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

Francesco Borsa1, Luca Fossati2, Tommi Koskinen3, Mitchell E. Young4 and Denis Shulyak5

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. In particular, the O i 777.4 nm triplet is used to probe airflow and aurora on the Earth and the oxygen abundance in stellar atmospheres, 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 ultrahot Jupiter KELT-9b, the hottest known giant planet. The synthetic spectrum computed employing novel non-local thermodynamic equilibrium radiative transfer calculations matches the data significantly better than that computed assuming local thermodynamic equilibrium. These non-local thermodynamic equilibrium radiative transfer calculations imply a mass-loss rate of $10^8 - 10^9$ kg s$^{-1}$, which exceeds the lower limit of $10^7 - 10^8$ kg s$^{-1}$ required to facilitate the escape of oxygen and iron from the atmosphere. Assuming a solar oxygen abundance, the non-local thermodynamic equilibrium model points towards the need for microturbulence and macroturbulence broadening of $3.0 \pm 0.7$ km s$^{-1}$ and $13 \pm 5$ km s$^{-1}$, 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 O i triplet, which is a powerful tool to constrain the key characteristics of exoplanetary atmospheres when coupled with forward modelling accounting for non-local thermodynamic equilibrium effects.

The atmosphere of KELT-9b is characterized by a temperature inversion with an upper atmospheric temperature of the order of 8,000–10,000 K (refs. 6,11,12,13,14). In particular, ref. 6 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 ionization states of atoms and molecules from those predicted by the Boltzmann and Saha equations, respectively. They predicted an upper-atmospheric temperature about 2,000 K hotter than that 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. Furthermore, it showed that, as a result of the high atmospheric temperature and of NLTE effects, the triplet of neutral oxygen at $\sim$777.4 nm should be almost as strong as the previously detected H$\alpha$ Balmer line.

We analysed public data downloaded from the Calar Alto archive, which were taken with the CARMENES spectrograph during three transits of KELT-9b. CARMENES is a fibre-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 O i triplet (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 normalized for this reference. All the fully in-transit spectra (that is, 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 artefacts, such as the Rossiter–McLaughlin (RM) effect or centre-to-limb variations (CLVs), can contaminate the transmission spectrum, causing line-profile deformation and false detections. 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 two-dimensional (2D) tomographic technique after averaging the O i triplet in velocity space (Fig. 2a) and by computing the $K_\delta - V_w$ map, where $K_\delta$ is the RV semi-amplitude of the planetary Keplerian motion and $V_w$ the systemic velocity, leading to a detection at a signal-to-noise ratio (S/N) 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$^{-1}$. The line contrast corresponds to an average effective planetary radius of $\sim 1.17 R_p$, with $R_p$ the white-light planetary radius. For a tidally locked rotating atmosphere, such as that expected for KELT-9b, an absorption feature at $\sim 1.17 R_p$ should have an FWHM of $\sim 13$ km s$^{-1}$. 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. Here, our model accounts for the spectral resolution of the instrument and planetary rotation (assuming tidal locking). In line with the data, the model...
The constraint on the microturbulence velocity is tighter than that on the macroturbulence velocity, because \( \nu_{\text{mac}} \) 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 \( \nu_{\text{mac}} \) that we obtained agrees well with that used by ref. 11 to fit the Ca ii 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. 1), further broadening it accounting for \( \nu_{\text{mac}} \) and \( \nu_{\text{mic}} \), values of \( 10 \pm 2\) km s\(^{-1}\) and \( 0 \pm 7\) km s\(^{-1}\), respectively, which we estimated as done for the NLTE spectrum. The LTE profile is significantly weaker and a worse fit compared with the spectrum computed in NLTE, as shown by the respective \( \chi^2 \) values of 275 (NLTE) and 297 (LTE), with a preference for the NLTE model at the 4.3 \( \sigma \) level (Methods). Unfortunately, the current modelling set-up of ref. 1 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 O i triplet is a 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 O i triplet can be used to study NLTE effects in the atmospheres of ultrahot 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 (see, for example, ref. 23). Radial expansion at a velocity of 5–6 km s\(^{-1}\) due to atmospheric escape in the line-formation region (5 \( \times \) 10\(^{-9}\)–5 \( \times \) 10\(^{-7}\) bar) could explain the observed broadening, but is

---

**Fig. 1 | Transmission spectrum of KELT-9b around the O i triplet.** The magenta line shows the NLTE model of ref. 8 broadened with the best-fit \( \nu_{\text{mic}} \) and \( \nu_{\text{mac}} \), while the blue line shows the same for the LTE model (also from ref. 8). Grey lines show individual measurements with related uncertainties; black points show a 0.1 Å binning, with error-bars calculated as the standard deviation in each bin. The vertical dotted lines mark the expected position of the O i triplet lines.

---

**Fig. 2 | Planetary reference frame confirmation.** 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. Crosses mark the position of the \( K_p \) determined in refs. 14,15. The green dot marks the best \( K_p \) position in the map. The solid line shows the 3\( \sigma \) confidence interval.
unlikely. The effective Jeans escape parameter reduces to zero at the Roche lobe at a pressure of about $10^{-10}$ bar. To a good approximation, the atmosphere escapes through the Roche lobe around the $L_1$ and $L_2$ points at sound speed. Because gas escaping at other latitudes around the planet is also directed towards the $L_1$ and $L_2$ 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$^{24,25}$. On the basis of this assumption, our model TP profile implies a mass-loss rate of $10^{-9}$–$10^{-8}$ kg s$^{-1}$, which translates to a radial velocity of the order of 100 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 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 ultrahot Jupiters predict wind speeds of several kilometres per second, even for planets that are substantially cooler than KELT-9b$^{26}$. Such wind speeds are in good agreement with the observations.

The $\text{O}\,\text{i}$ 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$^{27,28}$. Furthermore, the $\text{O}\,\text{i}$ triplet at ≈777.4 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 detection of oxygen from the ground also for exoplanets that orbit 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.

**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 on the nights of 2017 August 6, 2017 September 21 and 2018 June 16 with CARMENES$^9$. This is a fibre-fed high-resolution echelle spectrograph installed on the 3.5 m telescope at Calar Alto observatory. Its optical arm covers the $\approx 520$–960 nm wavelength range with an average resolving power of about 94,600. The data of each night (37, 56 and 140 spectra, respectively) cover the transit of the planet. The exposure times used were different, varying from 111 s to 400 s. S/N in order 39 of the spectrograph, where the oxygen triplet lies, ranges from 30 to 120.

We analysed the reduced archival data, which were extracted with the instrument pipeline version 2.01 (ref. 17). The pipeline, after correcting for bias, flat field and cosmic rays, performs a flat-relative optimal extraction$^{18}$ and wavelength calibration$^{18}$. The output is produced in order-by-order maps, and consists of flux, flux uncertainties and wavelength. Before performing our analysis, for each night we discarded spectra taken at an airmass higher than 2 and with 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 order 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 ($V\sin\theta \approx 110$ km s$^{-1}$), and thus line blending, the stellar $\text{O}\,\text{i}$ triplet appears as one broad feature (Supplementary Fig. 1). Throughout, we considered the system parameters given in ref. 16.

We extracted the transmission spectrum by following independently for each night a procedure based on that presented in ref. 17. 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 normalized 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 ($247.28$ km s$^{-1}$, ref. 18). The final rebinning is done on a wavelength step of 0.02 Å, 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 in-transit residual spectra.

**Telluric correction.** We performed telluric correction using a scaling relation between airmass and telluric line strength$^{17,19}$, rescaling all the normalized stellar spectra as if they were obtained at the airmass of the centre 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 Virginis, 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 microtelluric could be present in the wavelength range we analysed by employing the ESO Sky Model Calculator (https://www.eso.org/observing/etc/bini/gen/form/INS/MODE=swspectr+INS/NAM=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 $V_{\text{mac}}$ and barycentric Earth radial velocity for the different nights and convolving them with the instrumental broadening. Water microtellurics could partly overlap with the reddest line of the $\text{O}\,\text{i}$ triplet; we can thus not completely

---

**Fig. 3 | $X^2$ minimization map of the NLTE synthetic spectrum.** Lines show the contours of the 1σ, 2σ and 3σ confidence intervals for the values of $V_{\text{mic}} = 3.0 \pm 0.7$ km s$^{-1}$ and $V_{\text{mac}} = 13 \pm 5$ km s$^{-1}$. The colour bar shows the $X^2$ values.

**Fig. 4 | Atmospheric sound speed and Jeans escape parameter.** Sound speed (black) and Jeans escape parameter (red) as a function of pressure extracted from the NLTE atmospheric structure profile presented in ref. 8. The shaded area indicates the main formation region of spectral lines lying in the optical region.
excluding that this line could be partially contaminated. We however confirm that the first two lines of the triplet are instead free also from water microtellurics (Supplementary Fig. 2).

Given the position of these water microtellurics in the transmission spectrum (about 777.6–777.7 nm), we can argue that their possible variation could be the cause of the imperfect normalization on the right-hand side of the triplet. As a test to see the possible impact on our analysis, we recomputed the X0 map of Fig. 3 by comparing the planetary atmospheric models and the observed transmission spectrum only blueward of 7.755 Å, thus excluding the third line of the O1 triplet, finding that our final results remain unchanged, confirming the $k_{\text{mic}}$ and $e_{\text{mic}}$ values we obtained.

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

Stellar contamination in the transmission spectrum. The host star is not a homogeneous disk, but has a surface brightness that rotates and changes as a function of the distance from the centre. Effects such as CLV and stellar rotation (the RM effect) may lead to spurious signals in the transmission spectrum, possibly due to the exposure time, caused by the fact that the planet moves during each exposure, has a negligible impact on the overall broadening (that is, ~4.7, 3.5 and 1.3 km s$^{-1}$ on average for exposure times of 400, 300 and 111 s, respectively). We compared the obtained further transmission spectra models in the same way as described in ref. 8 applying different $v_{\text{mic}}$ values ranging between 1 and 14 km s$^{-1}$ (in steps of 1 km s$^{-1}$), further adding $e_{\text{mic}}$ broadening with $e_{\text{mic}} = 25$ km s$^{-1}$ (in steps of 1 km s$^{-1}$), finally looking for the $e_{\text{mic}} = 0$ pair leading to matching of the observations by using X$^2$ minimization. We obtained best-fitting values of $e_{\text{mic}} = 3.0 \pm 0.7$ km s$^{-1}$ and $e_{\text{mic}} = 13 \pm 5$ km s$^{-1}$ (Fig. 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. 4 to estimate the $e_{\text{mic}}$ and $e_{\text{mic}}$ best fitting the observations, obtaining $10 \pm 2$ km s$^{-1}$ and $0 \pm 7$ km s$^{-1}$, respectively (Supplementary Fig. 6). However, the LTE transmission spectrum minimizing the X$^2$ value is a significantly worse fit to the observation compared with the NLTE synthetic spectrum. We compared the obtained X$^2$ values of 275 (NLTE model) and 297 (LTE model) using a likelihood ratio test. With two degrees of freedom ($e_{\text{mic}}$ and $e_{\text{mic}}$), we find a P value of 1.67 × 10$^{-5}$, which corresponds to a 4.3σ preference for the NLTE model, assuming that the uncertainties have been correctly estimated.

Mass-loss rate. To aid the interpretation of these results, we employed the atmospheric model presented in ref. 4 to estimate the mass-loss rate. Figure 4 shows the effective Jeans escape parameter and sound speed in the atmosphere on the basis of 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

$$X = \frac{m}{\kappa T}$$

where $\Delta p$ is the gravitational potential difference between the pressure level and the Roche lobe and $m$ is the mass of an escaping particle. At the Roche lobe, $X = 0$ and we can estimate the mass-loss rate as

$$\frac{dM}{dt} \approx 4 \pi R_0^2 \rho \frac{C}{2 \sqrt{\pi}}$$

where $R_0$ is the radius at the L1 point, $\rho$ the density, $Y = 2R_0/3$ is the polar radius and $C$ is the thermal speed. The L1 point in our model is located at 2.6 $R_0$ with a pressure of 1.17 × 10$^{-10}$ bar, where the thermal speed is 14 km s$^{-1}$ and mass density is 1.17 × 10$^{-9}$ kg m$^{-3}$. The resulting value for the mass-loss rate is 5 × 10$^{-9}$ kg s$^{-1}$. We expect this estimate to be accurate roughly to an order of magnitude$^{10-12}$.

The crossover mass concept$^{13}$ can be used to derive the minimum mass-loss rate of hydrogen required to enable atomic oxygen and iron to escape the atmosphere$^{14}$:

$$\frac{dM}{dt} \approx 4 \pi m_{H} G M_{L} (M - 1) n d \frac{d}{\kappa T}$$

where $m_{H}$ is the mass of hydrogen, $M_{L}$ is the mass of the heavy element in units of $m_{H}$ and $n d$ is the product of the total number density and the mutual diffusion coefficient. Using values of $n d = 1.7 \times 10^{-10}$ m$^{-3}$ s$^{-1}$ for O–H$^+$ collisions and $n d = 1.4 \times 10^{-10}$ m$^{-3}$ s$^{-1}$ for O–H collisions gives a limiting mass-loss rate of 2–3 × 10$^{-9}$ kg s$^{-1}$ for neutral oxygen. The value obtained for iron is about 10$^{-7}$ 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 important, once the outflow velocity reaches a substantial fraction of the sound speed. On the basis of a mass-loss rate of
5 × 10^9 kg s^{-1} and atmospheric structure predicted by our model, this would happen at 10^{10}–10^{11} bar, given that Ω_{\text{SP}} is constant with radius r. This location is close to our upper boundary, and above the region probed by the oxygen line profile. Detailed models of the core 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.

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

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/idestj/EXOFASTv2.

Received: 17 June 2021; Accepted: 21 October 2021; Published online: 22 December 2021

References
1. Madhusudhan, N. Exoplanetary atmospheres: key insights, challenges, and prospects. Annu. Rev. Astron. Astrophys. 57, 617–663 (2019).
2. Mendillo, M., Withers, P. & Dalba, P. Atomic oxygen ions as ionospheric biomarkers on exoplanets. Nat. Astron. 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. Astron. Astrophys. 275, 269–282 (1993).
4. Przybilla, N., Becker, S. R., Kudritzki, R. P. & Venn, K. A. Non-LTE line formation for neutral oxygen. Model atom and first results on A-type stars. Astron. Astrophys. 359, 1085–1106 (2000).
5. Steffen, M. The photospheric solar oxygen project. IV. 3D-NLTE investigation of the 777 nm triplet lines. Astron. Astrophys. 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. Local thermodynamic equilibrium effects determine the upper atmospheric temperature structure of the ultrahot Jupiter KELT-9b. Astron. Astrophys. 653, A52 (2021).
9. Yan, E. & Henning, T. An extended hydrogen envelope of the extremely hot giant exoplanet KELT-9b. Nat. Astron. 2, 714–718 (2018).
10. Cauley, P. W. et al. Atmospheric dynamics and the variable transit of KELT-9 b. Astrophys. J. 157, 69 (2019).
11. Turner, J. D. et al. Detection of ionized calcium in the atmosphere of the ultrahot Jupiter KELT-9b. Astrophys. J. 888, L13 (2020).
12. Wytenbach, A. et al. Mass-loss rate and local thermodynamic state of the KELT-9 b thermosphere from the hydrogen Balmer series. Astron. Astrophys. 638, A87 (2020).
13. Hoeijmakers, H. J. et al. Atomic iron and titanium in the atmosphere of the ultrahot Jupiter KELT-9b. Nature 560, 453–455 (2018).
14. Hoeijmakers, H. J. et al. A spectral survey of an ultrahot Jupiter. Detection of the centre-to-limb variation and Rossiter–McLaughlin effect in transmission spectroscopy studies. Astron. Astrophys. 635, A206 (2020).
Competing interests
The authors declare no competing interests.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41550-021-01544-4.

Correspondence and requests for materials should be addressed to Francesco Borsa.
Peer review information Nature Astronomy thanks Matteo Brogi and Fei Yan for their contribution to the peer review of this work.
Reprints and permissions information is available at www.nature.com/reprints.
Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
© The Author(s), under exclusive licence to Springer Nature Limited 2022