phi^\top (\mathbf{M}_\psi^\top \Sigma_0^{-1} \mathbf{M}_\psi) \Delta \phi \right\}, \\ &= \sqrt{\frac{(2\pi)^{N_\phi} s^{2N_\phi}}{(2\pi)^{N_d} s^{2N_d}}} \frac{1}{\sqrt{|\Sigma_0| \times |\mathbf{M}_\psi^\top \Sigma_0^{-1} \mathbf{M}_\psi|}} \\ &\quad \times \exp \left\{ -\frac{1}{2s^2} \chi_{0, \phi = \bar{\phi}, \psi}^2 \right\}. \end{aligned} \quad (36)$$

Then, we can integrate over  $s$ , assuming improper prior  $\mathcal{P}(s) \propto 1/s^\alpha$

$$\begin{aligned} \int_s \mathcal{P}(s) \int_\phi P(\mathbf{d}|\psi, \phi, s^2) P(\phi) &\propto \frac{1}{\sqrt{|\Sigma_0| \times |\mathbf{M}_\psi^\top \Sigma_0^{-1} \mathbf{M}_\psi|}} \\ &\quad \times \int_s \frac{1}{s^{N_d - N_\phi + \alpha}} \exp \left\{ -\frac{1}{2s^2} \chi_{0, \phi = \bar{\phi}, \psi}^2 \right\}. \end{aligned} \quad (37)$$

We then use the identity ( $N > 1$ ,  $a > 0$ ),

$$\int_{x=0}^{\infty} \frac{1}{x^N} \exp \left\{ -\frac{a}{x^2} \right\} = \frac{1}{2a^{(N-1)/2}} \Gamma \left[ \frac{N-1}{2} \right]. \quad (38)$$

We derive that,

$$\begin{aligned} &\int_s \mathcal{P}(s) \int_\phi \mathcal{P}(\mathbf{d}|\psi, \phi, s^2) \mathcal{P}(\phi) \\ &\propto \frac{1}{\sqrt{|\Sigma_0| \times |\mathbf{M}_\psi^\top \Sigma_0^{-1} \mathbf{M}_\psi|}} \left( \frac{1}{\chi_{0, \phi = \bar{\phi}, \psi}^2} \right)^{\frac{N_d - N_\phi + \alpha - 1}{2}}, \end{aligned} \quad (39)$$

and therefore conclude that,

$$\mathcal{P}(\psi|\mathbf{d}) \propto \frac{\mathcal{P}(\psi)}{\sqrt{|\Sigma_0| \times |\mathbf{M}_\psi^\top \Sigma_0^{-1} \mathbf{M}_\psi|}} \left( \frac{1}{\chi_{0, \phi = \bar{\phi}, \psi}^2} \right)^{\frac{N_d - N_\phi + \alpha - 1}{2}}. \quad (40)$$

For numerical calculation, taking the logarithm gives the alternative expression

$$\begin{aligned} \log \mathcal{P}(\psi|\mathbf{d}) &= \text{Cst} + \log \mathcal{P}(\psi) - \frac{1}{2} \log (|\Sigma_0| \times |\mathbf{M}_\psi^\top \Sigma_0^{-1} \mathbf{M}_\psi|) \\ &\quad - \frac{N_d - N_\phi + \alpha - 1}{2} \log \chi_{0, \phi = \bar{\phi}, \psi}^2. \end{aligned} \quad (41)$$

Note that the minimized chi-square can be written as

$$\begin{aligned} \chi_{0, \phi = \bar{\phi}, \psi}^2 &= (\mathbf{d} - \mathbf{M}\bar{\phi})^\top \Sigma_0^{-1} (\mathbf{d} - \mathbf{M}\bar{\phi}), \\ &= \mathbf{d}^\top \Sigma_0^{-1} \mathbf{d} + \bar{\phi}^\top \mathbf{M}^\top \Sigma_0^{-1} \mathbf{M} \bar{\phi} - 2\mathbf{d}^\top \Sigma_0^{-1} \mathbf{M} \bar{\phi}, \\ &= \mathbf{d}^\top \Sigma_0^{-1} \mathbf{d} + [\bar{\phi}^\top \mathbf{M}^\top \Sigma_0^{-1} \mathbf{M} - \mathbf{d}^\top \Sigma_0^{-1} \mathbf{M}] \bar{\phi} \\ &\quad - \mathbf{d}^\top \Sigma_0^{-1} \mathbf{M} \bar{\phi}, \\ &= \mathbf{d}^\top \Sigma_0^{-1} \mathbf{d} - \mathbf{d}^\top \Sigma_0^{-1} \mathbf{M} \bar{\phi}, \\ &= \mathbf{d}^\top \Sigma_0^{-1} \mathbf{d} - \mathbf{d}^\top \Sigma_0^{-1} \mathbf{M} (\mathbf{M}^\top \Sigma_0^{-1} \mathbf{M})^{-1} \mathbf{M}^\top \Sigma_0^{-1} \mathbf{d}, \end{aligned} \quad (42)$$

where we used the fact that the derivative of the chi-square with respect to  $\phi$  vanishes, such that

$$\frac{d\chi^2}{d\phi} = \bar{\phi}^\top \mathbf{M}^\top \Sigma_0^{-1} \mathbf{M} - \mathbf{d}^\top \Sigma_0^{-1} \mathbf{M} = 0. \quad (43)$$

## ORCID iDs

Jean-Baptiste Ruffio  <https://orcid.org/0000-0003-2233-4821>

Bruce Macintosh  <https://orcid.org/0000-0003-1212-7538>

Quinn M. Konopacky  <https://orcid.org/0000-0002-9936-6285>

Travis Barman  <https://orcid.org/0000-0002-7129-3002>

Robert J. De Rosa  <https://orcid.org/0000-0002-4918-0247>

Jason J. Wang  <https://orcid.org/0000-0003-0774-6502>

Kielan K. Wilcomb  <https://orcid.org/0000-0002-9803-8255>

Ian Czekala  <https://orcid.org/0000-0002-1483-8811>

Christian Marois  <https://orcid.org/0000-0002-4164-4182>

## References

- Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, *A&A*, **558**, A33
- Baines, E. K., White, R. J., Huber, D., et al. 2012, *ApJ*, **761**, 57
- Baraffe, I., Chabrier, G., Barman, T. S., Allard, F., & Hauschildt, P. H. 2003, *A&A*, **402**, 701
- Barman, T. S., Konopacky, Q. M., Macintosh, B., & Marois, C. 2015, *ApJ*, **804**, 61
- Barman, T. S., Macintosh, B., Konopacky, Q. M., & Marois, C. 2011, *ApJ*, **733**, 65
- Bell, C. P. M., Mamajek, E. E., & Naylor, T. 2015, *MNRAS*, **454**, 593
- Birkby, J. L., de Kok, R. J., Brogi, M., et al. 2013, *MNRAS*, **436**, L35
- Birkby, J. L., de Kok, R. J., Brogi, M., Schwarz, H., & Snellen, I. A. G. 2017, *AJ*, **153**, 138
- Blunt, S., Ngo, H., Wang, J., et al. 2019, sblunt/orbitize: API Update to Prepare for RV+Astrometry Fits, Zenodo, doi:10.5281/zenodo.3337378
- Boehle, A., Larkin, J. E., Adkins, S. M., et al. 2016, *Proc. SPIE*, **9908**, 99082Q
- Booth, M., Jordán, A., Casassus, S., et al. 2016, *MNRAS*, **460**, L10
- Brogi, M., de Kok, R. J., Albrecht, S., et al. 2016, *ApJ*, **817**, 106

- Brogi, M., de Kok, R. J., Birkby, J. L., Schwarz, H., & Snellen, I. A. G. 2014, *A&A*, **565**, A124
- Brogi, M., Line, M., Bean, J., Désert, J. M., & Schwarz, H. 2017, *ApJL*, **839**, L2
- Brogi, M., & Line, M. R. 2019, *AJ*, **157**, 114
- Brogi, M., Snellen, I. A. G., de Kok, R. J., et al. 2012, *Natur*, **486**, 502
- Brogi, M., Snellen, I. A. G., de Kok, R. J., et al. 2013, *ApJ*, **767**, 27
- Bryan, M. L., Benneke, B., Knutson, H. A., Batygin, K., & Bowler, B. P. 2018, *NatAs*, **2**, 138
- Cantaloube, F., Mouillet, D., Mugnier, L. M., et al. 2015, *A&A*, **582**, A89
- de Kok, R. J., Brogi, M., Snellen, I. A. G., et al. 2013, *A&A*, **554**, A82
- Fabrycky, D. C., & Murray-Clay, R. A. 2010, *ApJ*, **710**, 1408
- Follert, R., Dorn, R. J., Oliva, E., et al. 2014, *Proc. SPIE*, **9147**, 914719
- Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, *PASP*, **125**, 306
- Gaia Collaboration 2018, *yCat*, **1345**, 0
- Gontcharov, G. A. 2006, *AstL*, **32**, 759
- Haffert, S. Y., Bohn, A. J., de Boer, J., et al. 2019, *NatAs*, **3**, 749
- Hoeijmakers, H. J., Ehrenreich, D., Heng, K., et al. 2018, *Natur*, **560**, 453
- Hughes, A. M., Wilner, D. J., Andrews, S. M., et al. 2011, *ApJ*, **740**, 38
- Hunter, J. D. 2007, *CSE*, **9**, 90
- Husser, T. O., Wende-von Berg, S., Dreizler, S., et al. 2013, *A&A*, **553**, A6
- Kaeufel, H.-U., Ballester, P., Biereichel, P., et al. 2004, *Proc. SPIE*, **5492**, 1218
- Kanodia, S., & Wright, J. 2018a, *RNAAS*, **2**, 4
- Kanodia, S., & Wright, J. T. 2018b, Barycorrpy: Barycentric velocity calculation and leap second management, Astrophysics Source Code Library, ascl:1808.001
- Kass, R. E., & Raftery, A. E. 1995, *J. Am. Stati. Assoc.*, **90**, 773
- Konopacky, Q. M., Barman, T. S., Macintosh, B. A., & Marois, C. 2013, *Sci*, **339**, 1398
- Konopacky, Q. M., Marois, C., Macintosh, B. A., et al. 2016, *AJ*, **152**, 28
- Krabbe, A., Gasaway, T., Song, I., et al. 2004, *Proc. SPIE*, **5492**, 1403
- Gravity Collaboration, Lacour, S., Nowak, M., et al. 2019, *A&A*, **623**, L11
- Larkin, J., Barczys, M., Krabbe, A., et al. 2006, *NewAR*, **50**, 362
- Lockhart, K. E., Do, T., Larkin, J. E., et al. 2019, *AJ*, **157**, 75
- Lockwood, A. C., Johnson, J. A., Bender, C. F., et al. 2014, *ApJL*, **783**, L29
- Lord, S. D. 1992, NASA Technical Memorandum, **103957**
- Lyke, J., Do, T., Boehle, A., et al. 2017, OSIRIS Toolbox: OH-Suppressing InfraRed Imaging Spectrograph pipeline, Astrophysics Source Code Library, ascl:1710.021
- Marois, C., Macintosh, B., Barman, T., et al. 2008, *Sci*, **322**, 1348
- Marois, C., Zuckerman, B., Konopacky, Q. M., Macintosh, B., & Barman, T. 2010, *Natur*, **468**, 1080
- Martin, E. C., Fitzgerald, M. P., McLean, I. S., et al. 2014, *Proc. SPIE*, **9147**, 914781
- Mathews, B., Kennedy, G., Sibthorpe, B., et al. 2014, *ApJ*, **780**, 97
- Mawet, D., Delorme, J. R., Jovanovic, N., et al. 2017, *Proc. SPIE*, **10400**, 1040029
- Metchev, S. A., Heinze, A., Apai, D., et al. 2015, *ApJ*, **799**, 154
- Mieda, E., Wright, S. A., Larkin, J. E., et al. 2014, *PASP*, **126**, 250
- Nugroho, S. K., Kawahara, H., Masuda, K., et al. 2017, *AJ*, **154**, 221
- O'Neil, K. K., Martinez, G. D., Hees, et al. 2019, *AJ*, **158**, 4
- Petit dit de la Roche, D. J. M., Hoeijmakers, H. J., & Snellen, I. A. G. 2018, *A&A*, **616**, A146
- Pueyo, L., Soummer, R., Hoffmann, J., et al. 2015, *ApJ*, **803**, 31
- Reidemeister, M., Krivov, A. V., Schmidt, T. O. B., et al. 2009, *A&A*, **503**, 247
- Rodler, F., Lopez-Morales, M., & Ribas, I. 2012, *ApJL*, **753**, L25
- Ruffio, J.-B., Macintosh, B., Wang, J. J., et al. 2017, *ApJ*, **842**, 14
- Salz, M., Czesla, S., Schneider, P. C., et al. 2018, *A&A*, **620**, A97
- Schwarz, H., Brogi, M., de Kok, R., Birkby, J., & Snellen, I. 2015, *A&A*, **576**, A111
- Schwarz, H., Ginski, C., de Kok, R. J., et al. 2016, *A&A*, **593**, A74
- Snellen, I. A. G., Brandl, B. R., de Kok, R. J., et al. 2014, *Natur*, **509**, 63
- Snellen, I. A. G., de Kok, R. J., de Mooij, E. J. W., & Albrecht, S. 2010, *Natur*, **465**, 1049
- Soummer, R., Hagan, J. B., Pueyo, L., et al. 2011, *ApJ*, **741**, 55
- Su, K. Y. L., Rieke, G. H., Stapelfeldt, K. R., et al. 2009, *ApJ*, **705**, 314
- Teachey, A., & Kipping, D. M. 2018, *SciA*, **4**, eaav1784
- Vanderburg, A., Rappaport, S. A., & Mayo, A. W. 2018, *AJ*, **156**, 184
- Vousden, W. D., Farr, W. M., & Mandel, I. 2016, *MNRAS*, **455**, 1919
- Wang, J., Mawet, D., Fortney, J. J., et al. 2018a, *AJ*, **156**, 272
- Wang, J. J., Graham, J. R., Dawson, R., et al. 2018b, *AJ*, **156**, 192
- Wertz, O., Absil, O., Gómez González, C. A., et al. 2017, *A&A*, **598**, A83
- Wilner, D. J., MacGregor, M. A., Andrews, S. M., et al. 2018, *ApJ*, **855**, 56
- Zucker, S. 2003, *MNRAS*, **342**, 1291
- Zuckerman, B., Rhee, J. H., Song, I., & Bessell, M. S. 2011, *ApJ*, **732**, 61
- Zurlo, A., Vigan, A., Galicher, R., et al. 2016, *A&A*, **587**, A57