

## Neutral Iron Emission Lines From The Day-side Of KELT-9b The GAPS Programme With HARPS-N At TNG XX

LORENZO PINO,<sup>1</sup> JEAN-MICHEL DÉSSERT,<sup>1</sup> MATTEO BROGI,<sup>2,3,4</sup> LUCA MALAVOLTA,<sup>5,6</sup> AURÉLIEN WYTTEBACH,<sup>7</sup>  
MICHAEL LINE,<sup>8</sup> JENS HOEIJMAKERS,<sup>9,10</sup> LUCA FOSSATI,<sup>11</sup> ALDO STEFANO BONOMO,<sup>3</sup> VALERIO NASCIMBENI,<sup>12</sup>  
VATSAL PANWAR,<sup>1</sup> LAURA AFFER,<sup>13</sup> SERENA BENATTI,<sup>13</sup> KATIA BIAZZO,<sup>6</sup> ANDREA BIGNAMINI,<sup>14</sup> FRANCESCO BORSA,<sup>15</sup>  
ILARIA CARLEO,<sup>16,12</sup> RICCARDO CLAUDI,<sup>12</sup> ROSARIO COSENTINO,<sup>17</sup> ELVIRA COVINO,<sup>18</sup> MARIO DAMASSO,<sup>3</sup>  
SILVANO DESIDERA,<sup>12</sup> PAOLO GIACOBBE,<sup>3</sup> AVET HARUTYUNYAN,<sup>17</sup> ANTONINO FRANCESCO LANZA,<sup>6</sup> GIUSEPPE LETO,<sup>6</sup>  
ANTONIO MAGGIO,<sup>13</sup> JESUS MALDONADO,<sup>13</sup> LUIGI MANCINI,<sup>19,3,20</sup> GIUSEPPINA MICELA,<sup>13</sup> EMILIO MOLINARI,<sup>21</sup>  
ISABELLA PAGANO,<sup>6</sup> GIAMPAOLO PIOTTO,<sup>5</sup> ENNIO PORETTI,<sup>17,15</sup> MONICA RAINER,<sup>22</sup> GAETANO SCANDARIATO,<sup>6</sup>  
ALESSANDRO SOZZETTI,<sup>3</sup> ROMAIN ALLART,<sup>10</sup> LUCA BORSATO,<sup>5</sup> GIOVANNI BRUNO,<sup>6</sup> LUCA DI FABRIZIO,<sup>17</sup>  
DAVID EHRENREICH,<sup>10</sup> ALDO FIOREZZANO,<sup>17</sup> GIUSEPPE FRUSTAGLI,<sup>15,23</sup> BAPTISTE LAVIE,<sup>10</sup> CHRISTOPHE LOVIS,<sup>10</sup>  
ANTONIO MAGAZZÙ,<sup>17</sup> DOMENICO NARDIELLO,<sup>5</sup> MARCO PEDANI,<sup>17</sup> AND RICCARDO SMAREGLIA<sup>14</sup>

<sup>1</sup>*Anton Pannekoek Institute for Astronomy, University of Amsterdam  
Science Park 904*

*1098 XH Amsterdam, The Netherlands*

<sup>2</sup>*Department of Physics, University of Warwick  
Coventry CV4 7AL, UK*

<sup>3</sup>*INAF - Osservatorio Astrofisico di Torino  
Via Osservatorio 20*

*I-10025, Pino Torinese, Italy*

<sup>4</sup>*Centre for Exoplanets and Habitability, University of Warwick  
Gibbet Hill Road*

*Coventry CV4 7AL, UK*

<sup>5</sup>*Dipartimento di Fisica e Astronomia Galileo Galilei, Università di Padova  
Vicolo dell'Osservatorio 3*

*I-35122, Padova, Italy*

<sup>6</sup>*INAF - Osservatorio Astrofisico di Catania  
Via S. Sofia 78*

*I-95123, Catania, Italy*

<sup>7</sup>*Leiden Observatory, Leiden University  
Postbus 9513*

*2300 RA Leiden, The Netherlands*

<sup>8</sup>*Arizona State University*

*Tempe, AZ, USA*

<sup>9</sup>*University of Bern, Center for Space and Habitability  
Gesellschaftsstrasse 6*

*CH-3012, Bern, Switzerland*

<sup>10</sup>*Observatoire astronomique de l'Université de Genève  
51 Chemin des Maillettes*

*1290 Versoix, Switzerland.*

<sup>11</sup>*Space Research Institute, Austrian Academy of Sciences  
Schmiedlstrasse 6*

*A-8041 Graz, Austria*

<sup>12</sup>*INAF - Osservatorio Astronomico di Padova  
Vicolo dell'Osservatorio 5*

*35122 Padova, Italy*

<sup>13</sup>*INAF Osservatorio Astronomico di Palermo  
Piazza del Parlamento, 1*

arXiv:2004.11335v1 [astro-ph.EP] 23 Apr 2020

*I-90134 Palermo, Italy*

<sup>14</sup>*INAF - Osservatorio Astronomico di Trieste  
Via Tiepolo 11  
34143 Trieste, Italy*

<sup>15</sup>*INAF - Osservatorio Astronomico di Brera  
Via E. Bianchi 46  
23807 Merate, Italy*

<sup>16</sup>*Astronomy Department and Van Vleck Observatory  
Wesleyan University  
CT 06459 Middletown, USA*

<sup>17</sup>*Fundación Galileo Galilei-INAF  
Rambla José Ana Fernández Pérez 7  
38712 Breña Baja, TF, Spain*

<sup>18</sup>*INAF - Osservatorio Astronomico di Capodimonte  
Salita Moiariello 16  
80131 Napoli, Italy*

<sup>19</sup>*Department of Physics, University of Rome Tor Vergata  
Via della Ricerca Scientifica 1  
I-00133 Rome, Italy*

<sup>20</sup>*Max Planck Institute for Astronomy  
Konigstuhl 17  
69117 Heidelberg, Germany*

<sup>21</sup>*INAF Osservatorio Astronomico di Cagliari & REM  
Via della Scienza, 5  
I-09047 Selargius CA, Italy*

<sup>22</sup>*INAF-Osservatorio Astrofisico di Arcetri  
Largo Enrico Fermi 5  
I-50125 Firenze, Italy*

<sup>23</sup>*Dipartimento di Fisica G. Occhialini, Università degli Studi di Milano-Bicocca  
Piazza della Scienza 3  
20126 Milano, Italy*

(Received April 24, 2020; Revised; Accepted 22 Apr 2020)

Submitted to ApJL

## ABSTRACT

We present the first detection of atomic emission lines from the atmosphere of an exoplanet. We detect neutral iron lines from the day-side of KELT-9b ( $T_{\text{eq}} \sim 4,000$  K). We combined thousands of spectrally resolved lines observed during one night with the HARPS-N spectrograph ( $R \sim 115,000$ ), mounted at the Telescopio Nazionale Galileo. We introduce a novel statistical approach to extract the planetary parameters from the binary mask cross-correlation analysis. We also adapt the concept of contribution function to the context of high spectral resolution observations, to identify the location in the planetary atmosphere where the detected emission originates. The average planetary line profile intersected by a stellar G2 binary mask was found in emission with a contrast of  $84 \pm 14$  ppm relative to the planetary plus stellar continuum ( $40 \pm 5\%$  relative to the planetary continuum only). This result unambiguously indicates the presence of an atmospheric thermal inversion. Finally, assuming a modelled temperature profile previously published (Lothringer et al. 2018), we show that an iron abundance consistent with a few times the stellar value explains the data well. In this scenario, the iron emission originates at the  $10^{-3}$ – $10^{-5}$  bar level.

*Keywords:* Exoplanet atmospheres — Exoplanet atmospheric composition — Hot Jupiters — High resolution spectroscopy

## 1. INTRODUCTION

Ultra-hot Jupiters are tidally locked gaseous giant planets that orbit their host stars in mere hours or days,

often reaching temperatures above 2,500 K in their permanent day-sides. Unlike for their cooler counterparts, equilibrium chemistry should provide an accurate description of their chemical network, and known condensates are likely secluded to their night-side (Kitzmann et al. 2018; Lothringer et al. 2018; Parmentier et al. 2018; Helling et al. 2019).

Detections of atomic metals at the day-night transition of their atmospheres (WASP-12b, Fossati et al. 2010; Haswell et al. 2012, KELT-9b, Hoeijmakers et al. 2018, 2019; Cauley et al. 2019; MASCARA-2b, Casasayas-Barris et al. 2019; WASP-121b, Sing et al. 2019; Gibson et al. 2020) show that heavy elements are not necessarily sequestered deep in the atmosphere of these planets. This may also indicate the presence of a shallow radiative-convective boundary (Thorngren et al. 2019). Iron is an element of particular interest. Indeed, its abundance is a proxy for the metallicity of stars, and thus a particularly relevant case for comparison between planetary and stellar metallicity. It was detected both in neutral and ionized form at the day-night transition in the atmosphere of KELT-9b (Hoeijmakers et al. 2018; Cauley et al. 2019; Borsa et al. 2019), probing pressures as low as a few  $\mu\text{bar}$  (Hoeijmakers et al. 2019). These lines likely originate within the extended atmosphere of the planet, also detected with  $\text{H}\alpha$  and  $\text{Ca II}$  lines (Yan & Henning 2018; Turner et al. 2020; Yan et al. 2019). Ionized iron was also found in the upper atmosphere of MASCARA-2b (Casasayas-Barris et al. 2019) and in the exospheres of WASP-12b and WASP-121b (Haswell et al. 2012; Sing et al. 2019). Yet, a detection of photospheric planetary iron lines is still missing.

In this paper we employ the high-resolution ( $R \sim 115,000$ ) spectrograph HARPS-N, mounted at the Telescopio Nazionale Galileo (TNG), to observe for the first time the thermal emission of an exoplanet with this instrument. To do so, we targeted KELT-9b monitoring the planet from quadrature to right before the planet is eclipsed behind the star. We describe our observations and data reduction in section 2 and Appendix A, including an adaptation of the line-weighted stellar binary masks method, traditionally used to extract radial velocities of exoplanets, to extract the signal of the planet via a cross-correlation function (CCF). We present the results of this analysis in Sec. 3. We then perform a second analysis of the extracted planetary CCF to derive atmospheric parameters of the planet based on models (Sec. 4, Sec. 5). To this aim, we introduce a new method to compare models and observations making use of the CCF technique with a line weighted binary mask, present a novel adaptation of the concept of contribution function to the context of cross-correlation analyses

(Section 4), and apply these tools to our observations (Sec. 5). We discuss the implications of our study in Section 6.

## 2. METHODS: TREATMENT OF DATA

### 2.1. Observations and data reduction

We observed the KELT-9 system in the framework of a Long-Term program (PI G. Micela) with HARPS-N and GIANO-B in GIARPS@TNG configuration (Claudi et al. 2017), as part of the GAPS project (Covino et al. 2013). For the present work, we only used the HARPS-N observations taken from the 22nd of July 2018 21:23 UT to the 23rd of July 2018 05:21 UT. The GIANO-B observations will be the target of a dedicated study. We collected 89 HARPS-N exposures, each with 180 seconds of integration. This is shorter compared to Hoeijmakers et al. (2018) and Hoeijmakers et al. (2019), who used an exposure time of 600 seconds. With this choice, the planet moved by at most  $2.25 \text{ km s}^{-1}$  during each exposure, which smeared the signal over 2.7 pixels. Considering the transit centred at phase 0, the planetary phases covered the range between 0.227 and 0.452, such that the planet was not occulted by the stellar disk.

We extracted and calibrated the spectra using the standard Data Reduction Software (DRS; version 3.7.1, Dumusque 2018). To avoid the increase of correlated noise from data interpolation, we performed our analysis on the individual echelle orders (e2ds spectra), after correcting for the blaze function. As previously reported by Borsa et al. (2019), our observations were affected by a malfunction of the Atmospheric Dispersion Corrector that caused a deformation of the spectral energy distribution due to chromatic losses, and a concomitant loss of efficiency in the blue part of the spectra across the night (see Fig. 1). While we mitigated this effect with a custom color-correction, following Malavolta et al. (2017), it is not possible to recover the lost signal-to-noise ratio at shorter wavelengths. We did not correct for telluric lines, because our analysis naturally excludes regions that are contaminated by them (Sec. 2.2). We then aligned the stellar spectra by removing the Barycentric Earth Radial Velocity motion, effectively shifting the spectra to the barycentric rest-frame of the solar system. This allowed us to build a high signal-to-noise ratio master stellar spectrum by: (1) rescaling every order to its average counts value and (2) computing a median in time for each order. The stellar motion induced by the planet amounts to about  $0.2 \text{ km s}^{-1}$  throughout the night, and does not significantly impact the shape of the stellar lines which are rotationally broadened by more than  $100 \text{ km s}^{-1}$ . Since the planet moved in radial velocity by more than one pixel per exposure for most

of the night, the resulting master spectrum contained the planetary lines only in minimal part. Each single e2ds spectrum was then divided by the master stellar spectrum, which removed the stellar lines. This procedure effectively provides the planet emission spectrum normalized to the stellar emission spectrum and planet continuum plus 1 (see Appendix A). A high-pass filter was then applied to each of the resulting rows to remove residual discontinuities and low-order variations due to imprecise blaze or color correction (see Appendix A). We found that the application of the high-pass filter enhanced the precision on the retrieved parameters by a factor of about 2. We finally applied a custom binary mask cross-correlation method (see Sec. 2.2).

## 2.2. Line-weighted binary mask CCF

With a temperature comparable to a K-dwarf, the atmosphere of KELT-9b should exhibit thousands of optical atomic transitions. The technique of cross-correlation (Baranne et al. 1979; Sparks & Ford 2002; Snellen et al. 2010) is best suited for their search (Hoeijmakers et al. 2019). The technique was applied with different flavours (e.g. template matching, binary mask), and consists in stacking these thousands of planetary lines to abate the photon noise, which hinders their detection.

We adopted a CCF technique with a weighted binary mask (Baranne et al. 1996; Pepe et al. 2002)<sup>1</sup>, where weights are attributed to individual spectral lines according to their information content (see Appendix B). It can employ the classic stellar binary masks used in the search of planets with the radial velocity method, as well as custom binary masks, and can be applied both to models and data. Compared to other cross-correlation schemes (Snellen et al. 2010), the binary mask approach preserves the contrast of the lines that it intercepts, which allows the comparison of the strength of different spectral features (Pino et al. 2018a). In practice, our technique extracts the average planet line normalized to the planet plus star continuum (which we call planet excess). This is similar to a least-squares deconvolution scheme (LSD, Donati et al. 1997), but without deconvolving the cross-correlation vector (with no loss of accuracy in the interpretation; Sec. 4). This average line profile can be used to directly extract observational properties of the planetary emission (Sec. 3), but the extraction of parameters of the atmospheric structure

requires the extra step of model comparison, for which we present a new method (Sec. 4).

Other works relied on similar definitions of the cross-correlation function (Hoeijmakers et al. 2019). However, they determined the weights on single pixels using model-injection, thus based on their information content, with the aim of reaching the highest signal-to-noise ratio on the planetary detection. In our approach, the binary mask attributes weights to single lines, as opposed to single pixels, with the advantage of reduced complexity and model-dependence. The consequently easier interpretation is obtained at the cost of a possible loss of signal-to-noise ratio, especially in the wings of the lines.

Since the planet has a temperature comparable to that of a star, in this work we adopted standard G2, K0 and K5 stellar masks provided by the DRS, optimized to extract radial velocities for planets orbiting stars for that spectral type. These masks are designed to exclude parts of the spectrum that are contaminated by telluric lines.

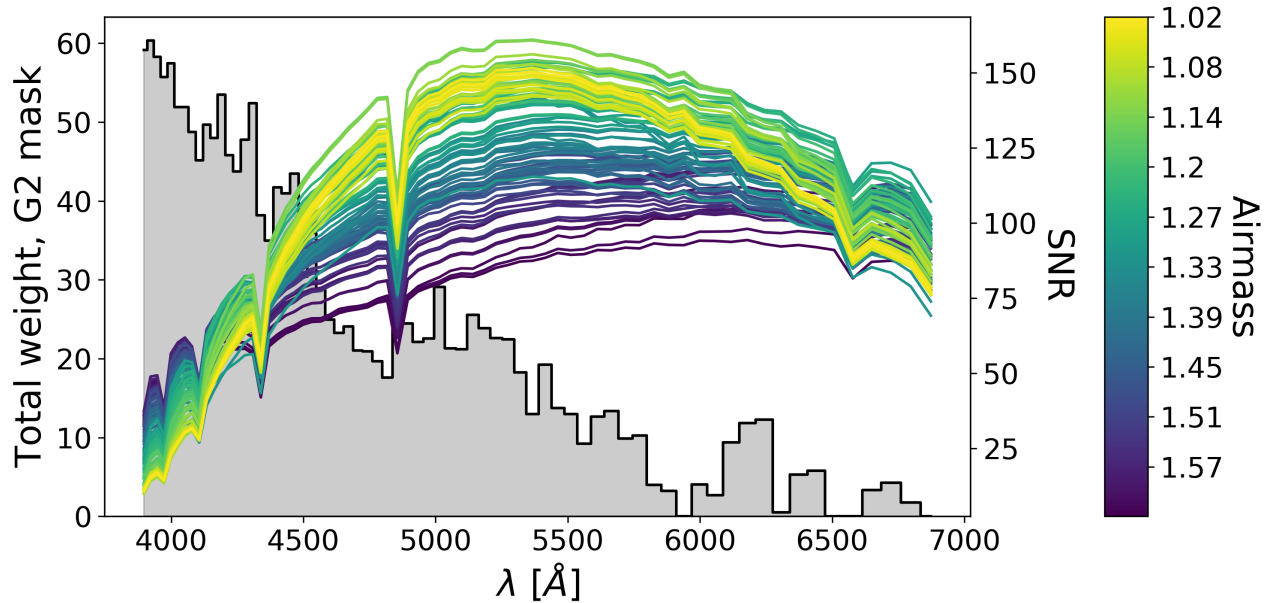
The results are mostly independent from the choice of the spectral type of the mask. This is likely because the masks share the strongest lines. Indeed, among the 1,000 strongest lines in each mask, the majority of the lines are closer than 0.001 Å, less than one tenth of a pixel. In percentage, the masks share 74.4% (G2 vs K0), 82.6% (G2 vs K5) and 84.4% (K0 vs K5) of the strongest lines. In the following, we discuss the G2 mask case.

A CCF is computed for every exposure. The result is an ‘exposure matrix’ which displays the planet trace in a diagram with radial velocity displacement from the stellar rest frame on the x-axis, and planetary phase (or exposure) on the y-axis (Fig. 2, upper panel). The fit (Sec. 4) was directly performed on this exposure matrix. However, we also display the results in the traditional  $K_p$ - $v_{\text{sys}}$  diagram, which visually highlights the presence or lack of a signal. In practice, we parametrized the planet orbit with a Keplerian velocity  $K_p$ , appropriate for a circular orbit, and moved to the corresponding planet rest frame. Only the correct  $K_p$  aligns the individual CCFs, that are then summed. The maximum is thus found at the global radial velocity of the system (systemic velocity,  $v_{\text{sys}}$ ). This is conveniently represented in the  $K_p$ - $v_{\text{sys}}$  diagram (Fig. 2, middle panels).

## 3. OBSERVATIONAL RESULTS

In Fig. 2, we present the result of applying a G2 binary mask to the planet-to-star flux ratio. In practice, what we see is the average planet emission line intersected by the G2 binary mask normalized to the planetary and stellar continua. This emission is interpreted

<sup>1</sup> We do not normalize by the standard deviation. As such, our scheme is a cross-covariance in the statistical sense, but we call it cross-correlation following Baranne et al. (1996) and the signal processing nomenclature.



**Figure 1.** Signal-to-noise ratio as a function of airmass and wavelength (solid curves), and total weight within the mask for each spectral bin (gray histogram; accounting for the number of spectral lines and their depth only). Spectra acquired at lower airmass are expected to have higher signal-to-noise ratio throughout the spectrum, due to a lower optical depth of Earth atmosphere, but the malfunction of the ADC modifies this behaviour. This effect is particularly severe in the blue, where most of the information on the planet lies, as quantified by the weight in the binary mask. Indeed, while no change in observing conditions were noticeable in our run, the bluest orders of the lowest airmass spectra have a signal-to-noise ratio which is half compared to airmass 1.6, the maximum reached within our run.

as due to the atmosphere of the planet.

The planetary atmospheric spectral feature, as seen through the G2 mask, has a contrast of  $(84 \pm 1)$  ppm relative to the continuum. We obtained this by fitting a Gaussian curve to the planetary signal integrated over the exposures assuming the best fit  $K_p$  (see Fig. 2, lower panel; Sec. 5.1). The formal error is likely underestimated due to the presence of correlated noise. By replacing the formal error with the standard deviation far from the planet signal (14 ppm, calculated at  $-200 \text{ km s}^{-1} < v_{\text{sys}} < -100 \text{ km s}^{-1}$ ), the signal-to-noise ratio of the detection is 6.

We then assumed that the continuum is the sum of the stellar and planetary continua (see Appendix A). We further assumed that the stellar and planetary continua are blackbodies at temperatures of 10,000 K and 4,570 K (Wong et al. 2019), respectively. The contrast relative to the planetary continuum is then

$$\frac{F_{\text{lines p}}}{F_{\text{cont, p}}} = 84 \cdot \left( 1 + \frac{F_{\text{cont, *}}}{F_{\text{cont p}}} \right) \text{ ppm}, \quad (1)$$

yielding  $(40 \pm 5)\%$ .

The planet excess appears in emission and not in absorption, which is an unambiguous sign of the presence of a thermal inversion in the atmosphere of the planet

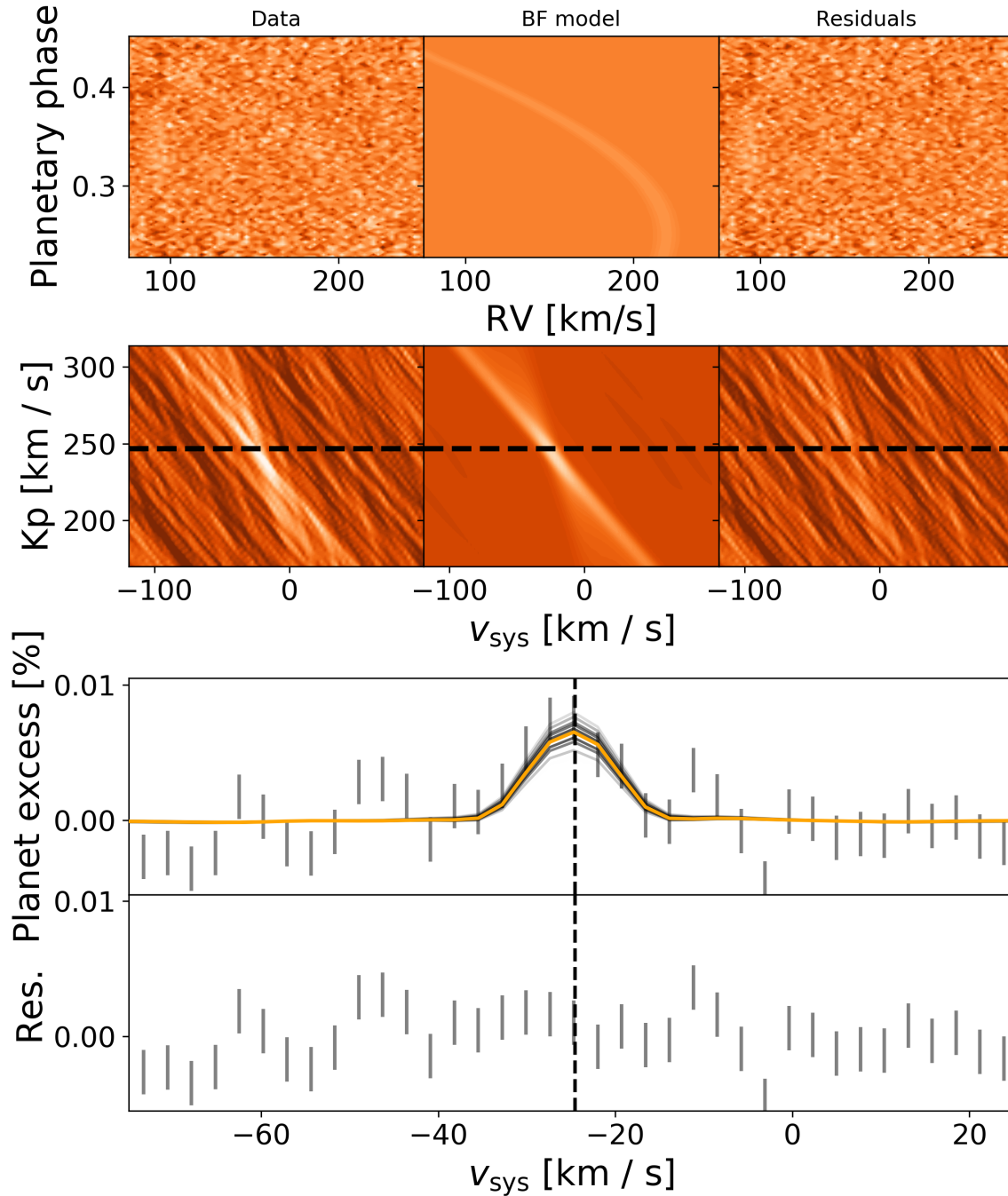
(see Schwarz et al. 2015; Nugroho et al. 2017; section 6.1).

#### 4. METHODS: EXTRACTING ATMOSPHERIC PARAMETERS OF KELT-9B

The next step is extracting the planetary parameters from the cross-correlation function. This requires two ingredients: (1) a parametrized model for the exoplanet atmosphere (Sec. 4.1) and (2) a cross-correlation to likelihood mapping (Sec. 4.2). We also adapt the concept of contribution functions to the line-weighted binary mask CCF, to identify the pressure range probed by our analysis (Sec. 4.3).

##### 4.1. Model grid of KELT-9b atmosphere

To compute planetary synthetic spectra, we developed a custom, line-by-line radiative transfer code that implements (1) opacities from the most important optical opacity sources in ultra hot Jupiters (Fe I, Fe II, Ti I, Ti II,  $\text{H}^-$ ; Kitzmann et al. 2018; Arcangeli et al. 2018; Lothringer & Barman 2019), (2) equilibrium chemistry. LTE is assumed throughout the planetary atmosphere, and  $\log g_p = 3.3$ . We neglected the reflected light component: from a theoretical standpoint, no reflective aerosols are expected in the atmosphere of the planet (Kitzmann et al. 2018); from an observational stand-



**Figure 2.** Observed and modelled average planet emission line intersected by the G2 mask, and residuals between data and best fit model. *Upper panel:* The exposure matrix in the region where we performed the fit. The curvature of the planetary trace is due to its overnight change in radial velocity compared to its host star. *Middle panel:*  $K_p$ - $v_{\text{sys}}$  diagram for data, best fit model and residuals. The color scale is the same across the three panels, showing that the residuals map is clean in the region where the planet excess is localized. A horizontal, black, dashed line indicates the best fit value for  $K_p$ . *Lower panel:* The average data, model and residuals in the best fit planetary rest frame  $K_p$ . Gray vertical lines are the data, with their uncertainties at 1 standard deviation, while the orange line is the model shown in the middle panel. Black lines are models deviating by less than  $2\sigma$  from the best fit, while varying the iron abundance, with transparency proportional to their deviation. The bottom half of the panel shows the residuals from the best fit with the same y-axis. A black dashed vertical line shows the best fit systemic velocity. The average planetary line is in emission, and has a contrast of 84 ppm compared to the continuum.

point, due to the fast rotation of the host star (Gaudi et al. 2017), reflected spectral lines are broadened and thus difficult to detect with our continuum-normalized technique, that removes the majority of their signal. Furthermore, due to the polar orbit of the planet (Gaudi et al. 2017), the reflected stellar atomic lines would show a variable broadening approximately ranging between the intrinsic broadening of the stellar lines (quadrature) and the rotational broadening of the star (eclipse, 112 km s<sup>-1</sup>), while the observed broadening is constant and consistent with the expected rotational broadening of the planet ( $\sim 6.63$  km s<sup>-1</sup>). We further detail the radiative transfer code in Appendix C. Here we illustrate the parameter space explored.

Our synthetic spectra can be expressed as:

$$S \left( \frac{\text{VMR}_{\text{Fe}}}{\text{VMR}_{\text{Fe}\star}}, \frac{\text{VMR}_{\text{Ti}}}{\text{VMR}_{\text{Ti}\star}}, v_{\text{sys}}, K_{\text{p}}, v_{\text{rot,p}}, \text{TP} \right), \quad (2)$$

where  $\text{VMR}_{\text{Fe}}/\text{VMR}_{\text{Fe}\star}$  and  $\text{VMR}_{\text{Ti}}/\text{VMR}_{\text{Ti}\star}$  are the planetary to stellar iron and titanium volume mixing ratio,  $v_{\text{rot,p}}$  is the planetary rotational velocity assumed constant in the atmospheric region probed by the planetary emission lines, TP is a suitable parametrization of the temperature pressure profile. At our precision level, we expect retrieved abundances to be degenerate with the rotation rate (due to broadening) and the temperature profile, so that a full exploration of the parameter space is necessary to provide accurate constraints on each parameter. This is beyond the scope of this paper. Instead, we focused on (1) determining which atomic species is mainly responsible for the observed average planetary emission line intersected by the G2 mask and (2) testing the hypothesis that the planet spectrum can be explained assuming abundances consistent with that of its host star. We could thus limit the parameter space by assuming that the planet is tidally locked ( $v_{\text{rot,p}} \cdot \sin i_p = 6.63$  km s<sup>-1</sup>). Furthermore, we fixed the thermal profile to the self-consistent temperature profile of KELT-9b that Lothringer et al. (2018) obtained by assuming a planetary metallicity equal to the stellar value and equilibrium chemistry. Under these reasonable assumptions, we produced three groups of models:

$$\begin{aligned} S_{\text{Fe, Ti}} & \left( \frac{\text{VMR}_{\text{Fe}}}{\text{VMR}_{\text{Fe}\star}}, \frac{\text{VMR}_{\text{Ti}}}{\text{VMR}_{\text{Ti}\star}}, v_{\text{sys}}, K_{\text{p}} \right), \\ S_{\text{Fe}} & \left( \frac{\text{VMR}_{\text{Fe}}}{\text{VMR}_{\text{Fe}\star}}, v_{\text{sys}}, K_{\text{p}} \right) \Big|_{\text{VMR}_{\text{Ti}}/\text{VMR}_{\text{Ti}\star}=0}, \\ S_{\text{Ti}} & \left( \frac{\text{VMR}_{\text{Ti}}}{\text{VMR}_{\text{Ti}\star}}, v_{\text{sys}}, K_{\text{p}} \right) \Big|_{\text{VMR}_{\text{Fe}}/\text{VMR}_{\text{Fe}\star}=0}. \end{aligned} \quad (3)$$

$S_{\text{Fe}}$  and  $S_{\text{Ti}}$  are obtained from  $S_{\text{Fe, Ti}}$  by removing titanium and iron, respectively. We fitted  $v_{\text{sys}}$  and  $K_{\text{p}}$ , and

simultaneously varied  $\text{VMR}_{\text{Fe}}/\text{VMR}_{\text{Fe}\star}$  between  $10^{-1}$  and  $10^3$ , and  $\text{VMR}_{\text{Ti}}/\text{VMR}_{\text{Ti}\star}$  between  $3 \cdot 10^{-3}$  and  $3 \cdot 10^3$ . Since the host star KELT-9 has a metallicity between 0.7 and 2.7 times solar, higher volume mixing ratios seem unlikely. Lower volume mixing ratios would not be detectable at our precision level, and would thus not suffice to explain the data.

#### 4.2. A new interpretation scheme for CCFs

The strength of the cross-correlation signal depends on the quality of the match between the binary mask, or the model if used directly, to the data. For example, this can be quantified through peak signal to noise ratio. However, this approach is not statistically sound and can therefore not be used to estimate confidence intervals on planet parameters (Brogi et al. 2016). Alternatives exist, such as the Welch  $T$ -test (Brogi et al. 2013) or  $\chi^2$ -comparison based on model injection into data (Brogi et al. 2016), but they are computationally expensive. To overcome these challenges, Brogi & Line (2019) presented a cross-correlation to likelihood mapping to perform the comparison in a statistically sound framework (see also Gandhi & Madhusudhan 2019), further generalized by Gibson et al. (2020), while Fisher et al. (2019) proposed a different method based on a random forest approach.

Here, we propose a novel method to directly compare the cross-correlation of models and data. The procedure consists in simulating end-to-end synthetic observations, including the effects of data reduction. In the case of HARPS-N, this is facilitated by the small contamination from telluric lines. Furthermore, HARPS-N is a very stable instrument, built to acquire precise radial velocity observations. Consequently, our data reduction procedure is relatively simple. We are thus able to simulate end-to-end the effect of the data reduction process on synthetic e2ds HARPS-N generated from our models. This enables a direct comparison using a likelihood function, in a procedure similar to what Kochukhov et al. (2010) have previously suggested to interpret LSD profiles. We cross-checked our new method with the likelihood-mapping by Brogi & Line (2019), finding good agreement (Appendix E).

The first step is simulating the exposure matrix described in Sec. 2.2 :

- We modelled the star using a PHOENIX model ( $T_{\text{eff}} = 10,000$  K,  $\log g = 4.0$ ), and applied rotational broadening ( $v_{\text{rot},\star} \cdot \sin i_{\star} = 111.8$  km s<sup>-1</sup>, Borsa et al. 2019, linear limb darkening coefficient  $\epsilon = 0.6$ ).
- We convolved each model emission spectrum of the exoplanet with a rotational kernel correspond-

ing to the tidally locked solution ( $v_{\text{rot,p}} \cdot \sin i_p = 6.63 \text{ km s}^{-1}$ ).

- For each exposure  $t_i$ , we Doppler shifted every spectrum for a given orbital solution ( $K_p, v_{\text{sys}}$ ). These velocities were parameters of the fit.

We then processed the simulated time-series through the procedure described in Appendix A. The result was a model exposure matrix for each set of parameters ( $K_p$  and  $v_{\text{sys}}$ ;  $\text{VMR}_{\text{Fe}}$  and  $\text{VMR}_{\text{Ti}}$ ), that we could directly fit to observations (see Fig. 2). Finally, we computed the Gaussian likelihood for radial velocities between  $75 \text{ km s}^{-1}$  and  $252 \text{ km s}^{-1}$ , within which the planet trace is expected to be found, by<sup>2</sup>

$$\log \mathcal{L} = \sum_i \left[ -\log \left( \sigma_i \sqrt{2\pi} \right) - \chi_i^2 / 2 \right], \quad (4)$$

where  $\sigma_i$  and  $\chi_i^2$  are the error and  $\chi^2$  associated to the data point  $i$ . We assumed that  $\sigma_i$  is constant in radial velocity over an exposure, and measured it as the dispersion far from the expected position of the planet (radial velocities between  $-200 \text{ km s}^{-1}$  and  $-100 \text{ km s}^{-1}$ ). The end result was a multi-dimensional  $\log \mathcal{L}$  grid. We then employed different flavours of the likelihood test ratio to assess the significance of each model, to compare the models and to extract confidence intervals (see Appendix D for practical details on how to do so).

This process is too slow to explore a large 4-dimensional grid of parameters. To speed it up, we found that: (1) rotational broadening can be included directly in the cross-correlated spectra; (2) instead of simulating all the exposures for each value of the couple ( $K_p, v_{\text{sys}}$ ), the model exposure matrix can be directly shifted to simulate different values of the couple ( $K_p, v_{\text{sys}}$ ) (see also Brogi & Line 2019). Practically, this assumes that the data reduction process effects on the planetary trace are independent of its  $K_p$  and  $v_{\text{sys}}$ . We tested that both approximations do not cause a significant variation of the likelihood distributions.

#### 4.3. Contribution function of the cross-correlation function

Since in our approach we are able to simulate the cross-correlation function of each model, for a given assumed atmospheric structure it is possible to directly study the location in pressure where the cross-correlation signal originates from. This can be described with a ‘contribution function to the cross-correlation function at the surface’. To our knowledge, this is the

first time that the contribution function is adapted to the context of high spectral resolution observations of planetary atmospheres. We define it here by analogy with the classic contribution function to the flux at the surface.

Following e.g. Irwin (2009) and Malik et al. (2019) we define the contribution functions as the contribution of each discrete layer in our model to the flux at the surface of the planetary atmosphere. In our case, we locate the ‘surface’ high-up in the optically thin region of the planet atmosphere, from which the photons escape and reach the observer. If every layer  $n$  emits an intensity  $\Delta_n I(\mu)$  in a direction  $\mu = \cos \theta$ , we can write:

$$I(\mu) = \sum_n [\Delta_n I(\mu) \exp(-\tau_n/\mu)], \quad (5)$$

where  $\tau_n$  represents the optical depth above layer  $n$ , and  $\Delta_n I$  is computed according to the linear in optical depth approximation (Toon et al. 1989). The  $n$ -th term in square brackets on the right-hand side of the equation is the contribution function of layer  $n$ .

We now apply the cross correlation at the left hand and right-hand side of Eq. 5. The sum over  $n$  atmospheric layers can be commuted with the sums contained in our definition of CCF (Eq. B9). As a result, we can write:

$$\begin{aligned} \text{CCF}(I(\mu)) &= \text{CCF} \left( \sum_n [\Delta_n I(\mu) \exp(-\tau_n/\mu)] \right) = \\ &= \sum_n \text{CCF} ([\Delta_n I(\mu) \exp(-\tau_n/\mu)]). \end{aligned} \quad (6)$$

By extension, the  $n$  terms in square brackets in the right-hand side of Eq. 6 represent the ‘contribution functions of the cross-correlation function’ of each layer. Physically, they represent the contribution to the intensity as a function of radial velocity rather than wavelength from every atmospheric layer.

With this definition, it is trivial to identify the pressure range that can be probed with a line-weighted binary mask CCF of high spectral resolution observations. Furthermore, for a given model, the contribution functions of the CCF inform us on which pressure layers more tightly constrain the radial velocity of the planet. By integrating over  $\mu$  one obtains expressions for the flux.

## 5. RESULTS FROM MODEL COMPARISON

In the following, we provide our interpretation of the average planet line intersected by the G2 mask based on model comparison.

### 5.1. Fit with line weighted binary mask

We first identified which among the models defined in Eq. 3 best explains the data. The model containing only

<sup>2</sup> The method can be used with any other likelihood function



lines from neutral and ionized titanium and no atmospheric iron,  $S_{\text{Ti}}$ , has maximum likelihood at the highest allowed abundances of titanium. This suggests that titanium lines are too weak to explain the observed emission lines even when  $\text{VMR}_{\text{Ti}} = 3,000 \cdot \text{VMR}_{\text{Ti},\star}$ . We then compared  $S_{\text{Ti}}$  to the full model  $S_{\text{Fe}, \text{Ti}}$  with a likelihood test ratio (see Appendix D), and found that it can be excluded at  $4.3\sigma$ . When limiting the maximum abundance of titanium to 100 times the stellar value, the model can be excluded at  $5.1\sigma$ . As a further indication that iron is necessary to explain the observed emission line, we calculated the difference in Bayesian Information Criterion (BIC, Liddle 2007) and found that  $\min[\text{BIC}(S_{\text{Ti}})] = \min[\text{BIC}(S_{\text{Fe}, \text{Ti}})] + 10$ . The difference increases to 17.5 when limiting the maximum abundance of titanium to 100 times the stellar value, indicating strong preference for the presence of iron.

In a similar fashion, we compared the model containing only lines from neutral and ionized iron and no atmospheric titanium,  $S_{\text{Fe}}$ , to the full model. In this case, the null hypothesis that  $S_{\text{Fe}}$  is the true model can not be excluded ( $< 0.1\sigma$ ). Furthermore, it is strongly favoured by the BIC test, with  $\min[\text{BIC}(S_{\text{Fe}})] = \min[\text{BIC}(S_{\text{Fe}, \text{Ti}})] - 8.7$ , which penalizes the presence of an additional parameter in  $S_{\text{Fe}, \text{Ti}}$ . We thus adopted  $S_{\text{Fe}}$  as our nominal model to derive planetary parameters (see Table 5.1).

The best fit parameters are  $K_p = 242 \text{ km s}^{-1}$ ,  $v_{\text{sys}} = -22.5 \text{ km s}^{-1}$ ,  $\text{VMR}_{\text{Fe}} = 30 \cdot \text{VMR}_{\text{Fe},\star}$ . The model is a very good match to the data, as evidenced by a reduced  $\chi^2 = 6128/5874 = 1.043$  and by residuals within the statistical fluctuations (Fig. 2). We computed the significance of the model by performing a likelihood test ratio, comparing it to the case of null detection  $\text{VMR}_{\text{Fe}} = 0$  (a straight line; see Appendix D). The detection is significant at  $6.15\sigma$ . As a further test, we computed that the BIC value of our best fit model is lower by 24.5 compared to the null detection, indicating a strong preference for the presence of iron. The  $1\sigma$  confidence intervals for the three parameters (see Appendix D) are  $238 \text{ km s}^{-1} < K_p < 247.5 \text{ km s}^{-1}$ ,  $-32 < v_{\text{sys}} < -19$  and  $10 < \text{VMR}_{\text{Fe}}/\text{VMR}_{\text{Fe},\star} < 150$  (compatible with a few times the stellar value at  $2\sigma$ ).

Finally, we compared our nominal model  $S_{\text{Fe}}$  with two models where we suppressed lines by neutral and ionized iron respectively. These two models are not formally nested in  $S_{\text{Fe}}$ , and we compared instead the significance yielded by the best fit parameters for each model. When only neutral iron is present, the results are nearly indistinguishable from the full model  $S_{\text{Fe}}$ , with a similar significance, best fit and confidence interval. On the other hand, when only ionized iron is present, the best fit is

**Table 1.** Comparison of models containing iron or titanium lines.

	$\Delta\text{BIC}$ with	LRT with
	$S_{\text{Fe}, \text{Ti}}$	$S_{\text{Fe}, \text{Ti}}$
$S_{\text{Fe}}$	-8.7	$< 0.1\sigma$
$S_{\text{Ti}}$	+10	$4.3\sigma$

**Notes.** The LRT metric indicates that a model containing neutral and ionized iron ( $S_{\text{Fe}}$ ) explains the data as well as a model containing also neutral and ionized titanium. On the other hand, a model containing only lines from neutral and ionized titanium ( $S_{\text{Ti}}$ ) does significantly worse. Furthermore, the BIC difference favours the model containing only neutral and ionized iron, and no titanium, due to the smaller number of free parameters. We thus adopt  $S_{\text{Fe}}$  as fiducial model.

**Table 2.** Comparison of models containing neutral iron lines, ionized iron lines or both.

	$\Delta\text{BIC}$ with	LRT with
	null detection	null detection
Neutral and ionized iron ( $S_{\text{Fe}}$ )	-25	$6.15\sigma$
Neutral iron only	-25	$6.15\sigma$
Ionized iron only	+5	$3.1\sigma$

**Notes.** The LRT metric indicates that a model containing only ionized iron has a lower significance compared to the null detection. Although the significance is still at the  $3\sigma$  level, this occurs at the upper limit of the allowed iron abundances (1,000 times solar), and the BIC test significantly disfavours this model compared to a flat line. Neutral iron is thus necessary to explain the data under our assumptions. Furthermore, The addition of ionized iron does not significantly improve the fit, or significantly change the inferred iron abundance.

found at the upper limit of  $\text{VMR}_{\text{Fe}} = 1,000 \cdot \text{VMR}_{\text{Fe},\star}$  and has a much lower significance of  $3.1\sigma$ . In this case, the BIC test favours the null detection, indicating that the ionized iron lines intersected by the G2 mask are too weak to explain the observed planetary feature alone (see Table 5.1).

We also applied the method by Brogi & Line (2019) to perform an independent test (see Appendix E). In this case, we fixed the abundance to its best-fit value, and retrieved  $K_p$  and  $v_{\text{sys}}$  and a scale factor which is a proxy for abundance. The results are in good agreement with those found with our novel framework (see Fig. 3, and Appendix E).

## 5.2. Reference frame of the signal

The comparison at face-value of the joint probability distributions and the marginalized 1D probabilities

reveals that our results are consistent with all literature values of the systemic velocity (Gaudi et al. 2017 adopted by Yan & Henning 2018, Hoeijmakers et al. 2019 and Borsa et al. 2019; see Table 5.2). While these authors reported individual precisions around  $0.1 \text{ km s}^{-1}$ , the measured values are significantly discrepant, spanning a range of about  $3 \text{ km s}^{-1}$ . Further analysis is required to pinpoint the origin of this discrepancy. We thus attributed an error of  $3 \text{ km s}^{-1}$  to the single measurements to account for an unknown systematic effect. With this assumption, the average  $v_{\text{sys}, \star} = -19 \pm 3 \text{ km s}^{-1}$  is compatible within one sigma with our result ( $\Delta v_{\text{sys}, \star} = 3.5^{+5.5}_{-4.5} \text{ km s}^{-1}$  and  $\Delta v_{\text{sys}, \star} = 1^{+3}_{-4} \text{ km s}^{-1}$  for the line weighted binary mask and the Brogi & Line 2019 approaches respectively). Furthermore, deviations between  $K_p$  measured from atomic metal lines in emission (our work) and in transmission (Hoeijmakers et al. 2019) are in agreement at the  $2\sigma$  level. However, the  $K_p$  value measured by Yan & Henning (2018) on the  $\text{H}\alpha$  line is in tension with the  $K_p$  measured on the metal lines ( $\Delta K_p = 27^{+7.5}_{-8} \text{ km s}^{-1}$  and  $\Delta K_p = 27.5 \pm 6 \text{ km s}^{-1}$  for the line weighted binary mask and Brogi & Line 2019 approaches respectively). We explored the possibility that this difference is of astrophysical origin, due to the fact that the hydrogen and iron lines probe different regions of the atmosphere. Yan & Henning (2018) report that the  $\text{H}\alpha$  line approaches but does not reach the Roche lobe. Furthermore, the  $\text{H}\alpha$  line has a symmetrical profile. Therefore, it is likely generated below the exosphere, in the part of the atmosphere gravitationally bound to KELT-9b. Any relative motion between the gas components probed by observations should thus be subsonic. By assuming the adiabatic coefficient of a monoatomic gas, the temperature profile by Lothringer et al. (2018) and the mean molecular weight from our model, we obtain that the sound speed ranges between  $6.5 \text{ km s}^{-1}$  and  $8.5 \text{ km s}^{-1}$ . If it was of astrophysical origin, the difference between the semi-amplitude measured by Yan & Henning (2018) and our measurement would thus be larger than the sound speed (although only marginally in the case of the line weighted binary mask), which is unlikely. Further dedicated work is necessary to consistently explain these observations.

## 6. DISCUSSION

### 6.1. A temperature inversion in the day-side of KELT-9b

The average planet line intersected by the G2 mask is in emission, which can only be explained if a thermal inversion is present in the atmosphere of KELT-9b. This conclusion is model-independent, since it only hinges on

**Table 3.** Literature and derived  $v_{\text{sys}}$  and  $K_p$  values.

	$v_{\text{sys}} [km s^{-1}]$	$K_p [km s^{-1}]$
Yan & Henning (2018)	$-20.6 \pm 0.1^a$	$269^{+6.5}_{-6}$
Borsa et al. (2019)	$-19.81 \pm 0.02$	–
Hoeijmakers et al. (2019)	$-17.7 \pm 0.1$	$234.24 \pm 0.9$
This work, G2 mask	$-22.5^{+3.5}_{-4.5}$	$242^{+5}_{-4}$
This work, Brogi & Line (2019) technique	$-20.5^{+2}_{-1.5}$	$241.5^{+3}_{-2}$

**Notes.** The error bars indicate  $1\sigma$  intervals reported in the literature, or on the 1D marginalized likelihoods. Our results are broadly consistent with the literature, with the exception of  $K_p$  measured by Yan & Henning (2018). When both are measured from the planetary spectrum, systemic velocity and Keplerian velocity are correlated, as evident from Fig. 3, where the 2D confidence intervals are reported.

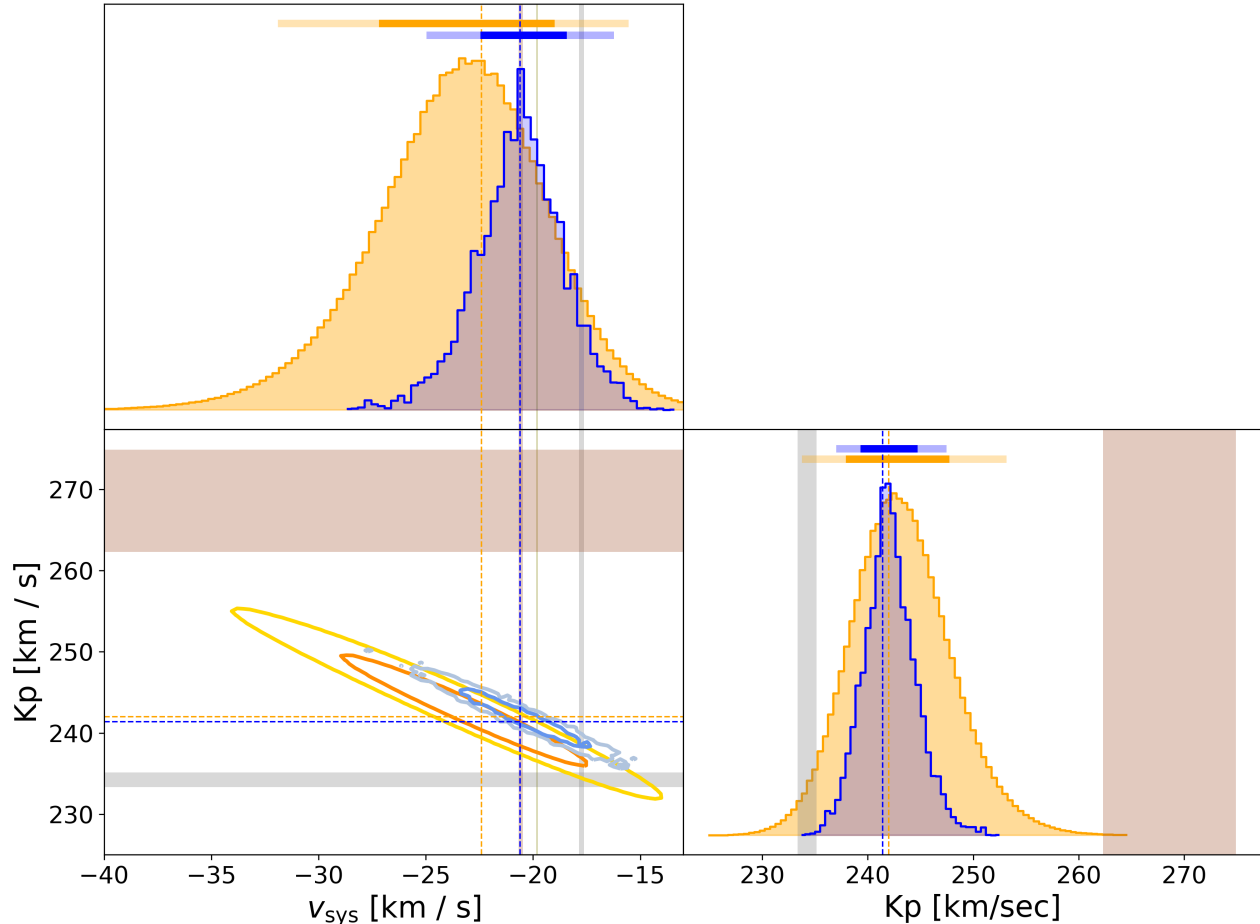
<sup>a</sup> Taken from Gaudi et al. (2017).

the sign of the planetary lines, which is preserved by our analysis.

We calculated the contribution functions to the CCF of the model adopting the thermal inversion by Lothringer et al. (2018) and solar iron abundance (Sec. 4.3). The emission from the neutral iron line cores originates between  $10^{-3}$  bar and  $10^{-5}$  bar (see Fig. 4). This is higher-up compared to the  $\sim 30$  mbar region probed by Hooton et al. (2018), who reported an evidence of inversion using ground-based photometry. It is also well within the inverted region of the atmosphere, found above the region of absorption of stellar irradiation and located between 1 and 100 mbar in the optical region probed by HARPS-N (Lothringer et al. 2018).

For hot Jupiters with equilibrium temperature larger than 1,600 K, molecules with strong optical opacities such as TiO and VO are expected to be in the gas phase causing a temperature inversion below 0.1 bar (Hubeny et al. 2003; Fortney et al. 2008). For the higher temperatures experienced by ultra-hot Jupiters, most molecules are dissociated, so these species become irrelevant for the thermal inversion. Instead, atomic metals and metal hydrides are capable of absorbing enough short wavelength irradiation to heat up the atmosphere. In particular, the bound-bound transitions of neutral iron absorbs significantly long-ward of  $0.3 \mu\text{m}$ , and the bound-free transitions absorbs the high-energy flux short-ward of  $0.3 \mu\text{m}$  (Sharp & Burrows 2007). This is enough to create a thermal inversion at 10 mbar (Lothringer et al. 2018). Higher up, around 0.5 mbar, iron is mostly found in its ionized form due to the high atmospheric temperature.

A second important factor that contributes to the for-



**Figure 3.** Marginalized likelihood distributions for  $v_{\text{sys}}$  and  $K_p$  for the line weighted binary mask (orange) and for the Brogi & Line (2019) approach (blue). Dark and light orange (blue) horizontal bars denote the  $1\sigma$  and  $2\sigma$  confidence levels. Orange (blue) dashed lines indicate the best-fit value. Shaded areas denote the literature values by Yan & Henning (2018) (sienna), Hoeijmakers et al. (2019) (gray) and (Borsa et al. 2019) (olive). Borsa et al. (2019) only measure  $v_{\text{sys}}$ . Our distributions for  $v_{\text{sys}}$  is consistent with the literature, while we deviate from the  $K_p$  value by Yan & Henning (2018) by about  $3\sigma$ .

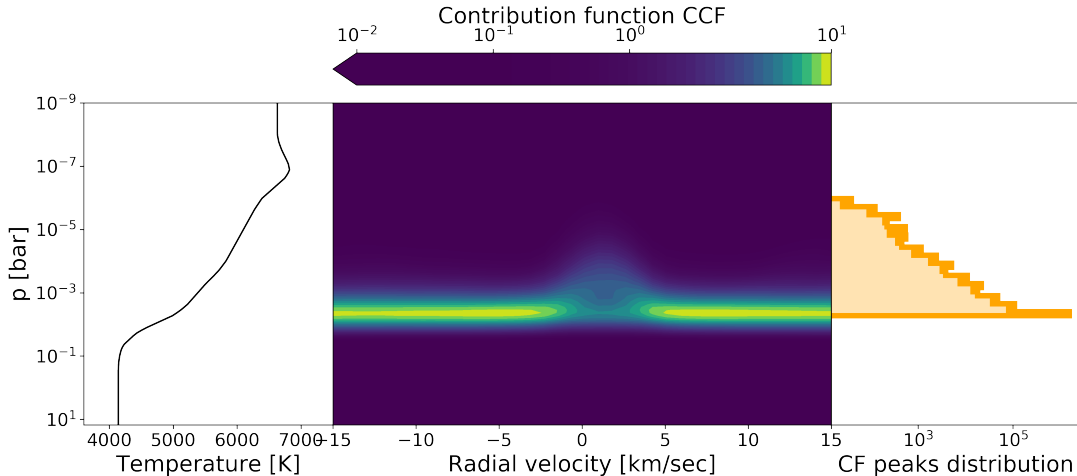
mation of thermal inversions is the lack of molecules with near-infrared opacities, able to radiatively cool the atmosphere. This can be caused by high C/O atmospheres (Mollière et al. 2015; Gandhi & Madhusudhan 2019) and/or by thermal dissociation (Lothringer et al. 2018; Parmentier et al. 2018; Arcangeli et al. 2018), with the latter scenario predicted to be important in ultra-hot Jupiters (Lothringer & Barman 2019; Malik et al. 2019).

### 6.2. On the chemical composition of KELT-9b

Ultimately, we conclude that the average KELT-9b emission line intersected by the G2 mask can be explained with neutral iron as predicted by equilibrium chemistry, with iron abundance compatible with a few times that of the host star. However, our results do not imply a lack of ionized iron lines or other species. Furthermore, with the current analysis, our confidence

intervals on  $\text{VMR}_{\text{Fe}}$  are likely too narrow. This is because (1) we fixed the thermal profile and rotation rate and (2) the choice of a specific mask inherently biases the results by selecting specific pixels within the spectrum.

Looking forward, an application of our method with additional line weighted masks sensitive to different lines, and with additional models exploring different thermal profiles, may provide an avenue to measure an  $[\text{Fe}/\text{H}]$  potentially representative of the whole planetary atmosphere. Indeed, iron condenses through a simple phase transition, passing to the liquid or solid state. When present, iron clouds effectively remove most of the iron from the atmosphere above them (Visscher et al. 2010). The mere presence of iron lines in the atmosphere of a planet indicates the likely absence of deep iron clouds, suggesting that the measured abundance may be repre-



**Figure 4.** Diagnostics of the contribution functions of a model assuming stellar iron abundance and the temperature-pressure profile by Lothringer et al. (2018) (shown in the *left panel*). *Central panel*: Cross-correlation of the contribution function, performed layer by layer. This indicates the relative contribution to the flux as a function of radial velocity rather than wavelength (see Appendix 4.3). The continuum of the cross-correlation function is located around a few bars, and is due to absorption by  $\text{H}^-$ . The peak of the CCF is mostly sensitive to pressure levels around  $10^{-3/-5}$  bar. *Right panel*: For every wavelength channel in our model, we look for the location in pressure of the maximum of the contribution function, and produce an histogram. The two separated peaks show that the continuum originates at the pressure of a few bars, and that the core of most of the iron lines are originated between pressures of  $10^{-3/-5}$  bar (mind the logarithmic scale of the counts).

sentative of the global iron abundance in the planetary atmosphere.

### 6.3. Comparison of transmission and emission spectroscopy of iron lines

The transmission spectrum of the planet atmosphere probes its terminator region, where lower temperatures are expected, which could reflect in different chemical properties of the atmosphere. Hoeijmakers et al. (2019) reported absorption from neutral iron at the terminator of KELT-9b at the millibar level by assuming the pressure level of the planetary continuum. This would be at a similar pressure compared to what we report here looking at the day-side emission line. Hoeijmakers et al. (2019) also reported the detection of ionized iron lines, which they estimated to be at the  $\mu\text{bar}$  level, higher up compared to the pressure level where neutral iron emission lines originate from in our scenario. The

combination of these results covers three orders of magnitude in pressure, although we highlight that we find no evidence for ionized iron with our analysis.

From a geometrical standpoint, transmission spectroscopy is sensitive to lower densities compared to emission spectroscopy. Therefore, the combination of the transmission and emission findings could suggest that neutral iron is depleted at around 0.1 millibar at the terminator compared to the day-side atmosphere of the planet. However, we emphasize that for both the emission and transmission studies, the pressure levels where spectral features originate were calculated by making assumptions regarding the temperature profile and gravity of the planet, and assuming a hydrostatic profile for the atmosphere. Further work to explore the effect of these assumptions is required to properly combine the data sets. Nevertheless, this comparison demonstrates the potential to characterize the 3D structure of the atmosphere of exoplanets by studying them at high spectral resolution both in transmission and emission.

## APPENDIX

### A. A PHYSICAL INTERPRETATION OF THE CROSS-CORRELATION FUNCTION

In this Appendix we aim to provide a physical understanding of the planetary excess observed in Fig. 2. This step is fundamental to properly set-up the simulations to be compared with data. We thus describe in mathematical detail (1) how the observations are related to the planetary and stellar spectrum and (2) the steps undertaken to normalize the spectral observations described in 2.1. The steps involved:

1. a color-correction, to mitigate chromatic losses that change the spectral shape observed over the night. This was particularly important for our observations, due to the failure of the ADC which corrects part of these effects at the telescope level;
2. a rescaling of the spectrum to its continuum in every order, to account for variations of the signal-to-noise overnight;
3. a normalization to the stellar spectrum, obtained directly from the data, to remove stellar lines.

#### A.1. Relation between observations and planetary and stellar spectra

HARPS-N records combined-light observations of the star and planet system at any given time in units of photoelectron counts  $C(\lambda_n, t_i)$ , split in orders  $m$  (e2ds spectra). In other words, the information they contain is the total energy deposited in each pixel  $n$  during the exposure  $i$ . On the other hand, both the PHOENIX models and our radiative transfer code output a spectral flux density, i.e. energy per unit wavelength per unit area per unit time  $F(\lambda_n, t_i)$ . We assume that these quantities are related by:

$$C(\lambda_n, t_i) = \text{LSF} * \{ [R_p^2 \cdot F_p(\lambda_n, t_i) + R_\star^2 \cdot F_\star(\lambda_n)] \cdot \mathcal{A}(\lambda_n, t_i) \cdot \mathcal{B}(t_i) \} (\lambda_n, t_i) \cdot \Delta t_i \cdot \frac{A_{\text{tel}}}{d^2} \cdot \Delta \lambda_n \cdot \mathcal{G}, \quad (\text{A1})$$

where  $\Delta t_i$  is the exposure time,  $A_{\text{tel}}/d^2$  is the ratio of the area of the telescope/distance of the system squared,  $\Delta \lambda_n$  is the wavelength range covered by the pixel  $n$ , and  $\mathcal{G}$  is a gain factor. We added two factors  $\mathcal{A}$  and  $\mathcal{B}$  to indicate chromatic losses ( $\mathcal{A}$ ; e.g. chromatic losses at the fibre entrance due to atmospheric dispersion) and overall flux losses ( $\mathcal{B}$ ; e.g. seeing variations, airmass effects). While  $\mathcal{B}$  is a simple scaling factor between the exposures,  $\mathcal{A}$  changes the shape of the spectrum in each exposure.

The relation is non-linear because the Line Spread Function of the spectrograph is convolved with the received spectral flux density, and the planet and star fluxes are already convolved with the respective rotational broadening kernel. In the rest of the discussion we assume that  $C(\lambda_n, t_i)$  is proportional to  $F(\lambda_n, t_i) = R_p^2 \cdot F_p(\lambda_n, t_i) + R_\star^2 \cdot F_\star(\lambda_n)$ , which we find true at a precision better than 0.1 parts-per-million (see also Pino et al. 2018b).

After having related observations and models, we turn to understanding how the data reduction process that we follow impacts the models in mathematical detail. With this next passage, we get a physical understanding of what the observed cross-correlation function (Fig. 2) means.

#### A.2. Preparation of spectra for cross-correlation

To combine the spectra in order to increase the signal-to-noise ratio, and properly extract the planet signal, the data reduction process aims at removing the time and wavelength dependence of  $\mathcal{A}(\lambda_n, t_i)$  and  $\mathcal{B}(t_i)$ .

The first step is color-correction, which removes the wavelength dependence of these multiplicative noise factors. Colour-correction is performed relative to a template, for which we used the first spectrum of the night, where the ADC was performing the best. We produced a low-resolution (LR) version of each spectrum, with one single point in every order. To remove temporal variations, every low-resolution spectrum is rescaled to its spectral order 48 ( $5,580 \text{ \AA} < \lambda < 5,640 \text{ \AA}$ ):

$$C_{\text{LR}}(\lambda_n, t_i) = \frac{\langle C(\lambda_n, t_i) \rangle_{\text{order } m}}{\langle C(\lambda_n, t_i) \rangle_{\text{order } 48}} = \frac{\langle F(\lambda_n, t_i) \mathcal{A}(\lambda_n, t_i) \rangle_{\text{order } m}}{\langle F(\lambda_n, t_i) \mathcal{A}(\lambda_n, t_i) \rangle_{\text{order } 48}} \quad (\text{A2})$$

where we used Eq. A1, angular brackets indicate average between pixels 1024 and 3072 of each order and we simplified several wavelength independent factors. By assuming that the factor  $\mathcal{A}(\lambda_n, t_i)$  is approximately a constant  $\mathcal{A}(t_i)_m$  over an order  $m$  and that the planet flux is small compared to the star, we obtain a residual curve for each exposure  $i$ :

$$\frac{C_{\text{LR}}(\lambda_n, t_i)}{C_{\text{LR, templ}}(\lambda_n)} = \frac{\mathcal{A}(t_i)_m}{\mathcal{A}_{\text{templ}, m}}. \quad (\text{A3})$$

Eq. A3 represents the variation of each spectrum compared to a template only due to the color effect, and needs to be removed from the spectra. We determined that an interpolation with a sixth order spline in wavelength at each  $\lambda_n$  for each exposure minimizes the residuals. We then obtain a color corrected version of  $C$  by dividing Eq. A1 by Eq. A3:

$$C_{\text{cc}}(\lambda_n, t_i) = \text{LSF} * \{ F(\lambda_n, t_i) \cdot \mathcal{A}_{\text{templ}}(\lambda_n) \cdot \mathcal{B}(t_i) \} \cdot \Delta t_i \cdot \frac{A_{\text{tel}}}{d^2} \cdot \Delta \lambda_n \cdot \mathcal{G}. \quad (\text{A4})$$

Some extra time-dependent (and wavelength independent) factors have been absorbed in  $\mathcal{B}(t_i)$ . We stress that color correction only ensures that the relative shape of spectra is the same and is not enough to perform spectrophotometry. Indeed, any deviation from the real shape of the spectrum is carried over to the other exposures through the factor  $\mathcal{A}_{\text{templ}}(\lambda_n)$ .

Now that the shape of the spectra is adjusted, it is possible to remove the overall flux level fluctuations  $\mathcal{B}(t_i)$ . This is done by rescaling each spectrum order by order to its average:

$$[\text{C}_{\text{cc,r}}(\lambda_n, t_i)]_m = \frac{[\text{C}_{\text{cc}}(\lambda_n, t_i)]_m}{\langle \text{C}_{\text{cc}}(\lambda_n, t_i) \rangle_m} = \frac{[\text{F}(\lambda_n, t_i)]_m}{\langle \text{F}(\lambda_n, t_i) \rangle_m}, \quad (\text{A5})$$

where we have used the independence of  $\mathcal{B}(t_i)$  from wavelength, and assumed that  $\mathcal{A}_{\text{templ}}(\lambda_n)$  can be brought out of the average within order  $m$ . At this point, the spectra have the same level in the continuum and can be combined.

Now, recall that  $\text{F} \propto R_p^2 \cdot \text{F}_p(\lambda_n, t_i) + R_\star^2 \cdot \text{F}_\star(\lambda_n)$ . While the star is assumed to be stable over the course of an observation, the planetary spectral lines move as a result of its Doppler motion, hence its time dependence. Our aim is now to remove  $R_\star^2 \cdot \text{F}_\star$  to isolate the planet signal. This is done by building a master spectrum  $\text{M}_\star$  containing only the stellar spectrum and the planetary continuum, and normalizing each exposure by the master spectrum. As common in the literature, we obtain the master spectrum with a median in time of the color corrected, rescaled spectra Eq. A5. Since the planet moves in time by about 0.5 – 3.5 pixels per exposure, its lines are mostly removed from the master. By splitting the planet flux in its line and continuum contribution ( $\text{F}_{\text{p,lines}}$  and  $\text{F}_{\text{p,cont}}$ ):

$$[\text{M}]_m = \text{med}_t [\text{C}_{\text{cc,r}}(\lambda_n, t_i)]_m \approx \frac{[\text{F}_\star(\lambda_n) + \text{F}_{\text{p,cont}}(\lambda_n)]_m}{\langle \text{F}_\star(\lambda_n) + \text{F}_{\text{p,cont}}(\lambda_n) \rangle_m}, \quad (\text{A6})$$

where we neglected the flux contained in the planetary spectral lines at the denominator. Finally, by dividing  $\text{C}_{\text{cc,r}}(\lambda_n, t_i)$  by the master spectrum, we obtain:

$$[\text{C}_{\text{cc,r,tn}}(\lambda_n, t_i)]_m = \frac{[\text{C}_{\text{cc,r}}(\lambda_n, t_i)]_m}{[\text{M}(\lambda_n)]_m} = \frac{R_p^2 \cdot [\text{F}_{\text{lines p}}(\lambda_n, t_i)]_m}{[R_\star^2 \cdot \text{F}_\star(\lambda_n) + R_p^2 \cdot \text{F}_{\text{p,cont}}(\lambda_n)]_m} + 1. \quad (\text{A7})$$

What we measure, is thus the planetary lines normalized to the stellar plus planetary continuum.

Finally, we applied a high-pass filter by computing the standard deviation of each pixel in time (i.e. across the full spectral sequence) and applying a threshold 3 times above the median level of the noise (the exact choice for the threshold level does not influence the final result). For each exposure and each order, we fitted a second-order polynomial to the spectra after rejecting strong outliers and masked pixels. We then divided the data by the fitted polynomial. Eventually, we applied the cross-correlation function.

The planet continuum itself can not be recovered. Indeed, the rescaling in Eq. A5 must be carried out order by order, because within one order  $\mathcal{A}(\lambda_n, t_i)$  is approximately constant. The same holds for the planetary continuum, which is thus removed from our analysis as a by-product. Alternative approaches use a polynomial normalization, with the same outcome. Recently, [Cauley et al. \(2019\)](#) claimed that they perform flux calibration on Echelle spectra similar to ours. Such an approach has a potentially enormous impact on the study of exoplanet atmospheres with this technique, because it would preserve the planetary continuum, which would already be detectable with currently achieved precisions ([Pino et al. 2018a](#)).

## B. LINE WEIGHTED, BINARY MASK CROSS-CORRELATION FUNCTION

Functionally, this CCF is a weighted average of a wavelength dependent signal  $S(\lambda)$ , in our case the planetary spectrum normalized to the continuum (Sec. A.2), on the spectral lines considered in the mask,

$$\text{CCF}(v) = \frac{\sum_{\text{orders}} w_{\text{order}} \sum_{i=1}^{N_{\text{lines}}} \int_{\text{order}} S(\lambda) \cdot M_i(\lambda)|_v \cdot w_i d\lambda}{\sum_{\text{orders}} w_{\text{order}} \sum_{i=1}^{N_{\text{lines}}} \int_{\text{order}} M_i(\lambda)|_v w_i}. \quad (\text{B8})$$

Within each order, to each of the  $N$  lines considered, we associate a binary mask  $M_i$  that has a value of 1 within a waveband  $0.82 \text{ km sec}^{-1}$  wide (1 HARPS-N nominal pixel) around each considered line shifted to account for a radial velocity  $v$ , 0 elsewhere. Each order is weighted according to the signal-to-noise ratio of the observations

( $w_{\text{order}} = 1/\sigma_{\text{order}}$ , where  $\sigma_{\text{order}}$  is the photometric dispersion of the order computed between pixels 1024 and 3072, and only orders with signal-to-noise ratio larger than 35 were kept), and each line is weighted ( $w_i$ ) according to its information content. In the case of the G2 mask that we used, this is the contrast of the spectral line, but different applications may require different weighting schemes. Since the width of the masks in the wavelength space changes with radial velocity, it is important to compute the normalization at every value of  $v$ .

Computationally, it is convenient to recast Eq. B8 to have an effective weight for every pixel in the detector. Practically, each binary mask can span one or more complete pixels and fractions of pixels at the edges. For a single line  $i$ , we can expand the integral by co-adding contributions from each pixel or pixel part that falls within the binary mask  $M_i$ . If we label each pixel by  $j$ , and call  $\Delta\lambda_j$  the width of the pixel in wavelength space, then pixels entirely within the mask contribute to the spectrum with  $\overline{\Delta\lambda_j} = \Delta\lambda_j$ , while pixels at the edges of the mask contribute with  $\overline{\Delta\lambda_j} < \Delta\lambda_j$ . Thus:

$$\text{CCF}(v) = \frac{\sum_{\text{orders}} \sum_{i=1}^{N_{\text{lines}}} \sum_{j=1}^{N_{\text{pixels in } M_i|v}} S(\lambda_j) \cdot (w_{\text{order}} \cdot w_i \cdot \overline{\Delta\lambda_j})}{\sum_{\text{orders}} \sum_{i=1}^N \sum_{j=1}^{N_{\text{pixels in } M_i|v}} (w_{\text{order}} \cdot w_i \cdot \overline{\Delta\lambda_j})}. \quad (\text{B9})$$

The term in parenthesis is the effective weight for each pixel in each order, and is a unique property of each mask considered. Written in this form, the calculation can be conveniently performed using matrix calculation.

We computed the CCF in each order of each exposure by sliding the binary mask between  $-400 \text{ km s}^{-1}$  and  $400 \text{ km s}^{-1}$  in steps of  $2.7 \text{ km s}^{-1}$  (1 nominal HARPS-N resolution element, containing about 3 nominal HARPS-N pixels). With this choice, we were entitled to treat each CCF point as statistically independent from the others, since their information comes from separate resolution elements. For each exposure, we then obtained a total CCF by summing the CCFs of each single order. With a similar procedure, we computed the normalization at the denominator in Eq. B9.

The peak of the CCF is found at a different position in every exposure, due to the planet motion around its host. The juxtaposition of all exposures provides a planet trace. We then assumed a circular orbit for the planet and shift the CCF in each exposure for different values of the tangential velocity of the planet  $K_p$ . For every combination, we interpolated the total CCFs in each exposure to a common velocity grid, and summed them. The resulting 1D CCF is maximized when the individual exposures are correctly aligned in the rest-frame of the planet.

### C. RADIATIVE TRANSFER CODE

We solved the radiative transfer equation in its integral form, employing a ‘‘linear in optical depth’’ approximation for the source function, which is valid for a non-scattering atmosphere (Toon et al. 1989). We employed 200 logarithmically spaced layers between  $10^5$  bar and  $10^{-12}$  bar, covering the full region where lines are generated with enough spatial resolution. This was verified with a step doubling procedure.

For a given temperature-pressure profile, we assumed equilibrium chemistry and calculated volume mixing ratios using the publicly available `FastChem` code version 2 (Stock et al. 2018; Stock, Kitzmann & Patzer in prep.). Our opacities are calculated by employing the VALD3 database (Piskunov et al. 1995; Ryabchikova et al. 1997; Kupka et al. 1999, 2000; Ryabchikova et al. 2015; Kurucz 2014; Bard et al. 1991; Bard & Kock 1994; Barklem et al. 2000; O’Brian et al. 1991; Fuhr et al. 1988; Kurucz 2010; Blackwell-Whitehead et al. 2006; Nitz et al. 1998; Lawler et al. 2013; Pickering et al. 2001; Martin et al. 1988; Bizzarri et al. 1993; Ryabchikova et al. 1994; Wood et al. 2013; Kurucz 2013; Barklem & Aspelund-Johansson 2005; Kroll & Kock 1987; Pauls et al. 1990; Blackwell et al. 1980; Baschek et al. 1970; Hannaford et al. 1992; Bridges 1973; Ryabchikova et al. 1999; Raassen & Uylings 1998). While the VALD3 database offers line lists for a variety of atomic and molecular species, we limited this study to Fe I, Fe II, Ti I, Ti II, expected to be the most spectrally active species in KELT-9b (Hoeijmakers et al. 2018). We computed opacity tables by broadening the lines with a Voigt profile accounting for thermal and natural broadening, and we used partition functions by Barklem & Collet (2016) to obtain opacities as a function of temperature, over a fine grid in wavelength ( $\Delta\lambda = 0.001 \text{ \AA}$ ) over the full HARPS-N range. At this resolution, the single lines in the atmosphere are resolved by a factor of 20 to 30, making our code effectively line-by-line. Our  $\text{H}^-$  bound-free opacity comes from John (1988), in particular their Eq. (4). We also note a possible imprecision in the units for  $\lambda_0$  and  $\alpha = hc/k_b$  in that paper, which appear to be inconsistent. If  $\lambda_0$  is taken in  $\mu\text{m}$  as the author suggests, the correct value for  $\alpha$  to insert in Eq. (3) is  $1.439 \cdot 10^4$  rather than  $1.439 \cdot 10^8$ . We validated our code by reproducing the position and depth of iron lines in a  $\log g = 4.5$ ,  $T_{\text{eff}} = 4,500 \text{ K}$  PHOENIX model (Husser et al. 2013), adopting the temperature profile provided in the ‘ATMOS.fits’ file. For such a star, iron lines are modelled in LTE, which we also assumed. We did not attempt to reproduce the pressure broadened wings and micro-turbulence broadening in the stellar spectrum, because HARPS-N is only sensitive to the core of the planetary iron lines and micro-turbulence is degenerate with rotational broadening at our level of precision. We also validated

the continuum in our model by reproducing it with petitRADTRANS (Mollière et al. 2019), finding agreement to within a few percent over the HARPS-N range.

#### D. SIGNIFICANCE OF THE DETECTION, MODEL COMPARISON AND CONFIDENCE INTERVALS

This Appendix presents practical details on how we treated  $\log \mathcal{L}$  to (1) assess the significance of our detection, (2) perform model comparison, (3) extract confidence intervals. All of these tasks can be performed using Wilk’s theorem (Wilks 1938). An extensive literature on the topic is available (e.g. Lampton et al. 1976 treats most of these problems in a very clear manner), and we specialize the discussion to our method. We also provide a practical method to marginalize the likelihood distribution.

Given a model  $S$  with  $p$  parameters, a model  $S_{\text{nested}}$  is nested to it if it can be obtained from  $S$  by fixing  $q < p$  parameters. In this case,  $\max \mathcal{L}(S) \geq \max \mathcal{L}(S_{\text{nested}})$ . Wilk’s theorem states that the likelihood ratio test (LRT) metric

$$\text{LRT} = -2 \ln \frac{\max \mathcal{L}(S_{\text{nested}})}{\max \mathcal{L}(S)} = -2 \ln [\max \mathcal{L}(S_{\text{nested}}) - \max \mathcal{L}(S)] \quad (\text{D10})$$

is distributed as a  $\chi^2$  distribution with  $q$  degrees of freedom under the null hypothesis that  $S_{\text{nested}}$  is true.

The application to model comparison is straightforward: in our case,  $S_{\text{Fe}}$  and  $S_{\text{Ti}}$  are nested in  $S_{\text{Fe, Ti}}$ , and  $q = 1$ . The survival function of a  $\chi^2$  distribution with 1 degree of freedom evaluated in LRT gives the probability that the measured LRT difference would be observed by chance alone. A high probability indicates that the null hypothesis that the nested model is true can be excluded. We convert this probability to  $\sigma$  values using a two tailed Gaussian distribution.

To assess the significance of the detection, we created a nested model with  $\text{VMR}_{\text{Fe}} = 0$ . We then compared this to our preferred model  $S_{\text{Fe}}$ . Using the properties of D10, we computed the probability at which the null hypothesis that a model without any spectral line can be excluded (again,  $q = 1$ ).

It is maybe less evident that the same theorem allows us to compute confidence intervals. A clear explanation is found in Cash (1976), which we summarize. Assume that we are interested in the confidence interval on parameter  $\theta$ , which can have values  $(\bar{\theta}_1, \bar{\theta}_2, \dots)$ . First, we fix  $\theta = \bar{\theta}_1$ , and look for the maximum likelihood by varying the rest of the parameters. Practically, this is a nested model with  $q = 1$ . Thus, we can apply Wilk’s theorem to compute the probability that the null hypothesis that  $\theta = \bar{\theta}_1$  is excluded. We then move to the next value of  $\theta$ , and repeat the operation. The locus of  $\theta$  values for which we obtain probabilities lower than a threshold  $\alpha$  gives the corresponding confidence interval.

Given a sufficiently fine grid of likelihoods, another equivalent option is to directly marginalize the likelihood. However, in general,  $\exp(\log \mathcal{L})$  can be a computationally untreatable number. We thus normalize the likelihood to its maximum prior to exponentiating, by computing

$$\bar{\mathcal{L}} = \exp[\log \mathcal{L} - \max \log \mathcal{L}] . \quad (\text{D11})$$

This quantity can then be marginalized, and correctly normalized a-posteriori. The contour levels can be obtained as percentiles of the resulting marginalized distribution.

#### E. CROSS-CORRELATION TO LIKELIHOOD MAPPING BY TO BROGI & LINE (2019)

To check the consistency of our method, we retrieved  $v_{\text{sys}}$  and  $K_p$  using the framework described in Brogi & Line (2019), and the best-fitting model computed and scaled as explained in Sec. 4 and Appendix A. In this scheme, the cross covariance  $R$  between data and the best-fitting model (rather than a binary mask) is computed. As such, the retrieval is model-dependent in line with its main application to determine atmospheric properties of exoplanets. Cross-covariance values are translated into log-likelihood via the formula:

$$\log(L) = -\frac{N}{2} \log[s_f^2 + s_g^2 - 2R], \quad (\text{E12})$$

where  $s_f$  and  $s_g$  are the data and model variance, respectively. A Markov Chain Monte Carlo is driven by the likelihood in Eq. E12, and run via the Python package `emcee`. Confidence intervals are determined by marginalising the posterior distributions and computing confidence intervals consistently with the line-weighted binary mask method (see Sec. D). We compared the likelihood distributions for  $K_p$  and  $v_{\text{sys}}$  obtained with the two methods in Fig. 3. The frameworks give results consistent at  $1\sigma$ . The significance of the detection with the framework by Brogi & Line (2019) is  $10.3\sigma$ , which is higher than the significance found with the line weighted binary mask case, and the confidence intervals on



$K_p$  and  $v_{\text{sys}}$  are consequently tighter. Possible explanations include (1) the larger amount of pixels and line shape information used in the Brogi & Line (2019) case, (2) the fact that, in the line weighted binary mask approach, we do not use a mask tailored to the planetary spectrum but rather a G2 stellar spectrum. A more detailed comparison will be target of dedicated work.

## REFERENCES

- Arcangeli, J., Désert, J.-M., Line, M. R., et al. 2018, *ApJL*, 855, L30
- Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, *A&A*, 558, A33
- Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, *AJ*, 156, 123
- Baranne, A., Mayor, M., & Poncet, J. L. 1979, *Vistas in Astronomy*, 23, 279
- Baranne, A., Queloz, D., Mayor, M., et al. 1996, *A&AS*, 119, 373
- Bard, A., Kock, A., & Kock, M. 1991, *Astron. and Astrophys.*, 248, 315, (BKK)
- Bard, A., & Kock, M. 1994, *Astron. and Astrophys.*, 282, 1014, (BK)
- Barklem, P. S., & Aspelund-Johansson, J. 2005, *Astron. and Astrophys.*, 435, 373, (BA-J)
- Barklem, P. S., & Collet, R. 2016, *A&A*, 588, A96
- Barklem, P. S., Piskunov, N., & O'Mara, B. J. 2000, *Astron. and Astrophys. Suppl. Ser.*, 142, 467, (BPM)
- Baschek, B., Garz, T., Holweger, H., & Richter, J. 1970, *Astron. and Astrophys.*, 4, 229, (BGHR)
- Bizzarri, A., Huber, M. C. E., Noels, A., et al. 1993, *A&A*, 273, 707, (BHN)
- Blackwell, D. E., Shallis, M. J., & Simmons, G. J. 1980, *Astron. and Astrophys.*, 81, 340, (BSScor)
- Blackwell-Whitehead, R. J., Lundberg, H., Nave, G., et al. 2006, *Monthly Notices Roy. Astron. Soc.*, 373, 1603, (BLNP)
- Borsa, F., Rainer, M., Bonomo, A. S., et al. 2019, arXiv e-prints, arXiv:1907.10078
- Bridges, J. M. 1973, in *Phenomena in Ionized Gases, Eleventh International Conference*, ed. I. Štoll, 418–+, (B)
- Brogi, M., de Kok, R. J., Albrecht, S., et al. 2016, *ApJ*, 817, 106
- Brogi, M., & Line, M. R. 2019, *AJ*, 157, 114
- Brogi, M., Snellen, I. A. G., de Kok, R. J., et al. 2013, *ApJ*, 767, 27
- Casasayas-Barris, N., Pallé, E., Yan, F., et al. 2019, *A&A*, 628, A9
- Cash, W. 1976, *A&A*, 52, 307
- Cauley, P. W., Shkolnik, E. L., Ilyin, I., et al. 2019, *AJ*, 157, 69
- Claudi, R., Benatti, S., Carleo, I., et al. 2017, *European Physical Journal Plus*, 132, 364
- Covino, E., Esposito, M., Barbieri, M., et al. 2013, *A&A*, 554, A28
- Donati, J. F., Semel, M., Carter, B. D., Rees, D. E., & Collier Cameron, A. 1997, *MNRAS*, 291, 658
- Dumusque, X. 2018, *A&A*, 620, A47
- Fisher, C., Hoeijmakers, H. J., Kitzmann, D., et al. 2019, arXiv e-prints, arXiv:1910.11627
- Fortney, J. J., Lodders, K., Marley, M. S., & Freedman, R. S. 2008, *ApJ*, 678, 1419
- Fossati, L., Haswell, C. A., Froning, C. S., et al. 2010, *ApJL*, 714, L222
- Fuhr, J. R., Martin, G. A., & Wiese, W. L. 1988, *Journal of Physical and Chemical Reference Data*, Volume 17, Suppl. 4. New York: American Institute of Physics (AIP) and American Chemical Society, 1988, 17, (FMW)
- Gandhi, S., & Madhusudhan, N. 2019, *MNRAS*, 485, 5817
- Gaudi, B. S., Stassun, K. G., Collins, K. A., et al. 2017, *Nature*, 546, 514
- Gibson, N. P., Merritt, S., Nugroho, S. K., et al. 2020, *MNRAS*, 220
- Hannaford, P., Lowe, R. M., Grevesse, N., & Noels, A. 1992, *Astron. and Astrophys.*, 259, 301, (HLGN)
- Haswell, C. A., Fossati, L., Ayres, T., et al. 2012, *ApJ*, 760, 79
- Helling, C., Gourbin, P., Woitke, P., & Parmentier, V. 2019, *A&A*, 626, A133
- Hoeijmakers, H. J., Ehrenreich, D., Heng, K., et al. 2018, *Nature*, 560, 453
- Hoeijmakers, H. J., Ehrenreich, D., Kitzmann, D., et al. 2019, *A&A*, 627, A165
- Hooton, M. J., Watson, C. A., de Mooij, E. J. W., Gibson, N. P., & Kitzmann, D. 2018, *ApJL*, 869, L25
- Hubeny, I., Burrows, A., & Sudarsky, D. 2003, *ApJ*, 594, 1011
- Husser, T. O., Wende-von Berg, S., Dreizler, S., et al. 2013, *A&A*, 553, A6
- Irwin, P. G. J. 2009, *Giant Planets of Our Solar System*, doi:10.1007/978-3-540-85158-5
- John, T. L. 1988, *A&A*, 193, 189
- Kitzmann, D., Heng, K., Rimmer, P. B., et al. 2018, *ApJ*, 863, 183

- Kochukhov, O., Makaganiuk, V., & Piskunov, N. 2010, *A&A*, 524, A5
- Kroll, S., & Kock, M. 1987, *Astron. and Astrophys. Suppl. Ser.*, 67, 225, (KK)
- Kupka, F., Piskunov, N., Ryabchikova, T. A., Stempels, H. C., & Weiss, W. W. 1999, *A&AS*, 138, 119
- Kupka, F. G., Ryabchikova, T. A., Piskunov, N. E., Stempels, H. C., & Weiss, W. W. 2000, *Baltic Astronomy*, 9, 590
- Kurucz, R. L. 2010, Robert L. Kurucz on-line database of observed and predicted atomic transitions, ,
- . 2013, Robert L. Kurucz on-line database of observed and predicted atomic transitions, ,
- . 2014, Robert L. Kurucz on-line database of observed and predicted atomic transitions, ,
- Lampton, M., Margon, B., & Bowyer, S. 1976, *ApJ*, 208, 177
- Lawler, J. E., Guzman, A., Wood, M. P., Sneden, C., & Cowan, J. J. 2013, *ApJS*, 205, 11
- Liddle, A. R. 2007, *MNRAS*, 377, L74
- Lothringer, J. D., & Barman, T. 2019, *ApJ*, 876, 69
- Lothringer, J. D., Barman, T., & Koskinen, T. 2018, *ApJ*, 866, 27
- Malavolta, L., Lovis, C., Pepe, F., Sneden, C., & Udry, S. 2017, *MNRAS*, 469, 3965
- Malik, M., Kitzmann, D., Mendonça, J. M., et al. 2019, *AJ*, 157, 170
- Martin, G., Fuhr, J., & Wiese, W. 1988, *J. Phys. Chem. Ref. Data Suppl.*, 17
- Mollière, P., van Boekel, R., Dullemond, C., Henning, T., & Mordasini, C. 2015, *ApJ*, 813, 47
- Mollière, P., Wardenier, J. P., van Boekel, R., et al. 2019, *A&A*, 627, A67
- Nitz, D. E., Wickliffe, M. E., & Lawler, J. E. 1998, *Astrophys. J. Suppl. Ser.*, 117, 313, (NWL)
- Nugroho, S. K., Kawahara, H., Masuda, K., et al. 2017, *AJ*, 154, 221
- O’Brian, T. R., Wickliffe, M. E., Lawler, J. E., Whaling, W., & Brault, J. W. 1991, *Journal of the Optical Society of America B Optical Physics*, 8, 1185, (BWL)
- Parmentier, V., Line, M. R., Bean, J. L., et al. 2018, *A&A*, 617, A110
- Pauls, U., Grevesse, N., & Huber, M. C. E. 1990, *Astron. and Astrophys.*, 231, 536, (PGHcor)
- Pepe, F., Mayor, M., Galland, F., et al. 2002, *A&A*, 388, 632
- Pickering, J. C., Thorne, A. P., & Perez, R. 2001, *Astrophys. J. Suppl. Ser.*, 132, 403, (PTP)
- Pino, L., Ehrenreich, D., Allart, R., et al. 2018a, *A&A*, 619, A3
- . 2018b, *A&A*, 619, A3
- Piskunov, N. E., Kupka, F., Ryabchikova, T. A., Weiss, W. W., & Jeffery, C. S. 1995, *A&AS*, 112, 525
- Raassen, A. J. J., & Uylings, P. H. M. 1998, *A&A*, 340, 300, (RU)
- Ryabchikova, T., Piskunov, N., Kurucz, R. L., et al. 2015, *PhyS*, 90, 054005
- Ryabchikova, T. A., Hill, G. M., Landstreet, J. D., Piskunov, N., & Sigut, T. A. A. 1994, *MNRAS*, 267, 697, (RHL)
- Ryabchikova, T. A., Piskunov, N. E., Kupka, F., & Weiss, W. W. 1997, *Baltic Astronomy*, 6, 244
- Ryabchikova, T. A., Piskunov, N. E., Stempels, H. C., Kupka, F., & Weiss, W. W. 1999, *Physica Scripta Volume T*, 83, 162, (T83av)
- Schwarz, H., Brogi, M., de Kok, R., Birkby, J., & Snellen, I. 2015, *A&A*, 576, A111
- Sharp, C. M., & Burrows, A. 2007, *ApJS*, 168, 140
- Sing, D. K., Lavvas, P., Ballester, G. E., et al. 2019, *AJ*, 158, 91
- Snellen, I. A. G., de Kok, R. J., de Mooij, E. J. W., & Albrecht, S. 2010, *Nature*, 465, 1049
- Sparks, W. B., & Ford, H. C. 2002, *ApJ*, 578, 543
- Stock, J. W., Kitzmann, D., Patzer, A. B. C., & Sedlmayr, E. 2018, *MNRAS*, 479, 865
- Thorngren, D., Gao, P., & Fortney, J. J. 2019, *ApJL*, 884, L6
- Toon, O. B., McKay, C. P., Ackerman, T. P., & Santhanam, K. 1989, *J. Geophys. Res.*, 94, 16287
- Turner, J. D., de Mooij, E. J. W., Jayawardhana, R., et al. 2020, *ApJL*, 888, L13
- Visscher, C., Lodders, K., & Fegley, Bruce, J. 2010, *ApJ*, 716, 1060
- Wilks, S. S. 1938, *The Annals of Mathematical Statistics*, 9, 60
- Wong, I., Shporer, A., Morris, B. M., et al. 2019, *arXiv e-prints*, arXiv:1910.01607
- Wood, M. P., Lawler, J. E., Sneden, C., & Cowan, J. J. 2013, *ApJS*, 208, 27
- Yan, F., & Henning, T. 2018, *Nature Astronomy*, 2, 714
- Yan, F., Casasayas-Barris, N., Molaverdikhani, K., et al. 2019, *arXiv e-prints*, arXiv:1911.00380

## ACKNOWLEDGMENTS

LP acknowledges helpful discussions with the exoplanet atmosphere team at Anton Pannekoek institute for astronomy, particularly K. Todorov and J. Arcangeli, and with Phil Uttley. We thank the anonymous referee, who's insightful feedback improved the quality and clarity of the manuscript. LP thanks P. Mollière for the prompt support with the validation of the radiative transfer code. The research leading to these results has received funding from the European Research Council (ERC) under the European Unions Horizon 2020 research and innovation programme (grant agreement no. 679633; Exo-Atmos). Based on observations made with the Italian *Telescopio Nazionale Galileo (TNG)* operated on the island of La Palma by the *Fundación Galileo Galilei* of the INAF at the *Observatorio Roque de los Muchachos*. This work has made use of the VALD database, operated at Uppsala University, the Institute of Astronomy RAS in Moscow, and the University of Vienna. This research made use of Astropy,<sup>3</sup> a community-developed core Python package for Astronomy ([Astropy Collaboration et al. 2013, 2018](#)). J.M.D acknowledges support by the Amsterdam Academic Alliance (AAA) Program. M.B. acknowledges support from the UK Science and Technology Facilities Council (STFC) research grant ST/S000631/1. LM acknowledges support from the University of Rome Tor Vergata through Mission: Sustainability 2017 fund. GS acknowledges financial support from Accordo ASI/INAF No. 2013-016-R.0 2013 July 9 and 2015 July 9. R.A. acknowledges the financial support of the SNSF. We acknowledge support from the INAF/FRONTIERA project through the Progetti Premiali” funding scheme of the Italian Ministry of Education, University, and Research.