



<b>Publication Year</b>	2022
<b>Acceptance in OA @INAF</b>	2022-03-16T15:29:29Z
<b>Title</b>	High-resolution detection of neutral oxygen and non-LTE effects in the atmosphere of KELT-9b
<b>Authors</b>	BORSA, Francesco; Luca Fossati; Tommi Koskinen; Mitchell E. Young; Denis Shulyak
<b>DOI</b>	10.1038/s41550-021-01544-4
<b>Handle</b>	<a href="http://hdl.handle.net/20.500.12386/31626">http://hdl.handle.net/20.500.12386/31626</a>
<b>Journal</b>	NATURE ASTRONOMY
<b>Number</b>	6

# High-resolution detection of neutral oxygen and non-LTE effects in the atmosphere of KELT-9b

Francesco Borsa<sup>1,\*</sup>, Luca Fossati<sup>2</sup>, Tommi Koskinen<sup>3</sup>, Mitchell E. Young<sup>4</sup>, and Denis Shulyak<sup>5</sup>

<sup>1</sup>INAF - Osservatorio Astronomico di Brera, Via E. Bianchi 46, 23807 Merate (LC), Italy

<sup>2</sup>Space Research Institute, Austrian Academy of Sciences, Schmiedlstrasse 6, A-8042 Graz, Austria

<sup>3</sup>Lunar and Planetary Laboratory, University of Arizona, 1629 East University Boulevard, Tucson, AZ 85721-0092, USA

<sup>4</sup>Department of Physics, University of Oxford, Denys Wilkinson Building, Keble Road, Oxford, OX1 3RH United Kingdom

<sup>5</sup>Instituto de Astrofísica de Andalucía - CSIC, c/ Glorieta de la Astronomía s/n, 18008 Granada, Spain

\*francesco.borsa@inaf.it

**Oxygen is a constituent of many of the most abundant molecules detected in exoplanetary atmospheres and a key ingredient for tracking how and where a planet formed<sup>1</sup>. In particular, the OI 777.4 nm triplet is used to probe airglow and aurora on the Earth<sup>2</sup> and the oxygen abundance in stellar atmospheres<sup>3,4,5,6</sup>, but has not been detected in an exoplanet atmosphere before. We present a definite ground-based detection of the neutral oxygen 777.4 nm triplet lines in the transmission spectrum of the ultra-hot Jupiter KELT-9b<sup>7</sup>, the hottest known giant planet. The synthetic spectrum computed employing novel non-local thermodynamic equilibrium (NLTE) radiative transfer calculations<sup>8</sup> matches the data significantly better compared to the one computed assuming local thermodynamic equilibrium. These NLTE radiative transfer calculations imply a mass-loss rate of  $10^8$ - $10^9$  kg s<sup>-1</sup>, which exceeds the lower limit of  $10^7$ - $10^8$  kg s<sup>-1</sup> required to facilitate the escape of oxygen and iron from the atmosphere. Assuming a solar oxygen abundance, the NLTE model points towards the need of microturbulence and macroturbulence broadening of  $3.0 \pm 0.7$  km s<sup>-1</sup> and  $13 \pm 5$  km s<sup>-1</sup>, respectively, indicative of the presence of fast winds in the middle and upper atmosphere. Present and upcoming high-resolution spectrographs will allow the detection in other exoplanets of the 777.4 nm OI triplet, which is a powerful tool to constrain the key characteristics of exoplanetary atmospheres when coupled with forward modelling accounting for NLTE effects.**

The atmosphere of KELT-9b<sup>9,10,11,12,13,14,15,16</sup> is characterised by a temperature inversion<sup>17,15,18</sup> with an upper atmospheric temperature of the order of 8000-10000 K<sup>19,11,12,18,8</sup>. In particular, Ref.<sup>8</sup> computed self-consistently the temperature-pressure (TP) profile and composition of the atmosphere of KELT-9b accounting for non-local thermodynamic equilibrium (NLTE) effects, which lead to deviations of the excitation and ionisation states of atoms and molecules from those predicted by the Boltzmann and Saha equations, respectively. They predicted an upper atmospheric temperature about 2000 K hotter than the one computed assuming local thermodynamic equilibrium (LTE). The NLTE transmission spectrum based on this TP profile provides an excellent match to the observed hydrogen Balmer lines<sup>8</sup>. Furthermore, it showed that, as a result of the high atmospheric temperature and of NLTE effects, the triplet of neutral oxygen at ~777.4 nm should be almost as strong as the previously detected H $\alpha$  Balmer line. We analysed public data downloaded from the Calar Alto archive, that were taken with the CARMENES spectrograph during three transits of KELT-9b. CARMENES<sup>20</sup> is a fiber-fed high-resolution echelle spectrograph installed on the 3.5 m telescope at Calar Alto observatory,

whose optical arm covers the 520-960 nm wavelength range with an average resolving power of about 94,600, and is thus well-suited for detecting neutral oxygen at the position of the 777.4 nm triplet.

By using a differential technique and comparing observations taken in- and out-of-transit, we can extract the transmission spectrum of the planetary atmosphere. For each of the three available transit datasets, we independently extracted the planetary transmission spectrum. A telluric correction was not required in the wavelength range covered by the OI triplet (see Methods). We created a stellar spectrum reference by averaging the out-of-transit measurements after moving to the stellar rest frame, then all the spectra were normalised for this reference. All the full-in-transit spectra (i.e., when the planetary disk is completely contained inside the stellar disk) of the three transits were then moved to the planetary rest frame and averaged to create the final planetary transmission spectrum (Fig. 1).

Spurious stellar artifacts, such as Rossiter-McLaughlin (RM) effect or center-to-limb variations can contaminate the transmission spectrum, causing line profile deformation<sup>21</sup> and false detections<sup>22</sup>. We verified that this is not the case here by taking these effects into account in our analysis (Methods). We further checked that the oxygen absorption signal originates from the planet by verifying that the absorption signal follows the expected velocity of the planet. This was done by using the 2D tomographic technique after averaging the OI triplet in velocity space (Fig. 2a) and by computing the  $K_p$ - $V_{\text{sys}}$  map, leading to a detection at a signal-to-noise ratio of 9.2 compatible with the planetary rest frame (Fig. 2b and Methods).

We fitted the line profiles present in the transmission spectrum with Gaussian functions in a Bayesian framework (Methods). We derived for the oxygen triplet an average line contrast of  $0.255 \pm 0.015$  %, and a full width at half maximum (FWHM) of  $20.7 \pm 1.8$  km s<sup>-1</sup>. The line contrast corresponds to an average effective planetary radius of  $\sim 1.17 R_p$ . For a tidally locked rotating atmosphere, such as that expected for KELT-9b, an absorption feature at  $\sim 1.17 R_p$  should have a FWHM of  $\sim 13$  km s<sup>-1</sup>. The larger FWHM obtained from the Gaussian fit points to the presence of additional broadening mechanism(s).

We compared the observations with the synthetic transmission spectrum computed with our NLTE model of the atmosphere, assuming a solar oxygen abundance<sup>8</sup>. Here, our model accounts for the spectral resolution of the instrument and planetary rotation (assuming tidal locking). In line with the data, the model shows prominent absorption by the triplet, but the observed line profiles are broader than the model lines (Methods). To find a better match between the observation and the model, we computed transmission spectra in the same way as described in Ref.<sup>8</sup> applying different microturbulence ( $v_{\text{mic}}$ ; i.e., velocity of gas on a scale smaller than the pressure scale height) and macroturbulence ( $v_{\text{mac}}$ ; i.e., velocity of gas on a scale larger than the pressure scale height) velocity values. We employed a  $X^2$  minimisation routine (Methods) to identify the best fitting  $v_{\text{mic}}$  and  $v_{\text{mac}}$  values and obtained  $v_{\text{mic}} = 3.0 \pm 0.7$  km s<sup>-1</sup> and  $v_{\text{mac}} = 13 \pm 5$  km s<sup>-1</sup> (Fig. 3). The constraint on the microturbulence velocity is tighter than that on the macroturbulence velocity, because  $v_{\text{mic}}$  variations also affect the strength of the lines. Furthermore, within  $2\sigma$ , the observations are also consistent with zero macroturbulence velocity. We note that assuming a higher abundance of oxygen is an alternative to increasing the microturbulence velocity, and thus a somewhat higher oxygen abundance in the model, and/or a lower ionization fraction of oxygen, could also help to improve the fit. A detailed exploration of the parameter space, however, is beyond the scope of the present work. Also,

the value of  $v_{\text{mic}}$  that we obtained agrees well with that used by Ref.<sup>11</sup> to fit the Call infrared triplet in the planetary transmission spectrum assuming solar abundance.

We also compared the observations with a model obtained assuming LTE in computing both atmospheric TP profile and transmission spectrum (Fig. 1, from Ref.<sup>8</sup>), further broadening it accounting for  $v_{\text{mic}}$  and  $v_{\text{mac}}$  values of  $10 \pm 2 \text{ km s}^{-1}$  and  $0 \pm 7 \text{ km s}^{-1}$ , respectively, that we estimated as done for the NLTE spectrum. The LTE profile is significantly weaker and a worse fit compared to the spectrum computed in NLTE, as shown by the respective  $X^2$  values of 275 (NLTE) and 297 (LTE), with a preference for the NLTE model at the  $4.3\sigma$  level (see Methods). Unfortunately, the current modelling setup of Ref.<sup>8</sup> does not enable fine tuning of the abundance of single elements, but the fact that the NLTE transmission spectrum fits well the hydrogen Balmer lines (Methods) and the OI triplet is strong indication that the model is capable of reproducing the average atmospheric properties in the region probed by these lines with solar abundances and that the OI triplet can be used to study NLTE effects in the atmospheres of ultra-hot Jupiters.

Understanding the possible origin of the macroturbulence velocity requires a deeper look at atmospheric structure and circulation. Figure 4 shows the sound speed and Jeans escape parameter as a function of atmospheric structure (Methods). Here, the effective Jeans escape parameter is based on the gravitational potential difference between the given pressure level and the Roche lobe boundary (e.g., Ref.<sup>23</sup>). Radial expansion at a velocity of  $5\text{--}6 \text{ km s}^{-1}$  due to atmospheric escape in the line formation region ( $5 \times 10^{-9} - 5 \times 10^{-5} \text{ bar}$ ) could explain the observed broadening, but is unlikely. The effective Jeans escape parameter reduces to zero at the Roche lobe at a pressure of about  $10^{-10} \text{ bar}$ . To a good approximation, the atmosphere escapes through the Roche lobe around the L1 and L2 points at sound speed. Because gas escaping at other latitudes around the planet is also directed towards the L1 and L2 points, we can estimate the global mass loss rate to within an order of magnitude by assuming that escape proceeds at sound speed through the entire Roche lobe surface<sup>24,25</sup>. Based on this assumption, our model TP profile implies a mass loss rate of  $10^8\text{--}10^9 \text{ kg s}^{-1}$ , which translates to a radial velocity of the order of  $100 \text{ m s}^{-1}$  in the line formation region. This exceeds the mass loss rate that would be required to enable the escape of neutral oxygen and iron from the atmosphere (Methods), but it is not sufficient to explain the broadening of the observed lines. A radial velocity of the order of  $5 \text{ km s}^{-1}$  in the line formation region would imply a much higher mass loss rate, with potentially disastrous consequences for the planet. Global circulation provides a more attractive alternative. Current circulation models for ultra-hot Jupiters predict wind speeds of several  $\text{km s}^{-1}$ , even for planets that are substantially cooler than KELT-9b<sup>26</sup>. Such wind speeds are in good agreement with the observations.

The OI detection in the atmosphere of KELT-9b shows that atomic oxygen can be detected and measured from the ground at optical wavelengths and not exclusively from space in the far-ultraviolet<sup>27,28</sup>. Furthermore, the OI triplet at  $\sim 777.4 \text{ nm}$  lies in a spectral window almost unaffected by telluric absorption and is accessible by several ground-based instruments attached to large telescopes, possibly enabling to detect oxygen from the ground also for other exoplanets orbiting stars fainter than KELT-9. Accounting for the relevant physical processes, including NLTE effects, enables the comparison of forward models with the observations to constrain the key characteristics of exoplanetary atmospheres, including the abundance of oxygen, mass loss, and velocity state, similarly to what is routinely done with stellar atmospheres.

## References

- [1] Madhusudhan, N. Exoplanetary Atmospheres: Key Insights, Challenges, and Prospects. *Annual Review of Astronomy and Astrophysics* **57**, 617 (2019).
- [2] Mendillo, M., Withers, P., and Dalba, P. A., Atomic oxygen ions as ionospheric biomarkers on exoplanets. *Nature Astronomy* **2**, 287-291 (2018).
- [3] Kiselman, D. The 777 nm oxygen triplet in the Sun and solar-type stars, and its use for abundance analysis. *Astronomy and Astrophysics* **275**, 269 (1993).
- [4] Przybilla, N., Butler, K., Becker, S. R., Kudritzki, R. P., and Venn, K. A. Non-LTE line formation for neutral oxygen. Model atom and first results on A-type stars. *Astronomy and Astrophysics* **359**, 1085-1106 (2000).
- [5] Steffen, M., The photospheric solar oxygen project. IV. 3D-NLTE investigation of the 777 nm triplet lines. *Astronomy and Astrophysics* **583**, A57 (2015).
- [6] Sitnova, T.M., Mashonkina, L.I. Influence of Inelastic Collisions with Hydrogen Atoms on Non-LTE Oxygen Abundance Determinations. *Astron. Lett.* **44**, 411–419 (2018).
- [7] Gaudi, B. S. *et al.* A giant planet undergoing extreme-ultraviolet irradiation by its hot massive-star host. *Nature*, **546**, 514–518 (2017).
- [8] Fossati, L., *et al.* Non-local thermodynamic equilibrium effects determine the upper atmospheric temperature structure of the ultra-hot Jupiter KELT-9b. *Astronomy and Astrophysics* **653**, A52 (2021).
- [9] Yan, F. and Henning, T. An extended hydrogen envelope of the extremely hot giant exoplanet KELT-9b. *Nature Astronomy*, **2**, 714-718 (2018).
- [10] Cauley, P. W., *et al.* Atmospheric Dynamics and the Variable Transit of KELT-9 b. *The Astrophysical Journal* **157**, 69 (2019).
- [11] Turner, J. D. *et al.* Detection of Ionized Calcium in the Atmosphere of the Ultra-hot Jupiter KELT-9b. *The Astrophysical Journal*, **888**, L13 (2020).
- [12] Wyttenbach, A., *et al.* Mass-loss rate and local thermodynamic state of the KELT-9 b thermosphere from the hydrogen Balmer series. *Astronomy and Astrophysics* **638**, A87 (2020).
- [13] Hoeijmakers, H. J. *et al.* Atomic iron and titanium in the atmosphere of the exoplanet KELT-9b. *Nature* **560**, 453–455 (2018).
- [14] Hoeijmakers, H. J., *et al.* A spectral survey of an ultra-hot Jupiter. Detection of metals in the transmission spectrum of KELT-9 b”, *Astronomy and Astrophysics* **627**, A165 (2019).
- [15] Pino, L., *et al.* Neutral Iron Emission Lines from the Dayside of KELT-9b: The GAPS Program with HARPS-N at TNG XX. *The Astrophysical Journal* **894**, L27 (2020).
- [16] Borsa, F. *et al.* The GAPS Programme with HARPS-N at TNG. XIX. Atmospheric Rossiter-McLaughlin effect and improved parameters of KELT-9b. *Astronomy and Astrophysics* **631**, A34 (2019).
- [17] Lothringer, J. D., Barman, T., and Koskinen, T., Extremely Irradiated Hot Jupiters: Non-oxide Inversions, H- Opacity, and Thermal Dissociation of Molecules. *The Astrophysical Journal* **866**, 27 (2018).
- [18] Fossati, L., *et al.* A data-driven approach to constraining the atmospheric temperature structure of the ultra-hot Jupiter KELT-9b. *Astronomy and Astrophysics* **643**, A131 (2020).
- [19] García Muñoz, A. and Schneider, P. C. Rapid Escape of Ultra-hot Exoplanet Atmospheres Driven by Hydrogen Balmer Absorption. *The Astrophysical Journal* **884**, L43 (2019).
- [20] Quirrenbach, A. *et al.* CARMENES instrument overview. *Proc. SPIE* **9147**, 91471F (2014).
- [21] Borsa, F. & Zannoni, A. Stellar contributions to the line profiles of high-resolution transmission spectra of exoplanets. *Astronomy and Astrophysics* **617**, A134 (2018).
- [22] Casasayas-Barris, N. *et al.* Is there Na I in the atmosphere of HD 209458b?. Effect of the centre-to-limb variation and Rossiter-McLaughlin effect in transmission spectroscopy studies. *Astronomy and Astrophysics* **635**, A206 (2020).

- [23] Lecavelier des Etangs, A., Vidal-Madjar, A., McConnell, J. C., and Hébrard, G. Atmospheric escape from hot Jupiters. *Astronomy and Astrophysics* **418**, L1–L4 (2004).
- [24] Li, S. L., Miller, N., Lin, D. N. C., and Fortney, J. J. WASP-12b as a prolate, inflated and disrupting planet from tidal dissipation. *Nature* **463**, 1054-1056 (2010).
- [25] Guo, J. H. Escaping Particle Fluxes in the Atmospheres of Close-in Exoplanets. II. Reduced Mass-loss Rates and Anisotropic Winds. *The Astrophysical Journal* **766**, 102 (2013).
- [26] Tan, X. and Komacek, T. D. The Atmospheric Circulation of Ultra-hot Jupiters. *The Astrophysical Journal* **886**, 26 (2019).
- [27] Vidal-Madjar, A. *et al.* Detection of Oxygen and Carbon in the Hydrodynamically Escaping Atmosphere of the Extrasolar Planet HD 209458b. *The Astrophysical Journal* **604**, 69 (2004).
- [28] Ben-Jaffel, L. and Ballester, G. E. Hubble Space Telescope detection of oxygen in the atmosphere of exoplanet HD 189733b. *Astronomy and Astrophysics* **553**, 52 (2013).

**Acknowledgements** F.B. acknowledges support from PLATO ASI-INAF agreement n. 2015-019-R.1-2018. This research has made use of the Spanish Virtual Observatory (<http://svo.cab.inta-csic.es>) supported by the MINECO/FEDER through grant AyA2017-84089.7. T.K. acknowledges support by the NASA Exoplanet Research Program grant 80NSSC18K0569. D.S. acknowledges the financial support from the State Agency for Research of the Spanish MCIU through the "Center of Excellence Severo Ochoa" award to the Instituto de Astrofísica de Andalucía (SEV-2017-0709). M.E.Y. acknowledges funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program under grant agreement No 805445.

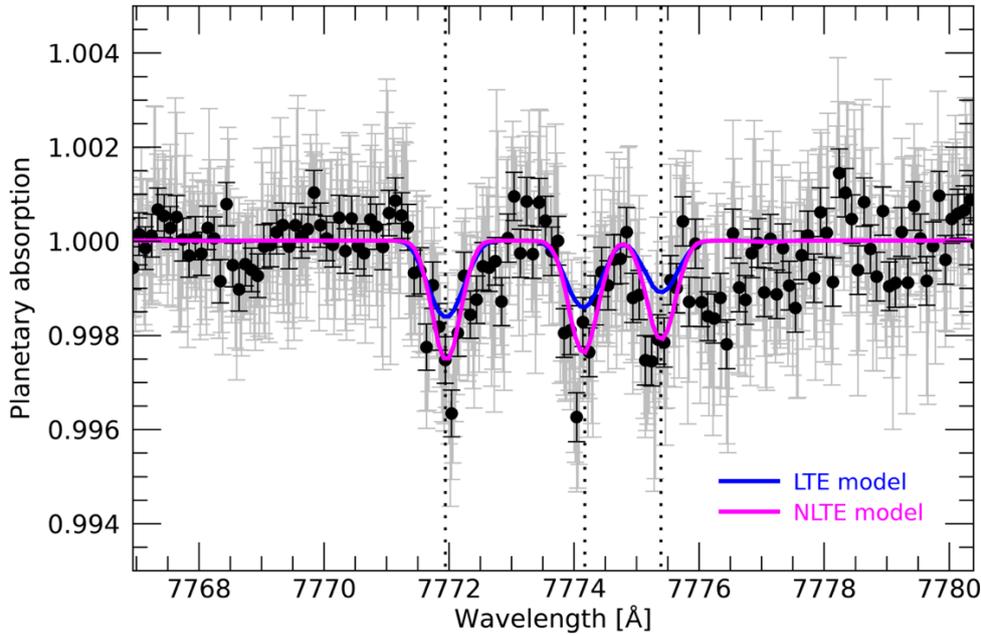
#### **Code availability**

The spectral reduction was done with a self-written IDL script. The stellar model spectrum used for the stellar contamination impact was obtained with the Spectroscopy Made Easy tool, which is publicly available from <http://www.stsci.edu/~valenti/sme.html>. The DE-MCMC routines were taken from EXOFAST, publicly available at <https://github.com/jdeast/EXOFASTv2>.

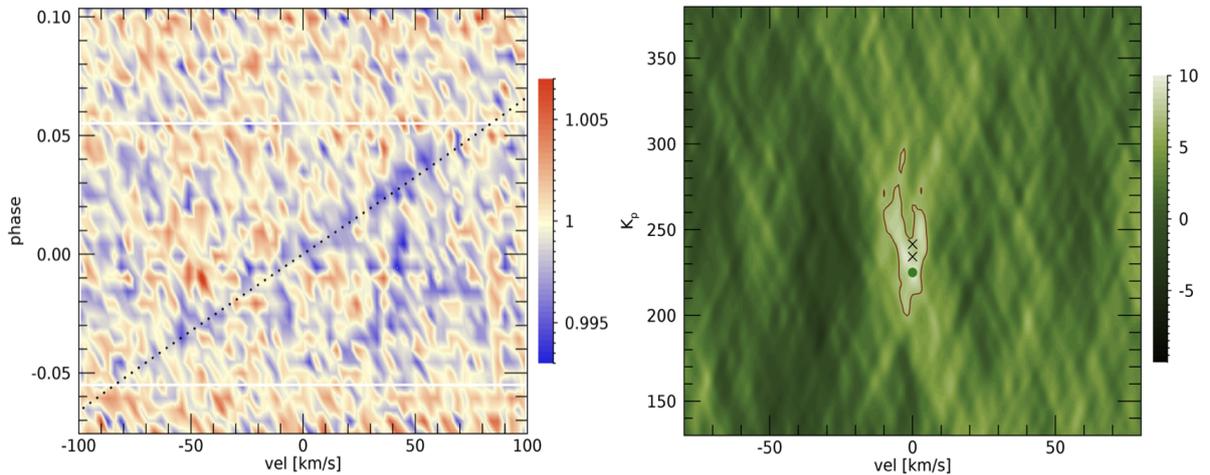
**Data availability** Data used in this work are publicly available from the Calar Alto archive at <http://caha.sdc.cab.inta-csic.es/calto/>.

**Author contributions statement** F.B. carried out the data analysis. L.F. and T.K. performed the theoretical calculations presented in this work. F.B., L.F., and T.K. contributed to the writing of the manuscript. All authors contributed to the interpretation of the data and the results.

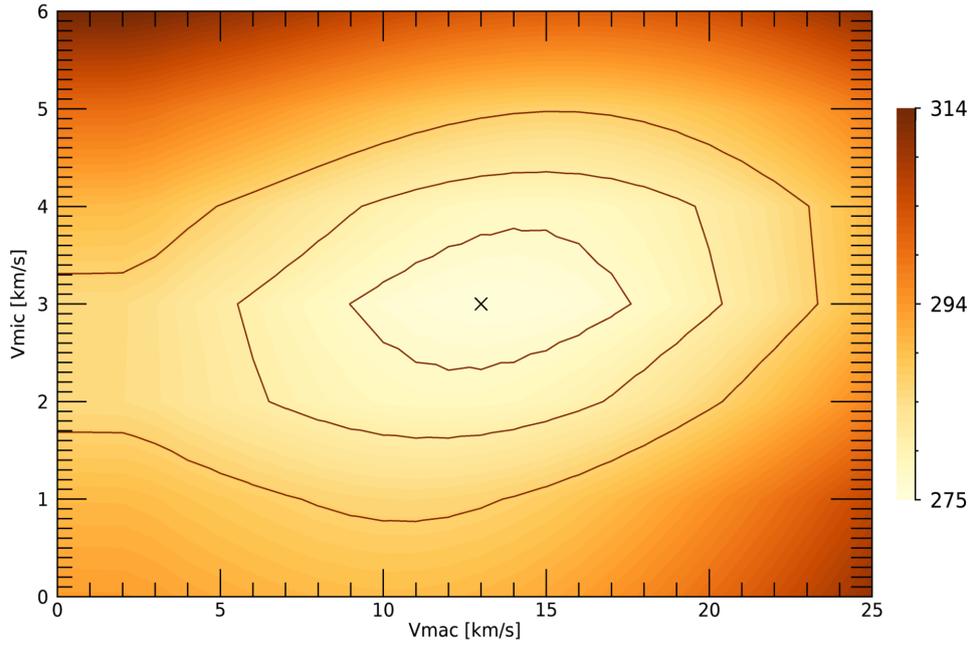
**Competing interests** The authors declare that they have no competing interests.



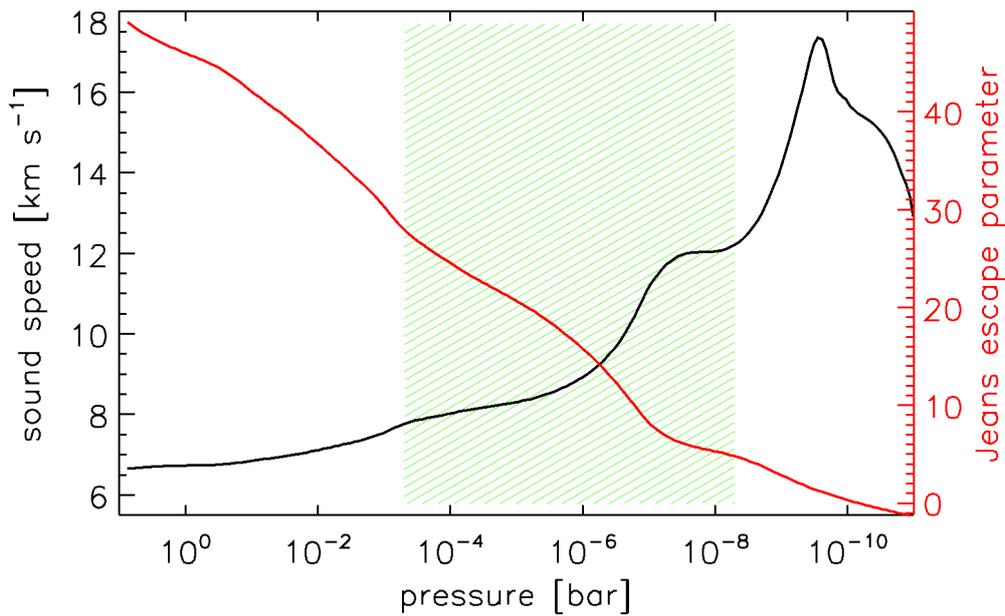
**Figure 1.** Transmission spectrum of KELT-9b around the OI triplet. The magenta line shows the NLTE model of Ref.<sup>8</sup> broadened with the best fit microturbulence ( $v_{\text{mic}}$ ) and macroturbulence ( $v_{\text{mac}}$ ) velocity, while the blue line shows the same for the LTE model (also from Ref.<sup>8</sup>). Grey lines show individual measurements with related uncertainties, black points show an 0.1 Angstrom binning. The vertical dotted lines mark the expected position of the OI triplet lines.



**Figure 2. a,** 2D contour map of the oxygen triplet averaged in velocity space, in the stellar rest frame. The black dotted line shows the expected motion of the planetary signal. The white horizontal lines mark the beginning and end of the transit. The colour scale shows the relative flux variation. **b,**  $K_p$ - $V_{\text{sys}}$  map for the OI triplet. Crosses mark the position of the  $K_p$  (the RV semi-amplitude of the planetary Keplerian motion) determined by Refs.<sup>14,15</sup>. The green dot marks the best  $K_p$  position in the map. The solid line shows the  $3\sigma$  confidence interval. The colour scale shows the signal-to-noise ratio of the detection.



**Figure 3.**  $\chi^2$  minimization map of the NLTE synthetic spectrum with contour lines of  $1\sigma$ ,  $2\sigma$  and  $3\sigma$  confidence intervals for the values of the microturbulence ( $v_{\text{mic}}$ ;  $3.0 \pm 0.7 \text{ km s}^{-1}$ ) and macroturbulence ( $v_{\text{mac}}$ ;  $13 \pm 5 \text{ km s}^{-1}$ ) velocities.



**Figure 4.** Sound speed (black) and Jeans escape parameter (red) as a function of pressure extracted from the NLTE atmospheric structure profile presented by Ref.<sup>8</sup>. The shaded area indicates the main formation region of spectral lines lying in the optical.

## Methods

### Data sample and reduction

We analysed spectroscopic public data downloaded from the Calar Alto archive (<http://caha.sdc.cab.inta-csic.es/calto/>), which contains data covering three transits of KELT-9b taken in the nights of 2017-08-06, 2017-09-21, and 2018-06-16 with CARMENES<sup>20</sup>. This is a fiber-fed high-resolution echelle spectrograph installed on the 3.5 m telescope at Calar Alto observatory. Its optical arm covers the ~520-960 nm wavelength range with an average resolving power of about 94,600. The data of each night (57, 56 and 140 spectra, respectively) cover the transit of the planet. The exposure times used were different, varying from 111 sec to 400 sec. The signal-to-noise ratio (S/N) in the order 39 of the spectrograph, where the oxygen triplet lies, ranges from 30 to 120.

We analysed the reduced archival data, that were extracted with the instrument pipeline version 2.01<sup>29</sup>. The pipeline, after correcting for bias, flat field, and cosmic rays, performs a flat-relative optimal extraction<sup>30</sup> and wavelength calibration<sup>31</sup>. The output is produced in order-by-order maps, and consists of flux, flux-uncertainties, and wavelength (in vacuum and Earth rest frame). Before performing our analysis, for each night we discarded spectra taken at an airmass higher than 2 and with a S/N lower than half of the average S/N of the night (4, 15 and 4 spectra, respectively).

### Transmission spectrum extraction

We focused just on the order number 39 of the spectrograph, after shifting it from vacuum to air wavelengths, and in particular around the 775-780 nm wavelength range, which contains the oxygen triplet (777.194, 777.417, 777.539 nm, in air wavelengths). Because of the fast stellar rotation ( $V_{\text{sin}i} \sim 110 \text{ km s}^{-1}$ ), thus line blending, the stellar OI triplet appears as one broad feature (Supplementary Fig. 1). Throughout, we considered the system parameters given by Ref.<sup>16</sup>.

We extracted the transmission spectrum by following independently for each night a procedure based on that presented by Ref.<sup>32</sup>. The spectra are first shifted to the stellar rest frame by using the barycentric Earth radial velocity given by the pipeline and a Keplerian model of the system, then each spectrum is continuum normalised by performing a linear fit between two wavelength ranges adjacent to the region of interest (776.0-776.3 nm and 778.2-778.4 nm). Finally, we divided all spectra by the average stellar spectrum created combining the out-of-transit spectra. At the end, we moved into the planetary rest frame by shifting all the residual spectra for the theoretical planetary radial velocity, using the planetary orbital  $K_P$  calculated from the orbital motion ( $247.28 \text{ km s}^{-1}$ , Ref.<sup>16</sup>). The final rebinning is done on a wavelength step of 0.02 Angstrom, chosen for being close to the mean resolution step in the range of interest. We then created the transmission spectrum with a weighted average of all full-in-transit residual spectra.

### Telluric correction

We performed telluric correction using a scaling relation between airmass and telluric line strength<sup>33,34,35</sup>, rescaling all the normalised stellar spectra as if they were obtained at the airmass of the center of the transit. We remark that no telluric line was standing beyond the noise level in the considered wavelength range. We further checked for telluric lines in this wavelength range in archival CARMENES spectra of the telluric standard 109 Vir, without finding any. We thus decided not to apply our telluric correction to avoid increasing the noise

in the final transmission spectrum, but the final results are independent from this choice. We further checked if any water micro-telluric could be present in the wavelength range we analysed employing the ESO Sky Model Calculator (<https://www.eso.org/observing/etc/bin/gen/form?INS.MODE=swspectr+INS.NAME=SKYCALC>). Rapid water vapour changes in the Earth atmosphere could make the depth of the telluric lines not follow the correlation with airmass, and could thus be missed by our telluric correction. We created models by selecting two very different values of precipitable water vapour (5 and 20 mm), and simulated variability between these values during the transits, shifting the telluric models using the values of  $V_{\text{sys}}$  and barycentric Earth radial velocity for the different nights and convolving them with the instrumental broadening. Water micro-tellurics could partly overlap with the reddest line of the OI triplet, we can thus not completely exclude that this line could be partly contaminated. We however confirm that the first two lines of the triplet are instead free also from water micro-tellurics (Supplementary Fig. 2).

Given the position of these water micro-tellurics in the transmission spectrum (about 777.6-777.7 nm), we can argue that their possible variation could be the cause of the imperfect normalisation on the right hand side of the triplet. As a test to see the possible impact on our analysis, we recomputed the  $X^2$  map of Fig. 3 by comparing the planetary atmospheric models and the observed transmission spectrum only blueward of 7775 Angstrom, thus excluding the third line of the OI triplet, obtaining that our final results remain unchanged, confirming the  $v_{\text{mic}}$  and  $v_{\text{mac}}$  values we found.

As for telluric emission, we checked fiber B of the instrument, which was pointing at the sky during the observations. No emission is present beyond the noise when co-adding the spectra in the telluric reference frame for each night.

### **Stellar contamination in the transmission spectrum**

The host star is not an homogeneous disk, but has a surface brightness which rotates and changes as a function of the distance from the center. Effects such as center-to-limb variation (CLV) and stellar rotation (the RM effect) may lead to spurious signals in the transmission spectrum, possibly causing false detections<sup>22</sup> and wrong line-profile estimates<sup>21</sup>.

We modeled these effects using the approach of Ref.<sup>36</sup> in which the star is modeled as a disk divided in sections of  $0.01 R_s$ , each section having its value of projected rotational velocity (propagated using the  $V \sin i$  value) and  $\mu$  ( $\mu = \cos(\theta)$ , with  $\theta$  being the angle between the normal to the stellar surface and the considered line of sight). A spectrum is assigned to each point of the grid, by quadratically interpolating on  $\mu$  and Doppler-shifting the model spectra created using Spectroscopy Made Easy<sup>37</sup> with ATLAS12 stellar atmospheric models (assuming solar abundances and LTE) and the line list from the VALD database<sup>38</sup>. The model spectra are created for 21 different  $\mu$  values and with null rotational velocity, and adapted to the resolving power of the instrument.

The stellar spectrum for each orbital phase during the transit of the planet is then calculated as the average of the non-occulted modeled sections. We then modeled the CLV and RM effect influence on the transmission spectrum by dividing all spectra by a master out-of-transit stellar spectrum, moving into the planetary rest frame and averaging all the full-in-transit spectra.

The magnitude of the stellar contamination given by CLV and RM effect is found to be well within the noise of the observations, demonstrating that the oxygen detection is not an artifact (Supplementary Fig. 3). Despite this, we applied the correction, normalising the transmission spectrum for the stellar contamination model.

## 2D tomographic confirmation

Spurious stellar contamination beyond the one given by CLV and rotation may still exist, as the stellar line profiles and/or not perfectly corrected tellurics may show variability during transit and cause spurious effects<sup>39,21,40,41</sup>. We thus verified that the signal belongs to the planetary rest frame by looking at the 2D tomographic map, where it is clear that the absorption signal follows the planetary rest frame and not the stellar one. To increase the S/N, we passed from wavelengths to velocities by using Doppler's law, centered on each line of the oxygen triplet, and then averaged the three lines. Although aliases are present due to the proximity of the lines, the detected absorption is of planetary nature (Fig. 2a).

We further checked the rest frame of the signal by computing the  $K_p$ - $V_{\text{sys}}$  map (Fig. 2b). This is created by cross-correlating the residual spectra in the stellar rest frame (i.e. the spectra after dividing by the average out-of-transit stellar spectrum) with a template modelled with three normalised Gaussian functions, centered to the rest frame of the OI triplet and considering the FWHM given by the instrumental resolution. We then averaged the in-transit cross-correlation functions after shifting them in the planetary rest frame, for a range of  $K_p$  values from 100 to 400 km s<sup>-1</sup> in steps of 1 km s<sup>-1</sup>. This is done by subtracting the planetary radial velocity calculated for each spectrum as  $V_p = K_p \times \sin 2\pi\phi$ , with  $\phi$  the orbital phase. We thus created the  $K_p$  versus  $V_{\text{sys}}$  map to verify the real planetary origin of the signal. We evaluated the noise by calculating the standard deviation of the map taken in the range  $K_p=[150:360]$  km s<sup>-1</sup> and  $V_{\text{sys}}=[20:75]$  km s<sup>-1</sup>, far from where any stellar or planetary signal is expected. We note that this noise sampling is not perfect, as the map shows strong aliasing and correlated noise. While the  $K_p$  can not be determined with high precision since the model only has three lines, the best  $K_p$  position of the signal is compatible with the one determined in other works analysing metals in the atmosphere of KELT-9b<sup>14,15</sup>, thus further validating the planetary nature of the OI signal. The detection is at S/N=9.2.

We also performed a comparison between the in-trail ( $|v| < 6$  km s<sup>-1</sup>) and out-of-trail ( $|v| > 20$  km s<sup>-1</sup>) values shown in Fig. 2 in the planetary rest frame. A Welch t-test on the data rejects the hypothesis that the two samples are drawn from the same parent distribution at the  $8.4\sigma$  level, although we remind the reader that the strong aliasing slightly biases this statistics making the distributions not perfectly Gaussian.

## Planetary absorption lines Gaussian fit

The final transmission spectrum presents an imperfect normalisation just on the right of the oxygen triplet, i.e. at  $\sim 777.65$  nm, that is possibly caused by micro-telluric contamination (Sect. Telluric correction). Indeed, the synthetic spectrum of Ref.<sup>8</sup>, which accounts for the lines of all metals up to Zn, does not show any feature at this wavelength.

We fitted the sum of three Gaussian functions to the OI absorption features (Supplementary Fig. 4), imposing the same full width at half maximum (FWHM) for each line to avoid the imperfect normalisation redward of the triplet from influencing the fit of the 777.539 nm line. The fit was performed in a Bayesian framework employing a differential evolution Markov chain Monte Carlo (DE-MCMC) technique<sup>42,43</sup>, running ten DE-MCMC chains of 100,000 steps and discarding the burn-in. The medians and the 15.86% and 84.14% quantiles of the posterior distributions were taken as the best values and  $1\sigma$  uncertainties. The results of the fit are shown in Supplementary Table 1. The line contrasts were translated into an average effective planetary radius of  $R_{\text{eff}} \sim 1.17 R_p$  by assuming  $(R_{\text{eff}}/R_p)^2 = (\delta+h)/\delta$ , where  $\delta$  is the transit depth and  $h$  is the line contrast.

Unfortunately we are not sensitive to any possible asymmetry of the lines given the S/N available. When fitting the lines with skewed Gaussian functions, the skew parameter remains unconstrained.

### Line broadening of the model

We compared the observed transmission spectrum of the OI triplet to the NLTE synthetic transmission spectrum that guided our observations<sup>8</sup>. After having added instrumental (using a Gaussian kernel with FWHM of 3.17 km s<sup>-1</sup>) and rotational broadening (assuming tidal locking at a radius of 1.17 R<sub>p</sub> and a rotational profile calculated as in Ref.<sup>39</sup>) to the synthetic line profiles, the synthetic spectrum appeared to be narrower than the observation. Line broadening due to the exposure time, caused by the fact that the planet moves during each exposure, has a negligible impact on the overall broadening (i.e. ~4.7, 3.5, and 1.3 km s<sup>-1</sup> on average for exposure times of 400, 300, and 111 sec, respectively).

Therefore, we recomputed further transmission spectra models in the same way as described in Ref.<sup>8</sup> applying different microturbulence velocity ( $v_{\text{mic}}$ ) values ranging between 1 and 14 km s<sup>-1</sup> (in steps of 1 km s<sup>-1</sup>), further adding macroturbulence velocity ( $v_{\text{mac}}$ ) broadening with  $v_{\text{mac}} < 25$  km s<sup>-1</sup> (in steps of 1 km s<sup>-1</sup>), finally looking for the  $v_{\text{mic}}$  and  $v_{\text{mac}}$  pair leading to match the observations by using X<sup>2</sup> minimization. We obtained best fitting values of  $v_{\text{mic}} = 3.0 \pm 0.7$  km s<sup>-1</sup> and  $v_{\text{mac}} = 13 \pm 5$  km s<sup>-1</sup> (Figure 3). Remarkably, the inclusion of these broadening values does not worsen the fit of the hydrogen Balmer lines (Supplementary Fig. 5).

We followed the same procedure on the LTE transmission spectrum of Ref.<sup>8</sup> to estimate the  $v_{\text{mic}}$  and  $v_{\text{mac}}$  values best fitting the observations obtaining  $10 \pm 2$  km s<sup>-1</sup> and  $0 \pm 7$  km s<sup>-1</sup>, respectively (Supplementary Fig. 6). However, the LTE transmission spectrum minimising the X<sup>2</sup> value is a significantly worse fit to the observation compared to the NLTE synthetic spectrum. We compared the obtained X<sup>2</sup> values of 275 (NLTE model) and 297 (LTE model) by using a likelihood ratio test. With two-degrees of freedom ( $v_{\text{mic}}$  and  $v_{\text{mac}}$ ), we find a p-value of  $1.67 \times 10^{-5}$ , that corresponds to a  $4.3\sigma$  preference for the NLTE model when assuming that the uncertainties have been correctly estimated.

### Mass loss rate

To aid the interpretation of these results, we employed the atmospheric model presented by Ref.<sup>8</sup> to estimate the mass loss rate. Figure 4 shows the effective Jeans escape parameter and sound speed in the atmosphere based on our model TP profile. We remind the reader that the sound speed represents the largest possible atmospheric radial velocity below the L1 point. The effective Jeans escape parameter is

$$X = \frac{m\Delta\phi}{kT}$$

where  $\Delta\phi$  is the gravitational potential difference between the pressure level and the Roche lobe. At the Roche lobe,  $X = 0$  and we can estimate the mass loss rate as

$$\frac{dM}{dt} \simeq 4\pi Y R_{L1} \rho \frac{C}{2\sqrt{\pi}}$$

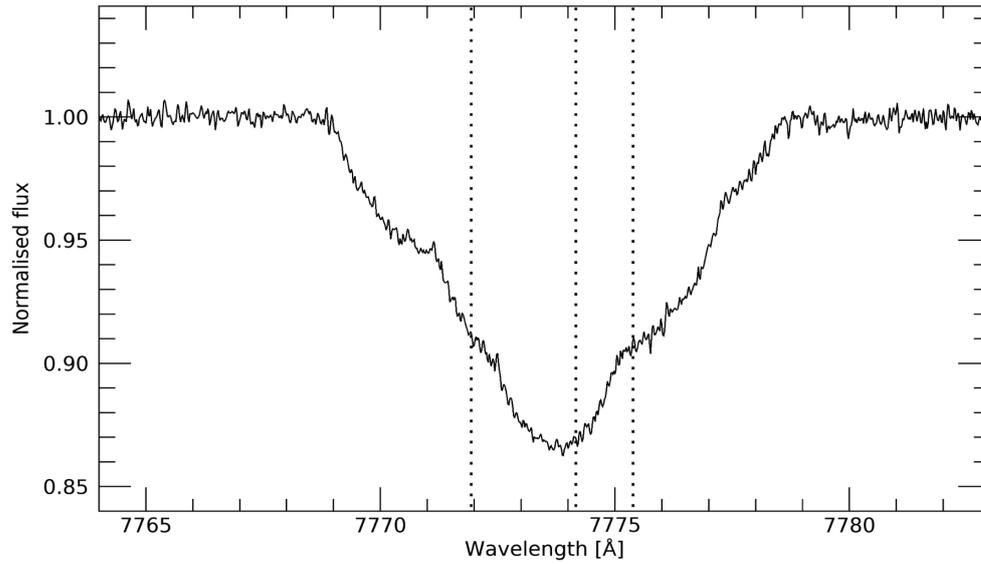
where  $R_{L1}$  is the radius at the L1 point,  $Y = 2R_{L1}/3$  is the polar radius and  $C$  is the thermal speed. The L1 point in our model is located at 2.6 R<sub>p</sub> with a pressure of  $1.17 \times 10^{-10}$  bar where the thermal speed is 14 km s<sup>-1</sup> and mass density is  $1.17 \times 10^{-13}$  kg m<sup>-3</sup>. The resulting value for the mass loss rate is  $5 \times 10^8$  kg s<sup>-1</sup>. We expect this estimate to be accurate roughly to an order of magnitude<sup>24,25</sup>.

The cross-over mass concept<sup>44</sup> can be used to derive the minimum mass loss rate of hydrogen required to enable atomic oxygen and iron to escape the atmosphere<sup>45</sup>:

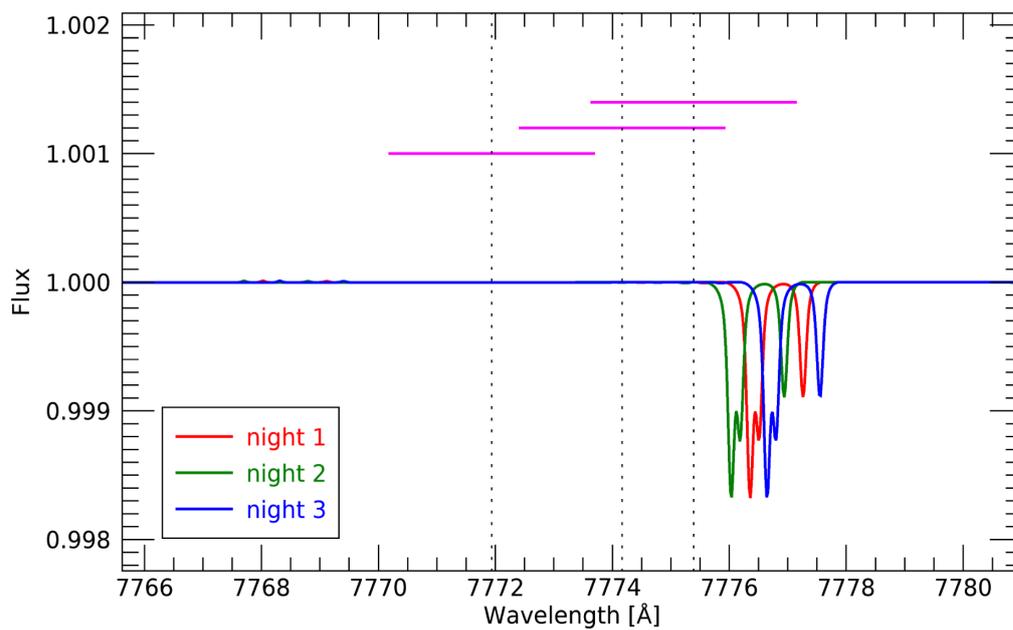
$$\frac{dM}{dt} \simeq 4\pi m_H^2 G M_p (M_c - 1) \frac{nD}{kT}$$

where  $m_H$  is the mass of hydrogen,  $M_c$  is the mass of the heavy element in units of  $m_H$  and  $nD$  is the product of the total number density and the mutual diffusion coefficient. Using values of  $nD \simeq 1.7 \times 10^{22} \text{ m}^{-1} \text{ s}^{-1}$  for O-H<sup>+</sup> collisions or  $nD \simeq 1.4 \times 10^{22} \text{ m}^{-1} \text{ s}^{-1}$  for O-H collisions gives a limiting mass loss rate of  $2\text{-}3 \times 10^7 \text{ kg s}^{-1}$  for neutral oxygen. The value obtained for iron is about  $10^8 \text{ kg s}^{-1}$ .

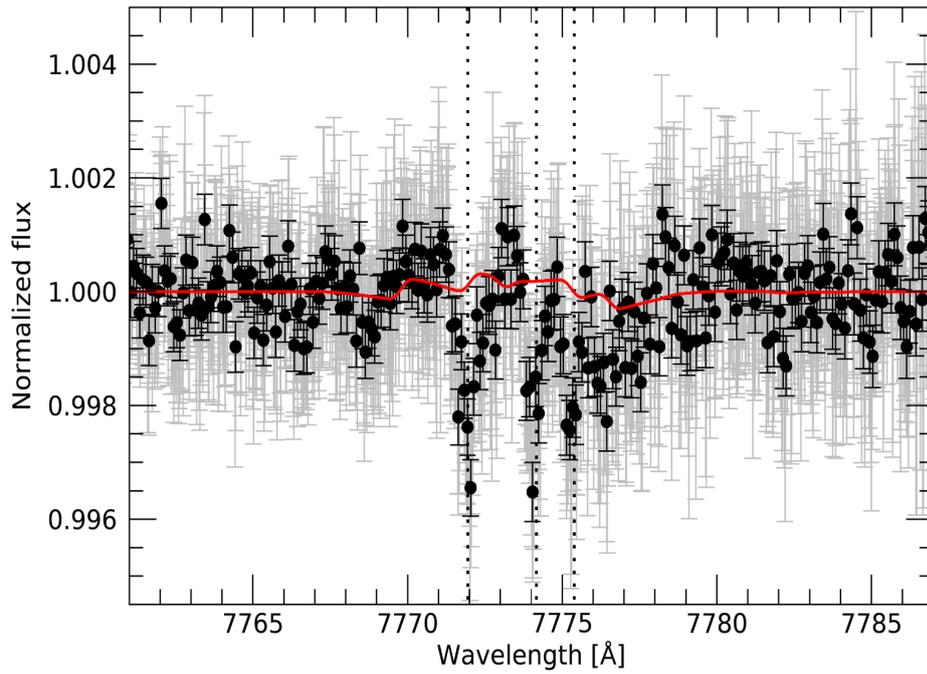
We note that our atmosphere model is hydrostatic and does not include escape. The atmosphere begins to deviate from hydrostatic equilibrium and adiabatic cooling due to expansion becomes significant once the outflow velocity reaches a substantial fraction of the sound speed. Based on a mass loss rate of  $5 \times 10^8 \text{ kg s}^{-1}$  and atmospheric structure predicted by our model, this would happen at  $10^{-10}\text{-}10^{-9}$  bar, given that  $4\pi\rho v r^2$  is constant with radius. This location is close to our upper boundary, and above the region probed by the oxygen line profile. Detailed models of escape that include heavy elements are required to calculate the mass loss rate more precisely and interpret transit signatures that probe higher altitudes in the upper atmosphere.



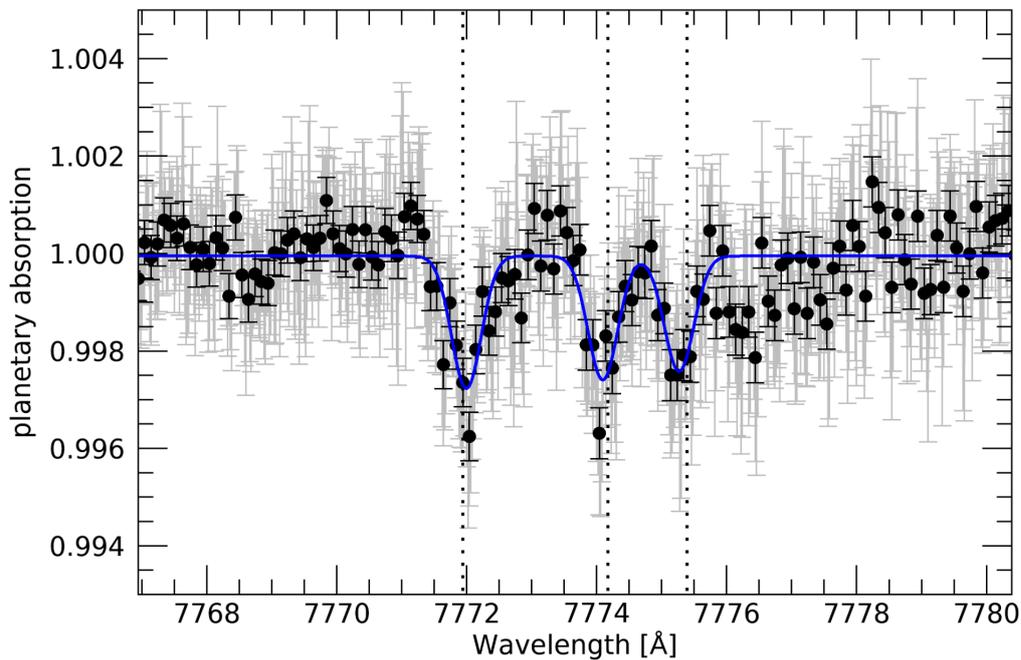
**Supplementary Figure 1.** Normalised master stellar spectrum in the region of the oxygen triplet. The vertical dotted lines mark the position of the three lines of the triplet.



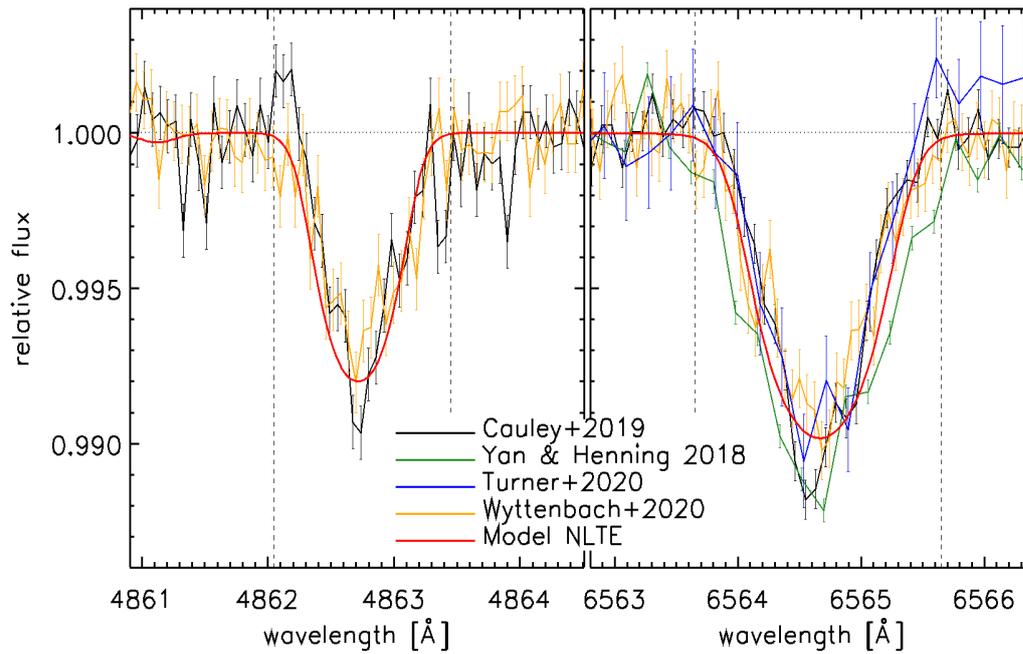
**Supplementary Figure 2.** Position of telluric contamination possibly caused by strong precipitable water vapour changes in the Earth atmosphere toward the target during the observations, in the telluric rest frame shifted for the  $V_{\text{sys}}$ . The magenta horizontal lines show the wavelength range spanned by the planetary motion during the transit for the three lines of the OI triplet. The vertical dotted lines mark the position of the three lines of the triplet.



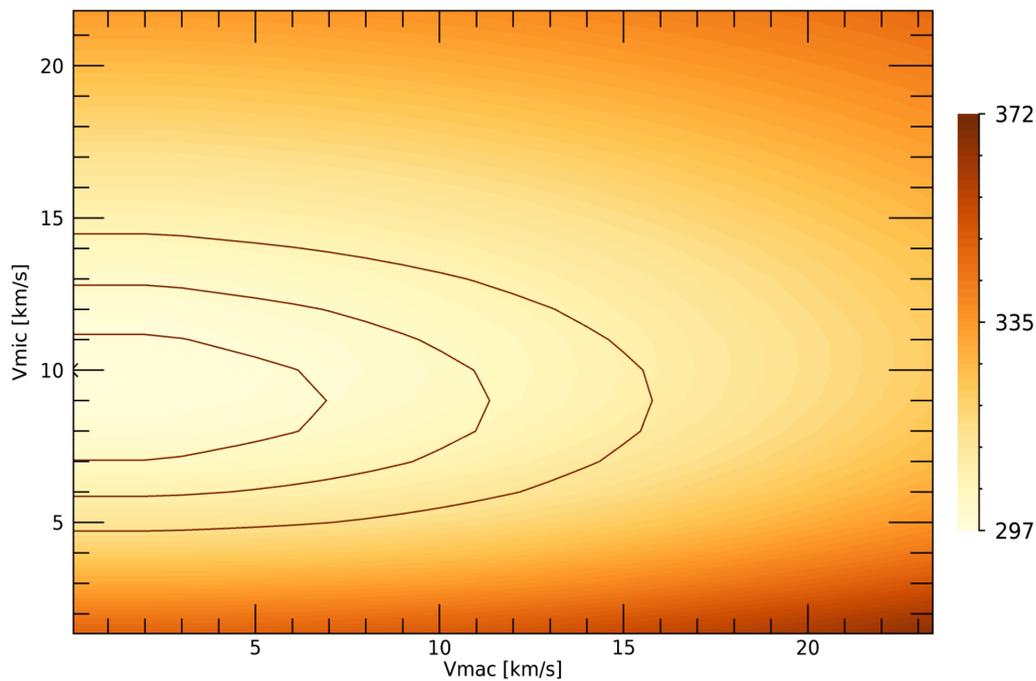
**Supplementary Figure 3.** Raw transmission spectrum of KELT-9b around the OI triplet, before correcting for the stellar contamination of center-to-limb variation and RM effect (red line). The vertical dashed lines indicate the expected position of the OI triplet in the planetary rest frame.



**Supplementary Figure 4.** Transmission spectrum of KELT-9b around the OI triplet. The blue line shows the Gaussian best fits to each line. Black points show an 0.1 Angstrom binning. The vertical dotted lines mark the expected position of the OI triplet lines.



**Supplementary Figure 5.** Comparison between the NLTE synthetic transmission spectrum accounting for  $v_{\text{mic}}=3 \text{ km s}^{-1}$  and  $v_{\text{mac}}=13 \text{ km s}^{-1}$  (red) and the observations<sup>9,10,11,12</sup> in the region of the H $\alpha$  (right) and H $\beta$  (left) Balmer lines.



**Supplementary Figure 6.**  $X^2$  minimization map of the LTE synthetic spectrum with contour lines of  $1\sigma$ ,  $2\sigma$  and  $3\sigma$  confidence intervals for the values of the microturbulence ( $v_{\text{mic}}$ ) and macroturbulence ( $v_{\text{mac}}$ ) velocities.

**Supplementary Table 1.** Summary of the results obtained from the Gaussian fits of the OI triplet. The third line could be partly influenced by water micro-tellurics.

$\lambda_{\text{air}}$ [Angstrom]	contrast [%]	center [km s <sup>-1</sup> ]	FWHM [km s <sup>-1</sup> ]
7771.94	0.272±0.024	2.1±1.1	20.7±1.8
7774.17	0.254±0.027	-2.9±1.1	
7775.39	0.237±0.027	-4.7±1.2	
triplet	0.255±0.015	-1.7±0.7	20.7±1.8

- [29] Caballero, J. A. *et al.* CARMENES: data flow. *Observatory Operations: Strategies, Processes, and Systems V, Proc. SPIE* 9910. (2016).
- [30] Zechmeister, M., Anglada-Escudé, G., and Reiners, A. Flat-relative optimal extraction. A quick and efficient algorithm for stabilised spectrographs. *Astronomy and Astrophysics* **561**, 59 (2014).
- [31] Bauer, F. F., Zechmeister, M., and Reiners, A. Calibrating echelle spectrographs with Fabry-Pérot etalons. *Astronomy and Astrophysics* **581**, 117 (2015).
- [32] Wyttenbach, A., Ehrenreich, D., Lovis, C., Udry, S. & Pepe, F. Spectrally resolved detection of sodium in the atmosphere of HD 189733b with the HARPS spectrograph. *Astronomy and Astrophysics* **577**, A62 (2015).
- [33] Snellen, I. A. G., Albrecht, S., de Mooij, E. J. W., and Le Poole, R. S. Ground-based detection of sodium in the transmission spectrum of exoplanet HD 209458b. *Astronomy and Astrophysics* **487**, 357 (2008).
- [34] Vidal-Madjar, A., *et al.* The Earth as an extrasolar transiting planet. Earth's atmospheric composition and thickness revealed by Lunar eclipse observations. *Astronomy and Astrophysics* **523**, 57 (2010).
- [35] Astudillo-Defru, N. and Rojo, P. Ground-based detection of calcium and possibly scandium and hydrogen in the atmosphere of HD 209458b. *Astronomy and Astrophysics* **557**, 56 (2013).
- [36] Yan, F., Pallé, E., Fosbury, R. A. E., Petr-Gotzens, M. G., and Henning, T. Effect of the stellar absorption line centre-to-limb variation on exoplanet transmission spectrum observations. *Astronomy and Astrophysics* **603**, 73 (2017).
- [37] Piskunov, N. and Valenti, J. A. Spectroscopy Made Easy: Evolution. *Astronomy and Astrophysics* **597**, 16 (2017).
- [38] Ryabchikova, T. *et al.* A major upgrade of the VALD database. *Phys. Scr.* **90**, 054005 (2015).
- [39] Brogi, M. *et al.* Rotation and Winds of Exoplanet HD 189733 b Measured with High-dispersion Transmission Spectroscopy. *The Astrophysical Journal* **817**, 106 (2016).
- [40] Palle, E. *et al.* Transmission spectroscopy and Rossiter-McLaughlin measurements of the young Neptune orbiting AU Mic. *Astronomy and Astrophysics* **643**, 25 (2020).
- [41] Benatti, S., *et al.* Constraints on the mass and atmospheric composition and evolution of the low-density young planet DS Tuc A b. *Astronomy and Astrophysics* **650**, A66 (2021).
- [42] Ter Braak, C. J. F. A Markov Chain Monte Carlo version of the genetic algorithm Differential Evolution: easy Bayesian computing for real parameter spaces. *Statistics and Computing* **16**, 239 (2006).
- [43] Eastman, J., Gaudi, B. S., and Agol, E. EXOFAST: A Fast Exoplanetary Fitting Suite in IDL. *Publications of the Astronomical Society of the Pacific* **125**, 923 (2013).
- [44] Hunten, D. M., Pepin, R. O., and Walker, J. C. G. Mass fractionation in hydrodynamic escape. *Icarus* **69**, 532 (1987).
- [45] Koskinen, T. T., Yelle, R. V., Lavvas, P., and Y-K. Cho, J. Electrodynamics on Extrasolar Giant Planets. *The Astrophysical Journal* **796**, 16 (2014).