Non-local thermodynamic equilibrium effects determine the upper atmospheric temperature structure of the ultra-hot Jupiter KELT-9b

L. Fossati1, M. E. Young2,1, D. Shulyak3, T. Koskinen4, C. Huang4, P. E. Cubillos1, K. France5,6, and A. G. Sreejith1

1 Space Research Institute, Austrian Academy of Sciences, Schmiedlstrasse 6, 8042 Graz, Austria
e-mail: luca.fossati@oeaw.ac.at
2 Department of Physics, University of Oxford, Denys Wilkinson Building, Keble Road, Oxford, OX1 3RH, UK
3 Max-Planck Institut für Sonnensystemforschung, Justus-von-Liebig-Weg 3, 37077 Göttingen, Germany
4 Lunar and Planetary Laboratory, University of Arizona, 1629 East University Boulevard, Tucson, AZ 85721-0092, USA
5 Laboratory for Atmospheric and Space Physics, University of Colorado, 600 UCB, Boulder, CO 80309, USA
6 Center for Astrophysics and Space Astronomy, University of Colorado, 389 UCB, Boulder, CO 80309, USA

Received 15 March 2021 / Accepted 7 June 2021

ABSTRACT

Context. Several observational and theoretical results indicate that the atmospheric temperature of the ultra-hot Jupiter KELT-9b in the main line formation region is a few thousand degrees higher than predicted by self-consistent models.

Aims. Our aim was to test whether non-local thermodynamic equilibrium (NLTE) effects are responsible for the presumably higher temperature.

Methods. We employed the Cloudy NLTE radiative transfer code to self-consistently compute the upper atmospheric temperature-pressure (TP) profile of KELT-9b, assuming solar metallicity and accounting for Roche potential. In the lower atmosphere, we used an updated version of the HELIOS radiative-convective equilibrium code to constrain the Cloudy model.

Results. The Cloudy NLTE TP profile is ≈2000 K hotter than that obtained with previous models assuming LTE. In particular, in the 1–10−7 bar range the temperature increases from ≈4000 to ≈8500 K, remaining roughly constant at lower pressures. We find that the high temperature in the upper atmosphere of KELT-9b is driven principally by NLTE effects modifying the Fe and Mg level populations, which strongly influence the atmospheric thermal balance. We employed Cloudy to compute LTE and NLTE synthetic transmission spectra on the basis of the TP profiles computed in LTE and NLTE, respectively, finding that the NLTE model generally produces stronger absorption lines, particularly in the ultraviolet, than the LTE model (up to 30%). We compared the NLTE synthetic transmission spectrum with the observed Hα and Hβ line profiles obtaining an excellent match, thus supporting our results.

Conclusions. The NLTE synthetic transmission spectrum can be used to guide future observations aiming at detecting features in the KELT-9b transmission spectrum. Metals, such as Mg and Fe, and NLTE effects shape the upper atmospheric temperature structure of KELT-9b, and thus affect the mass-loss rates derived from it. Finally, our results call for checking whether this is the case also for cooler planets.

Key words. radiative transfer – planets and satellites: atmospheres – planets and satellites: gaseous planets – planets and satellites: individual: KELT-9b

1. Introduction

Ultra-hot Jupiters, which are planets whose continuum is dominated by H− and where the optical and infrared spectral ranges are dominated by metal absorption and lack of molecular absorption (Arcangeli et al. 2018; Parmentier et al. 2018; Kitzmann et al. 2018; Lothringer et al. 2018), are becoming prime targets for atmospheric characterization. This is mostly due to their high atmospheric temperature, which ensures large pressure scale heights and hence large (detectable) atmospheric signals in transmission and emission, and to an absence of aerosols, at least on the day-side. Furthermore, these planets typically orbit rather bright stars, favouring atmospheric characterisation observations.

About half a dozen ultra-hot Jupiters have been detected and observed to date, mostly in transmission and employing high-resolution spectroscopy from the ground, though some (space-based) low-resolution spectroscopic observations have also been obtained and analysed. However, the interest of the community in ultra-hot Jupiters has significantly increased following the detection of KELT-9b, also known as HD 195689 b, which is the hottest exoplanet orbiting a non-degenerate star known to date (Gaudi et al. 2017). Both TESS and Spitzer phase curve observations have been employed to measure the day- and night-side planetary temperatures, obtaining values of roughly 4600 and 3040 K, respectively (Mansfield et al. 2020; Wong et al. 2020).

KELT-9b is one of the most targeted planets for ground-based high-resolution transmission spectroscopy observations. Both hydrogen Balmer lines and metal features have been detected in the planetary transmission spectrum (e.g. Yan & Henning 2018; Hoeijmakers et al. 2018, 2019; Yan et al. 2019; Borsa et al. 2019; Cauley et al. 2019; Turner et al. 2020; Pino et al. 2020; Wytenbach et al. 2020). Similar results have also been obtained for a few other (not so extreme) ultra-hot Jupiters, such as WASP-33b (e.g. von Essen et al. 2019; Yan et al. 2019, 2021; Cauley et al. 2021; Nugroho et al. 2020a), WASP-121b (e.g. Sing et al. 2019; Gibson et al. 2020; Ben-Yami et al. 2020;
Bourrier et al. 2020; Cabot et al. 2020; Hoeijmakers et al. 2020b), WASP-76b (e.g. Seidel et al. 2019; Ehrenreich et al. 2020; Tabernero et al. 2020; Edwards et al. 2020; Fu et al. 2021), WASP-189b (e.g. Yan et al. 2020), and MASCARA-2b (e.g. Casasayas-Barris et al. 2018, 2019; Hoeijmakers et al. 2020a; Nugroho et al. 2020b; Stangret et al. 2020).

In the case of KELT-9b the observations led to constraining the planetary atmospheric temperature-pressure (TP) profile and the composition from the detection of metal features. Lothringer et al. (2018) employed the PHOENIX stellar and planetary atmosphere code to perform forward modelling of the TP profile of synthetic ultra-hot Jupiters of different temperatures assuming local thermodynamic equilibrium (LTE). They concluded that the upper atmospheric layers of these planets are characterised by a temperature inversion. According to the PHOENIX calculations, for KELT-9b this temperature inversion should lead to upper atmospheric temperatures of the order of about 6500 K (see also Lothringer & Barman 2020). Pino et al. (2020) detected FeI emission from secondary eclipse ground-based high-resolution observations and arrived at the conclusion that the planetary atmosphere is indeed characterised by an inverted TP profile, confirming the modelling predictions. However, Pino et al. (2020) were unable to quantify the temperature inversion. García Muñoz & Schneider (2019) modelled the planetary upper atmosphere assuming a pure hydrogen composition and accounting for non-LTE (NLTE) effects. They argued that the upper atmospheric temperature should be of the order of 15 000 K, but also mentioned that the inclusion of metals in the model would likely decrease it by 1000–2000 K.

The detection of the hydrogen Balmer lines provides valuable constraints on the TP profile and abundances in the upper atmosphere (García Muñoz & Schneider 2019; Turner et al. 2020; Wytenbach et al. 2020; Fossati et al. 2020). Furthermore, this information allows us to constrain the energetics driving atmospheric escape (Yan & Henning 2018; Fossati et al. 2018; Wytenbach et al. 2020) even though the Balmer lines do not probe the atmosphere beyond the planetary Roche lobe (Turner et al. 2020). Accounting for NLTE effects, both García Muñoz & Schneider (2019) and Turner et al. (2020) showed that fitting the observed hydrogen Balmer lines requires an upper atmospheric temperature significantly hotter than 6500 K, which is the value predicted by PHOENIX LTE forward models (Lothringer et al. 2018; Fossati et al. 2018; Lothringer & Barman 2020). Wytenbach et al. (2020) employed a retrieval technique to constrain the atmospheric TP profile, impact of NLTE effects, and mass-loss rate. They obtained an atmospheric temperature of 9600 ± 1200 K, a density of excited hydrogen (i.e. in the n = 2 state) relative to the total amount of hydrogen of about 10⁻¹¹, and a mass-loss rate of the order of 10⁻¹⁴ g s⁻¹. However, they assumed an isothermal TP profile, which Fossati et al. (2020) showed to be an extreme and likely unphysical assumption. Furthermore, they considered the ratio of LTE to NLTE hydrogen densities as a free parameter, which led to significant degeneracies among the output parameters.

In an attempt to overcome the assumptions of Wytenbach et al. (2020), Fossati et al. (2020) constructed a large sample of artificial TP profiles characterised by a temperature inversion of varying shape and location; they were looking for those leading to best fit the observed hydrogen Balmer lines. They accounted for NLTE effects in the spectral synthesis calculations employing the Cloudy code (Ferland et al. 2017). They found that the family of TP profiles best fitting the observations is characterised by an inverted temperature profile starting at pressures higher than 10⁻⁴ bar, and that the upper atmospheric temperature could be as high as 10 000 K. Their results also showed that the LTE assumption leads to overestimating the strength of the hydrogen Balmer lines. Driven by the significant metal ionisation in the upper atmosphere extracted from the Cloudy simulations, Fossati et al. (2020) suggested that metal photoionisation might be responsible for driving the properties of the TP profile, and in particular for the ≈3000–4000 K difference between the TP profile best fitting the hydrogen Balmer lines and what was predicted by previous models.

Cloudy is a code designed to simulate physical conditions within gas clouds ranging from the intergalactic medium to the high-density LTE limit (Ferland et al. 1998, 2013, 2017). Within the exoplanet field Cloudy has been employed to compute the physical conditions of planetary escaping gas interacting with the stellar radiation and wind (Turner et al. 2016), to obtain heating and cooling rates to be used in exoplanetary hydrodynamic simulations (Salz et al. 2015, 2016, 2018, 2019), and to compute transmission spectra for a wide range of planets (Salz et al. 2018, 2019; Turner et al. 2020; Cubillos et al. 2020; Fossati et al. 2020; Young et al. 2020a,b). In this work we extend some of previous Cloudy exoplanet modelling work, and in particular build on the results of Fossati et al. (2020) by employing forward modelling both in LTE and NLTE to construct TP profiles of KELT-9b with the aim of exploring in detail the origin of the upper atmospheric heating. We also employ the LTE and NLTE TP profiles to generate transmission spectra and explore the impact of the LTE assumption on the interpretation of the observations. We finally compare the NLTE transmission spectrum with the observed hydrogen Balmer lines, but leave the comparison with the observed metal lines to future dedicated works.

This paper is organised as follows. In Sect. 2 we describe the adopted modelling scheme and tools. In Sect. 3 we present the results obtained from forward NLTE modelling of the atmospheric structure of KELT-9b, while in Sect. 4 we discuss the results and compare them with the observed Hα and Hβ line profiles. We draw the conclusions of this work in Sect. 5.

2. Atmospheric modelling

We generated synthetic TP profiles for the ultra-hot Jupiter KELT-9b combining the results obtained employing the HELIOS code (Malik et al. 2017, 2019) for the lower atmosphere (i.e. P ≥ 10⁻³ bar) and the Cloudy code (Ferland et al. 2017) for the upper atmosphere (i.e. P ≤ 10⁻⁴ bar). The motivation for combining the results of two distinct codes comes from the fact that HELIOS uses opacity tables computed under LTE assumption, which cannot be changed in the current version of the code, and NLTE effects are important in the upper atmosphere (e.g. Fossati et al. 2020, see below). Instead, Cloudy can account for NLTE effects, but it is unreliable for total gas densities higher than roughly 10¹⁵ cm⁻³, hence in the lower atmosphere. This density limit is driven by the approximations involved in the treatment of three-body recombination–collisional ionisation for the heavy elements (see Sect. 3 of Ferland et al. 2017). However, tests run with Cloudy indicate that this most strongly affects calculations in which Cloudy is allowed to compute the temperature structure at high gas densities, while those with a fixed temperature profile appear to be significantly less affected (Young et al. 2020a; Fossati et al. 2020). This occurs because the approximations mostly impact the heating and cooling functions that are involved in calculating the temperature structure, and not the spectra.
Here we describe in detail the modelling conducted with HELIOS and Cloudy. For all calculations, we considered the system parameters given by Borsa et al. (2019) and listed in Fossati et al. (2020, Table 2). The stellar spectral energy distribution we considered as input to compute the TP profiles is that presented by Fossati et al. (2018), effective temperature of 10000 K.

2.1. Lower atmosphere: $P \geq 10^{-4}$ bar

To calculate the structure of the dense layers of the planetary atmosphere, we employed the HELIOS\textsuperscript{1} radiative-convective equilibrium code (Malik et al. 2017, 2019). The inputs for the HELIOS code are planetary mass, radius, and atmospheric abundances; orbital semi-major axis; and stellar radius and effective temperature. Because HELIOS does not include photochemistry, we assumed equilibrium abundances, which we obtained with the FASTCHEM\textsuperscript{2} code (Stock et al. 2018). Chemical equilibrium is a reasonable approximation because the concentrations of many major species in deep atmospheric layers of very hot planets are close to their chemical equilibrium values (see e.g. Shulyak et al. 2020). Our simulation setup includes 100 atmospheric layers distributed logarithmically between 100 and $10^{-3}$ bar. The equilibrium temperature of the planet is controlled by the heat redistribution parameter $f$, which accounts for the (presumably unknown) redistribution efficiency between the day and night sides of the planet. By adjusting this parameter we can fine tune the desired day-side temperature of the planet. We set $f$ equal to 0.63, the value for which the temperature at pressures higher than $10^{-4}$ bar best reproduces the TP profile computed with the PHOENIX code (Fossati et al. 2018; Lothringer & Barman 2020). Furthermore, with this $f$ value, at pressures of around $10^{-1}$–$10^{-2}$ bar the temperature lies between 4500 and 5000 K, in agreement with the measured day-side temperature of about 4600 K (Gaudi et al. 2017; Mansfield et al. 2020).

Figure 1 presents the HELIOS TP profile, in comparison to the PHOENIX TP profile.

The hot temperature of the KELT-9b atmosphere requires several additional opacity sources to be added compared to that present in the public version of the HELIOS code. First, we included atomic line opacity due to neutral and singly ionised atoms, namely C\textsc{i}-\textsc{ii}, C\textsc{r}-\textsc{ii}, Fe\textsc{i}-\textsc{ii}, K\textsc{i}-\textsc{ii}, Mg\textsc{ii}-\textsc{ii}, Na\textsc{i}-\textsc{ii}, O\textsc{ii}-\textsc{ii}, and Si\textsc{i}-\textsc{ii}. These are the elements that were found to contribute most to the line opacity throughout the planetary atmosphere. The original line lists are those produced by R. Kurucz\textsuperscript{3} (Kurucz 2018). For each atom we generated cross sections for different sets of temperatures, pressures, and wavelengths using the HELIOS-K\textsuperscript{4} package (Grimm & Heng 2015; Grimm et al. 2021). The molecular line opacity includes such molecules as CH\textsubscript{2}, CO\textsubscript{2}, CO, H\textsubscript{2}O, HCN, NH\textsubscript{3}, OH, SiO, TiO, and VO. The pretabulated cross sections of these molecules were taken from the public opacity database for exoplanetary atmospheres\textsuperscript{5}. The generated cross sections were then weighted with the number density of the corresponding species computed under the LTE assumption and using the K\textsc{table} code distributed together with the HELIOS installation.

Next we extended HELIOS with new continuum opacity sources. These include bound-free and free-free transitions of $^3$He\textsuperscript{-}, free-free transitions of He\textsuperscript{+}, Rayleigh scattering on H and H\textsuperscript{+} atoms, and Thomson scattering on free electrons. Additionally, we included bound-free opacity for the following atoms: H\textsc{i}, C\textsc{i}-\textsc{ii}, C\textsc{r}-\textsc{ii}, O\textsc{i}-\textsc{ii}, Mg\textsc{ii}-\textsc{ii}, Al\textsc{i}, Si\textsc{i}-\textsc{ii}, and Fe\textsc{ii}. Finally, we included the continuum opacity of the OH and CH molecules. All relevant numerical routines were taken from the LLM models stellar model atmosphere code (see Shulyak et al. 2004, and references therein). The other continuum opacities include Rayleigh scattering by H\textsubscript{2}O, H\textsubscript{2}, and CO\textsubscript{2} molecules, and collisionally induced absorption due to H\textsubscript{2}O, H\textsubscript{2}, and H\textsubscript{2}-He following Borysow et al. (2001), Borysow (2002), and Borysow & Fromhold (1989). This is the same set of tools that Fossati et al. (2020) employed to compute synthetic hydrogen Balmer line profiles assuming LTE.

2.2. Upper atmosphere: $P \leq 10^{-4}$ bar

2.2.1. Cloudy

We modelled the planetary upper atmosphere employing Cloudy\textsuperscript{6} (version 17.02), which is a microphysics spectral synthesis code designed to simulate physical conditions within an astrophysical plasma, predicting the emitted and absorbed spectrum, further accounting for (photo)chemistry and NLTE effects (Felten et al. 1998, 2013, 2017). Cloudy accounts for NLTE effects by explicitly computing level populations and ionisation without resourcing to the Boltzmann and Saha equations, and can consider non-Planck radiation field and non-Maxwellian velocity distribution. We note that the gas in the planetary atmosphere considered here is Maxwellian, because it is not subject to extremely intense non-thermal radiation due to cosmic rays, for example (e.g. Felten et al. 2016). Cloudy computations include a wide range of atomic (i.e. all elements up to Zn) and molecular species, and are valid across

\textsuperscript{1} https://github.com/exoclime/HELIOS
\textsuperscript{2} https://github.com/exoclime/FastChem
\textsuperscript{3} http://kurucz.harvard.edu/linelists/gfall/
\textsuperscript{4} https://github.com/exoclime/HELIOS-K
\textsuperscript{5} https://dace.unige.ch/opacity
\textsuperscript{6} https://gitlab.nublado.org/cloudy/cloudy/-/wikis/home
a wide interval of plasma temperatures ($3 \leq 10^5$ K) and densities ($< 10^{15}$ cm$^{-3}$), covering the parameter space of what is expected in upper planetary atmospheres. Within Cloudy the temperature of the gas corresponds to the electron temperature.

To enable computing radiative transfer accounting for NLTE effects, Cloudy does not employ pre-computed opacity tables, because in NLTE the opacity is itself a function of the radiation field and therefore the opacities must be recomputed every time the atmospheric temperature changes from one iteration to another. Instead, Cloudy computes the opacities line-by-line; specifically, at each given frequency point the code extracts the line opacity from all available lines and adds the continuum contribution to it to obtain the total opacity. The opacity is then computed in this way at each frequency point in the considered wavelength range and for a large number of frequencies ($> 10^5$). The frequency spacing is nearly exponentially increasing with increasing wavelength (see Cloudy User Manual 2 for more details). The spectral resolution for the Cloudy computations is given by the user and the minimum is 33. In this work, we considered a resolution of 100 000.

The code predicts the thermal, ionisation, and chemical structure of a gas cloud, as well as its emitted and transmitted spectrum. Cloudy predicts these quantities from specifying just (a) the shape and intensity of the external radiation field illuminating a cloud (i.e. in our case the stellar spectral energy distribution), (b) the chemical composition of the gas (we assume solar composition), and (c) the geometry of the gas, including its radial extent and the dependence of the gas density as a function of the distance to the radiation source. Cloudy does this by simultaneously solving the equations of statistical and thermal equilibrium, the equations that balance ionisation-neutralisation processes, and the heating-cooling processes (Osterbrock & Ferland 2006). Cloudy treats all ionisation stages and includes as recombination mechanisms charge-exchange, radiative recombination, and dielectronic recombination processes. The ionisation mechanisms accounted for in the code comprise photoionisation from valence, inner shells, and excited states, as well as collisional ionisation by both thermal and supra-thermal electrons (Voronov 1997; Dere 2007), and charge transfer. Cloudy also accounts for molecular photodissociation.

Cloudy calculations are one-dimensional and assume plane-parallel geometry and hydrostatic equilibrium. For the specific case of KELT-9b, the latter assumption is valid within the main line formation region of optical and infrared spectral lines. This is because deviations from hydrostatic equilibrium become important once the outflow velocity reaches a significant fraction of the sound speed (i.e. near the sonic point), which lies close to the top boundary of the range of pressures considered here ($10^{-11} \leq P \leq 10^{-12}$ bar; see below), and thus well above the line formation region of optical and infrared spectral lines (e.g. Koskinen et al. 2013b; Turner et al. 2020; Fossati et al. 2020). Cloudy calculations are iterative and the iterations are controlled by convergence criteria on the local pressure, on the temperature (i.e. heating-cooling balance), and on the electron density, for each considered layer.

Cloudy builds up its calculations on a large database comprising model atoms (e.g. energy levels, level state definition, statistical weight of each energy level, Einstein coefficients, line oscillator strengths, collision rates) for all elements up to Zn. The information included in the model atoms is taken from Lykins et al. (2015, largely based on NIST® for the atomic part), the CHIANTI database (Dere et al. 1997; Landi et al. 2012), and the Leiden Atomic and Molecular Database (LAMDA; Schöier et al. 2005). When computing spectral lines Cloudy accounts for natural, thermal, and Stark broadening. Therefore, broadening through collisions with neutrals (i.e. van der Waals broadening) is not included in the code. Furthermore, Cloudy can also account for line broadening by turbulence, which modifies line opacities and the resulting optical depths, and adds a component of ram pressure to the total pressure. However, in this work, we do not include any turbulence in the calculations.

2.2.2. Cloudy for Exoplanets

Cloudy is a general-purpose code adaptable to a great variety of gas clouds; the user has to first set up the geometry of the cloud and its basic physical conditions, namely the hydrogen density profile as a function of the radial distance to the external light source and the basic chemical composition (i.e. which elements to consider in addition to hydrogen, their abundance, and which molecules to take into account). To this end, we developed a PYTHON interface, called Cloudy for Exoplanets (CIE), that writes Cloudy input files on the basis of input parameters given by the user, runs Cloudy, and reads Cloudy output files, using the information contained there to set up a new Cloudy calculation in an iterative procedure until the temperature profile has converged as described below. The details of how CIE sets up Cloudy input files and the iteration procedure are described below.

In the simplest case a Cloudy simulation requires as input the total hydrogen density and physical extent of the atmosphere being simulated, and a source of illumination. In addition, to compute the temperature and chemical structure of an exoplanetary atmosphere, CIE requires the planetary and stellar masses ($M_p$ and $M_\ast$) and radii ($R_p$ and $R_\ast$), the orbital semi-major axis (a), the orbital period ($\tau$), an approximate planetary equilibrium temperature ($T_\text{eq}$), the optical continuum pressure level ($p_0$), and the bulk atmospheric metallicity ($[M/H]$). With this information, CIE creates an initial sub-stellar one-dimensional atmospheric structure by assuming an isothermal atmosphere of $T = T_\text{eq}$, scaled solar abundance ratios, and no ionisation or excitation, and then computes the radius scale over a given pressure range assuming hydrostatic equilibrium as

$$r_{i+1} = -H_i \ln \left( \frac{P_{i+1}}{P_i} \right) + r_i, \quad (1)$$

where $r_i$ is the radius from the planet’s centre of the $i$th level, $H_i$ is the pressure scale height of the $i$th level, and $P_i$ is the atmospheric pressure of the $i$th level. The pressure scale height is calculated as

$$H_i = \frac{kT_{i+\frac{3}{2}}}{m_{i+\frac{3}{2}}g_i}, \quad (2)$$

where $k$ is Boltzmann constant, $T_{i+\frac{3}{2}}$ is the average atmospheric temperature between the $i$th and $i$th+1 levels, $m_{i+\frac{3}{2}}$ is the average mean molecular weight of the atmosphere between the $i$th and $i$th+1 levels, and $g_i$ is the gravitational acceleration experienced by atmospheric particles at the $i$th level.

To solve Eq. (1), it is necessary to adopt an appropriate reference $p_0$-$R_0$ pair to set the value of the planetary radius corresponding to the optical continuum pressure level ($p_0$). There is a well-known pressure-radius reference location degeneracy in this choice (e.g. Lecavelier Des Etangs et al. 2008;
Benneke & Seager 2012; de Wit & Seager 2013; Griffith 2014; Heng & Kitzmann 2017; Bétrémieux & Swain 2017). We note that Heng & Kitzmann (2017) extends this degeneracy into a three-way degeneracy by including the water abundance. However, Welbanks & Madhusudhan (2019) later showed that the degeneracy between planetary radius and its reference pressure is well characterised irrespective of the water abundance, and has little effect on abundance estimates. Therefore, the \(p_0-R_p\) choice, assuming it is reasonably close to reality, does not have a significant impact on the results (see below).

The gravitational acceleration accounts for the effects of the Roche potential, formulated as

\[
g_{\text{rot}} = -\frac{GM_p}{r_{\text{i}}^2} = \frac{GM_\ast (r_{\text{i}} \lambda^2 - a \lambda + r_i \xi^2 + r_i \nu^2)}{[a - r_i \nu] \lambda^2 + (r_i \xi)^2 + (r_i \nu)^2]}^{\frac{3}{2}} + \omega^2 (r_{\text{i}} \lambda^2 - \mu a \lambda + r_i \xi^2),
\]

where \(\nu = \frac{\omega^2 \mu \xi}{\nu p}\), \(\omega = \frac{\mu \lambda}{r_{\text{i}}^2}\), and \(\lambda = \cos (\phi) \sin (\theta)\), \(\xi = \sin (\phi) \sin (\theta)\). \(\nu = \cos (\phi) \cos (\theta)\), \(\nu_p = -\sin (\phi) \sin (\theta)\), for co-latitude \(0 \leq \theta \leq \pi\) and longitude \(0 \leq \phi \leq 2\pi\). This simplifies to

\[
g_{\text{rot}} = -\frac{GM_\ast}{r_{\text{i}}^2} + \frac{GM_{\ast}}{(a - r_i \nu)^2} + \omega^2 \left(r_{\text{i}} - a - \frac{M_{\ast}}{M_\ast + M_p}\right)
\]

for \(\theta = \pi/2\) and \(\phi = 0\) at the substellar point.

This initial radius scale and density profile are passed to Cloudy, which computes the thermal and chemical equilibrium solution of the atmosphere, and the resulting output is used to generate a new atmospheric structure following the above method, but without assuming an isothermal and neutral atmosphere. In practice, CIE updates the hydrogen density structure of the planetary atmosphere given as input to Cloudy employing the temperature and mean molecular weight as a function of pressure obtained from the previous Cloudy run. The process is repeated for a number of iterations set by the user, after which the converged temperature structure, chemical solution, and radius scale of the atmosphere are recorded, as is the radius scale at the terminator, computed from the final structure according to the above method using polar latitude and longitude coordinates. Experience indicates that the temperature structure stabilises following two CIE iterations, but we run up to four CIE iterations reaching a <5 K difference between the TP profiles computed in the last two CIE runs.

The Cloudy atmospheric model indicates that in the case of KELT-9b the maximum gas density above which Cloudy results become unreliable lies around the 1 mbar level. Therefore, for the CIE calculations, we set \(p_0\) equal to 0.175 bar, the pressure level at which the HELIOS model gives an optical depth of 2/3 integrating over the KELT bandpass filter, whose barycentre lies at \(\approx 6000\ \text{Å}\) (Pepper et al. 2007). Furthermore, to solve Eq. (1) and to be consistent with the estimate of \(p_0\), we adopted \(r(p_0)\) equal to the observed transit radius \(R_p = 1.936 R_{\text{Jup}}\) given by Borsa et al. (2019), which is based on the transit depth measurement obtained from the KELT light curves (Gaudí et al. 2017). We tested the impact of this choice on the obtained TP profile by recomputing the Cloudy TP profile employing smaller radii of 1.6, 1.65, 1.7, 1.75, 1.8, and 1.85 \(R_{\text{Jup}}\) instead of 1.936 \(R_{\text{Jup}}\), always obtaining an identical result (Fig. A.1).

For all Cloudy calculations we took into account all elements up to Zn and only hydrogen molecules (i.e. H\(_2\), H\(_3^+\), and H\(_3^+\)) because other molecules play a minor role in the hot environment of the KELT-9b atmosphere in determining the atmospheric physical properties (e.g. Lothringer & Barman 2019).

Cloudy computations do not allow us to consider atmospheric heat redistribution; therefore, we mimicked \(f\) by adding a scaling factor to the input stellar spectral energy distribution that we varied until the Cloudy and HELIOS TP profiles matched at around the \(10^{-4}\)-bar level. This is close to the highest pressure at which Cloudy calculations are still valid and NLTE effects are small (Fossati et al. 2020), making Cloudy and HELIOS results comparable. In this part of the atmosphere the two TP profiles follow the same shape. We computed TP profiles employing a range of scaling factors, and obtained that the one leading to the best match of the HELIOS profile around the \(10^{-4}\)-bar level is 0.48. We also compared the TP profiles computed with different values of the scaling factor obtaining that the TP profile at pressures below \(10^{-5}\) bar, thus in the main line formation region, is only weakly dependent on it, and increases by about 200 K with the scaling factor going from 0.4 to 1. Therefore, we conclude that the general results presented here can be considered robust against the choice of the scaling factor. The adopted TP profile obtained with CIE, and accounting for NLTE effects, is shown in Fig. 1.

### 3. Results

We defined the final TP profile over the \(10^{-10} - 1\) bar pressure range, dividing it into 180 layers equally spaced in log \(p\). This pressure range is wide enough to fully contain the atmospheric formation region of ultraviolet to infrared spectral lines. It further ensures that the atmosphere is transparent to light at ultraviolet to infrared wavelengths at the top of the atmosphere and opaque in the same wavelength range at the bottom of the atmosphere. We ran models considering a smaller and larger number of atmospheric layers, but the results are identical to those obtained considering 180 layers (see Fig. B.1).

We joined the HELIOS and Cloudy TP profiles in the \(1.4 \times 10^{-4}\) and \(6.0 \times 10^{-5}\) bar pressure range, which is where the shape of the two profiles looks most similar. In this pressure range, we derived the composite TP profile by averaging between the HELIOS and Cloudy results, further interpolating on the final pressure scale. At pressures higher than \(1.4 \times 10^{-4}\) bar and lower than \(6.0 \times 10^{-5}\) bar, we set the composite TP profile to correspond to those obtained with HELIOS and Cloudy, respectively, also interpolating them on the final pressure scale. Figure 1 shows the composite TP profile, in comparison to the HELIOS, Cloudy, and PHOENIX TP profiles.

By construction, at high pressures the composite TP profile is in agreement with the PHOENIX TP profile and with the measured day-side temperature. The profile further presents a rather steep temperature increase between about 1 and \(10^{-3}\) bar.
in which the temperature goes from ≈4200 to ≈8400 K, and it remains roughly constant at lower pressures. In the upper atmosphere, the temperature is about 1800–2500 K higher than predicted by HELIOS and PHOENIX. The shape of the composite TP profile goes in the direction of the empirical TP profile (Fossati et al. 2020) found to best match the observed hydrogen Balmer lines. In particular, in the lower atmosphere the TP profile empirically obtained by Fossati et al. (2020) is about 1000–2000 K hotter than obtained here, but that profile was unconstrained by the observations outside of the main line formation region, specifically at pressures higher and lower than roughly $10^{-3}$ and $10^{-8}$ bar, respectively (Fig. 1). The two profiles concur that a rather steep temperature rise starts at the 10 mbar pressure level. The main difference lies at the top of the upper atmosphere, where the TP profile presented here is about 2000 K cooler than the empirical TP profile of Fossati et al. (2020) and is roughly isothermal at pressures lower than $10^{-6}$ bar, instead of constantly increasing with decreasing pressure. Within the main line formation region the temperature changes by almost 3000 K, which indicates that the assumption of an isothermal atmosphere, for example taken by Wytenbach et al. (2020) and Yan et al. (2021), is most likely unphysical and should be avoided.

To also obtain a homogeneous chemical atmospheric structure, the composite TP profile is passed to CIE to generate input for a single iteration while fixing the temperature structure in the Cloudy simulation. The top panel of Fig. 2 shows the details of the atmospheric composition for hydrogen-bearing species (neutral hydrogen H, protons H+, molecular hydrogen H$_2$, H$^+$, H$_2^+$, electrons e$^-$) with respect to the total hydrogen density as a function of pressure. The atmosphere is dominated by neutral hydrogen at pressures higher than about $10^{-7}$ bar, while protons are the most abundant species at lower pressures. In the lower atmosphere, at pressures higher than about $10^{-3}$ bar, H and H$_2$ dominate the hydrogen atmospheric composition, with the latter decreasing rapidly with decreasing pressure. The H$^+$ hydrogen species is the fourth most abundant in the lower atmosphere and dominates the spectral shape of the continuum (e.g. Arcangeli et al. 2018). The H$_2$ abundance increases quickly with decreasing pressure starting from the $10^{-5}$ bar level. This is mostly due to thermal ionisation, because hydrogen photoionisation occurs at higher altitudes, and it is believed to be small due to the shape of the stellar spectral energy distribution that is characterised by a very weak emission of extreme ultraviolet radiation (Fossati et al. 2018).

The middle and bottom panels of Fig. 2 show instead the mixing ratio as a function of pressure for some of the species most relevant in terms of abundance and observability at ultraviolet and/or optical wavelengths. As a consequence of the solar composition assumption, He is the second most abundant species throughout most of the atmosphere. Carbon and oxygen behave in a similar way; they remain mostly in neutral form up to the $10^{-5}$ and $10^{-7}$ bar level, respectively, being then mostly singly ionised at lower pressures. This result suggests that the several C (neutral and ionised) and O (mostly neutral) resonance lines lying in the far-ultraviolet might be detectable in transmission, for example with HST. Because of their similar configurations, Na and K behave alike and are singly ionised throughout most of the atmosphere.

As a result of their similar ionisation potentials, Mg, Si, and Fe behave alike, with the neutral species dominating at pressures higher than about 5 mbar and the singly ionised species dominating at lower pressures, up to the $10^{-5}$ bar level where the doubly ionised species take over. Single spectral lines of these species, particularly Mg and Fe, have been directly detected in transmission spectra of KELT-9b (e.g. Cauley et al. 2019). The near-ultraviolet wavelength range contains several strong Mg, Si, and Fe lines that could be detectable with both HST and CUTE (Fleming et al. 2018). The distribution of Ca atoms presents a clear separation, with CaI dominating at pressures higher than about 0.1 bar, CaII dominating up to the $10^{-5}$ bar level, and CaIII dominating at lower pressures. However, the CaI density decreases rather slowly with decreasing pressure above the $10^{-5}$ bar level, namely in the main
line formation region, which probably explains why the Ca I infrared triplet has been detected in transmission (Turner et al. 2020).

The density profiles shown in Fig. 2 are remarkably similar to those obtained by Fossati et al. (2020) and extracted with Cloudy on the basis of one of the empirical TP profiles they found to fit the H α and H β lines better. The main difference lies in the slightly higher pressure levels at which ionised species become more abundant than neutral species. This is driven by the cooler temperature, obtained here through forward modelling, compared to what Fossati et al. (2020) obtained though a grid approach.

4. Discussion

4.1. Origin of the upper atmospheric heating

4.1.1. NLTE vs. LTE temperature-pressure profile

Fossati et al. (2020) found that the family of TP profiles best fitting the hydrogen Balmer lines is characterised by an inverted temperature profile with an upper atmospheric temperature of the order of 10 000 K. The high ionisation fraction of metals in the upper atmosphere found on the basis of the TP profiles best fitting the hydrogen Balmer lines led them to propose metal photoionisation as the upper atmospheric heating mechanism. However, thanks to the implementations described in Sect. 2.1, the HELIOS TP profile has been computed accounting for metal photoionisation. Therefore, the origin of the extra heating lies somewhere else.

One of the aspects separating HELIOS from Cloudy is the assumption of LTE opacity in the HELIOS model calculations. We note that the PHOENIX TP profile was also computed assuming LTE. Therefore, under the assumption of radiative equilibrium, NLTE effects may be a viable cause for the high upper atmospheric temperature. In particular, the overpopulation and/or underpopulation of specific energy levels of species playing a major role in the atmospheric opacity, and thus heating and cooling, may significantly affect the shape of the TP profile. Therefore, we made use of Cloudy, which has the capability of running computations in both LTE and NLTE, to identify the impact of NLTE effects on the TP profile. We note that the LTE TP profile still accounts for photoionisation.

Figure 3 shows the TP profiles computed with Cloudy in LTE and NLTE, also in comparison with those obtained with PHOENIX and HELIOS assuming LTE. At the top of the atmosphere, the LTE TP profile ends at somewhat higher pressure compared to the NLTE TP profile because Cloudy experienced convergence problems at very low gas densities and assuming LTE. In general, the shape of the LTE Cloudy TP profile resembles that of the PHOENIX and HELIOS TP profiles, but it is slightly cooler and does not show the rather narrow temperature bump around 10^-8 bar that characterises the HELIOS TP profile.

There is a significant difference between the LTE and NLTE Cloudy TP profiles. At pressures higher than about 10^-4 bar the two profiles are similar, but diverge significantly at lower pressures. In the upper atmosphere, thus in the main formation region of major spectral lines (Turner et al. 2020; Fossati et al. 2020), the TP profile computed accounting for NLTE effects is more than 1000 K hotter than for LTE, with the largest difference reaching more than 2000 K at pressures ranging between about 5 x 10^-8 and 10^-10 bar.

4.1.2. Atmospheric heating and cooling processes

We explored the Cloudy models to identify the species and energy levels responsible for the extra heating and/or lack of cooling driving the difference between the LTE and NLTE TP profiles. Figures 4 and 5 show respectively the heating and cooling contribution provided by the three most important heating and cooling processes in the atmosphere as a function of pressure extracted from the Cloudy output obtained from the NLTE and LTE runs. In both NLTE and LTE cases, metal line absorption is the largest heating contribution at pressures lower than ~10^-4 bar, except for the very top of the atmosphere in the NLTE case where the photoionisation of hydrogenic species (i.e. photionisation of excited H I, in this case) is the major heating contribution, though metal line absorption is still significant. Absorption from H II becomes progressively more important as pressure increases, but also at these deeper layers metal line absorption heating is significant, and is still the main heating contributor in the LTE case.

Although Cloudy does not give information on the total heating rate due to each species, we inferred this for Fe in NLTE by taking the output obtained after the last CIE iteration and then running a further iteration considering only Fe I (i.e. Fe I is not allowed to ionise), or considering only Fe II (i.e. Fe II is not allowed to ionise or recombine), or removing Fe completely from the list of considered atmospheric species. For the cases in which we fixed Fe corresponding to Fe I/Fe II, we employed the Fe I/Fe II density profile obtained from the model that accounts for all elements. Figure 6 shows the temperature and the total heating rate as a function of pressure obtained in these four cases. By removing Fe the heating rate decreases significantly in the 10^-5–10^-10 bar pressure range, which is precisely the portion of atmosphere that most differs between the LTE and NLTE case. Figure 6 clearly indicates that most of the heating in this part of the atmosphere is caused by Fe II.

The heating rate around the nbar level and above is comparable to what is typically obtained from hydrodynamic modelling of classical hot Jupiters (e.g. Yelle 2004; Koskinen et al. 2013a). However, at higher pressures, particularly when including Fe, the heating rate is much higher. This is due to the metals included in Cloudy that are mostly missing from current hydrodynamic upper atmosphere models. If Fe reaches the upper atmosphere, this result demonstrates that accounting for Fe and NLTE effects becomes critically important for correctly modelling the atmospheric energy balance. This is particularly relevant for planets
Heating contribution (in %) of the three most important heating processes at each atmospheric layer as a function of pressure extracted from the NLTE (top) and LTE (bottom) Cloudy runs. The relevant heating processes occurring in the upper atmosphere are hydrogen photoionisation (red; photoionisation of H$^+$ lying in the ground state), metal line absorption (blue), H$^+$ absorption (green), Compton heating (i.e. electron absorption; yellow), photoionisation of hydrogenic species (photoionisation of excited H; magenta), and H$_2^+$ absorption (brown). Metal absorption is the main heating contributor at pressures ranging between about $5 \times 10^{-5}$ and $10^{-10}$ bar, where the atmospheric temperature increases steeply and reaches its maximum. The heating fraction of any given process is artificially set equal to zero in the atmospheric regions in which the heating process is not among the three most important ones.

Fig. 5. Cooling contribution (in %) of the three most important cooling processes or species at each atmospheric layer as a function of pressure obtained from the Cloudy NLTE (top) and LTE (bottom) runs. The label “FF” indicates free-free cooling from H and He. The cooling fraction of any given species is artificially set equal to zero in the atmospheric regions where the species producing the cooling is not one of the three most important. Magnesium is the species most contributing to the cooling of the upper atmosphere.

Excluding Fe from the Cloudy calculation led to an almost 2000 K cooler TP profile at pressures ranging between $10^{-5}$ and $5 \times 10^{-10}$ bar, which Fig. 4 shows being heated predominantly by metal line absorption. This indicates that there must be numerous Fe lines, likely rising from overpopulated energy levels, contributing to the heating in this part of the atmosphere. Even in the absence of Fe the temperature rises steeply at the top of the atmosphere, which is due to the increasing importance of H$_2^+$ photoionisation heating with decreasing pressure (Fig. 4, top left panel).

4.1.3. Role of Mg and Fe in the TP profile

To quantify the role of Fe and Mg in heating and cooling the planetary atmosphere, we computed TP profiles with CIE in NLTE excluding either Fe or Mg from the list of considered elements. This is different from what we show in Fig. 6; in that case the calculation was based on just one Cloudy iteration for which the starting point was the TP profile computed accounting for all elements, while here it is the full CIE run that does not consider Fe among the list of elements. This is to ensure the maximum possible consistency when comparing results. Figure 7 compares the Cloudy TP profiles obtained considering all elements up to Zn with those computed excluding Fe or Mg.

Excluding Fe from the Cloudy calculation led to an almost 2000 K cooler TP profile at pressures ranging between $10^{-5}$ and $5 \times 10^{-10}$ bar, which Fig. 4 shows being heated predominantly by metal line absorption. This indicates that there must be numerous Fe lines, likely rising from overpopulated energy levels, contributing to the heating in this part of the atmosphere. Even in the absence of Fe the temperature rises steeply at the top of the atmosphere, which is due to the increasing importance of H$_2^+$ photoionisation heating with decreasing pressure (Fig. 4, top left panel).

Fig. 6. Temperature (left) and total heating rate (right) as a function of pressure obtained accounting for all elements (black; this is the same TP profile shown in Figs. 1 and 3), considering Fe only in the form of FeI (blue; i.e. Fe is not allowed to ionise), considering Fe only in the form of FeII (green; i.e. Fe is not allowed to recombine or ionise), and after removing Fe among the list of considered elements (red).
Excluding Mg from the computation of the TP profile has the opposite effect of excluding Fe, though to a smaller extent. The largest difference between the TP profiles computed with and without Mg is found in the 10^{-7}–10^{-10} bar pressure range, with the maximum difference of about 500 K being at a pressure of about 5 \times 10^{-8} bar. As expected, this is the location at which the top panel of Fig. 5 indicates that Mg is the most important cooling species.

We performed similar calculations excluding each of the other elements one at a time obtaining differences that are significantly smaller than those obtained by removing Mg or Fe, and in most cases the differences are negligible. Furthermore, the Mg and Fe abundance profiles extracted from the LTE and NLTE models are similar. Therefore, the overpopulation of certain levels of Fe and/or the underpopulation of certain levels of Mg is likely to be responsible for the large difference between the NLTE and LTE TP profiles.

This is confirmed by the departure coefficients (b) we extracted from the Cloudy NLTE run for Fe, FeI, FeII, MgI, and MgII. The departure coefficients are defined as

$$b = \frac{n_{\text{NLTE}}}{n_{\text{LTE}}}$$

where $n_{\text{NLTE}}$ and $n_{\text{LTE}}$ are the densities of a given atom lying in a certain level in NLTE and LTE, respectively. The $n_{\text{LTE}}$ profiles are those obtained through the Boltzmann equation.

Figures C.1–C.5 show the departure coefficients we obtained for FeI, FeII, FeIII, and MgI, for the first 80 energy levels, and for MgII for the first 20 energy levels (those included in the 17.02 Cloudy distribution). Throughout the upper atmosphere, the considered Fe energy levels on average are underpopulated, with the population increasing with pressure. We note that the Fe underpopulation, particularly of the upper energy levels, might partially be an artefact of an incomplete Fe model atom, although for each element we considered all energy levels available in the 17.02 Cloudy distribution. Mashonkina et al. (2011) showed that for late-type stars a more complete Fe model atom facilitates recombination from FeII, decreasing the underpopulation of the upper levels of Fe.

The first ~60 energy levels of FeI are systematically underpopulated by a factor of about ten at pressures lower than ~10^{-7} bar. Higher FeI energy levels are instead slightly underpopulated. Figure C.2 also shows that at pressures higher than about 10^{-5} bar FeI can be treated in LTE. The FeII departure coefficients behave similarly to those of FeI, with an average moderate underpopulation and with the level population increasing with pressure. At pressures higher than ~10^{-8} bar, FeII can be treated in LTE. Figures C.4 and C.5 show that the considered MgI and MgII energy levels are significantly underpopulated.

Therefore, the overpopulation of the lower levels of FeI and the underpopulation of the lower levels of MgI and MgII is the likely cause of the significant difference between the LTE and NLTE TP profiles, and thus also of the difference between the Cloudy NLTE TP profile and the HELIOS and PHOENIX TP profiles that have been computed assuming LTE. The importance of FeI overpopulation in the atmospheric heating increases when considering that FeII is the dominant Fe species at pressures between 10^{-7} and 10^{-10} bar (Fig. 2), thus in the pressure range in which the temperature increases most significantly and reaches its maximum. Furthermore, the lines rising from the lower, overpopulated energy levels lie in the (near-)ultraviolet, that is the wavelength band in which the stellar radiation is most intense.

### 4.2. Importance of NLTE effects

We further explored the importance of accounting for NLTE effects by computing the planetary transmission spectra in both LTE and NLTE. To compute the transmission spectrum in LTE, we first constructed a composite TP profile by joining the Cloudy LTE TP profile in the upper atmosphere with the HELIOS TP profile in the lower atmosphere, as described in Sect. 2. In the case of the LTE Cloudy TP profile, we found that the factor scaling the stellar flux and leading to the best match of the HELIOS TP profile around the 10^{-4} bar pressure level is 0.67, which is very close to the HELIOS (i.e. LTE) f value (Section 2.1). To improve the accuracy in the calculation of the transmission spectra, particularly in setting the continuum level (see below), we resampled the composite LTE and NLTE TP profiles in the 10^{-7}–10^{-13} bar pressure range and divided it into 50 layers equally spaced in log $p$.

We computed the transmission spectra employing Cloudy and the algorithm described in Young et al. (2020a) and Fossati et al. (2020). To obtain the radius profile of the atmosphere (Eq. (1)), we determined the reference pressure–radius reference point ($p_0$ and $R_0$) by fitting iteratively for the reference pressure $p_0$ constrained by the observed transit radius considering the KELT filter bandpass, whose barycentre lies at about 6000 Å ($R_p = 1.936 R_{\text{Jup}} = R_0$). It is necessary to perform the $p_0$–$R_0$ calibration again because of the different geometry involved in computing transmission spectra, compared to what we considered for computing the TP profile described in Sect. 2.2 (i.e. from emission geometry to transmission geometry). Following the procedure described in Fossati et al. (2020), we determined $p_0$ in an iterative procedure by integrating the transmission model spectrum over the KELT bandpass filter, adjusting $p_0$ until the band-integrated model planetary radius matches the observed transit radius of $R_p = 1.936 R_{\text{Jup}}$. In particular, at each iteration we determined $p_0$ looking for the pressure at which the optical depth integrated over the KELT bandpass filter reaches 0.56 (Lecavelier Des Etangs et al. 2008). In this way, following two iterations, we obtained that $p_0$ lies at a pressure of 0.059 and 0.028 bar for the NLTE and LTE cases, respectively. We obtained two different values of $p_0$ in LTE and NLTE (because the continuum optical depth depends on the entire photosphere that is being probed), which have different physical characteristics in LTE and NLTE. Furthermore, the two $p_0$ values were obtained by computing the radiative transfer in different ways (i.e. assuming LTE or accounting for NLTE effects).
We note that we assumed that the reference radius $R_0$ corresponds to the transit radius when computing both the TP profile at the substellar point and the transmission spectrum. However, this assumption has a negligible impact on the results. In contrast, when computing the transmission spectrum, we make a further, more coarse assumption, namely that the TP profile computed at the substellar point is valid across the entire planet (i.e. both day and night sides; see Fossati et al. 2020, for a discussion about this assumption for KELT-9b). As described in Sect. 4.3, it is possible that this assumption has a non-negligible impact on the line shapes in the transmission spectra.

Figure 8 presents the LTE and NLTE transmission spectra in the 1100–11 000 Å range, the NLTE correction, and the transit depth difference as a function of wavelength. Similar plots, but zooming into shorter wavelength ranges for better visibility can be found in Appendix D. Throughout the considered wavelength range the LTE assumption leads mostly to underestimated absorption-line strengths, while the opposite occurs for a much smaller number of features. NLTE effects, with deviations from LTE consistently above 15%. Interestingly, the NLTE transmission spectrum indicates that both Fe II and a strong Al II resonance line. This strong blend presents a temperature difference between the underlying LTE and NLTE TP profiles. The majority of the spectral lines at short wavelengths come from ionised species, which are more abundant in the NLTE model as a consequence of the generally higher atmospheric temperature of the NLTE TP profile compared to the LTE TP profile, particularly in the line forming region. Furthermore, the difference in the LTE and NLTE TP profiles leads to different scale heights, which also affect the features in the transmission spectrum. Figure 9 demonstrates this by showing the comparison between the transmission spectra computed in LTE and NLTE on the basis of the NLTE TP profile. Figures 8 and 9 show, for example, for the Hα line that the NLTE correction has moved from about 10% to zero. This is because Cloudy is not set up to do LTE calculations involving the hydrogen $n = 2$ level (i.e. Lyman and Balmer lines), and therefore the NLTE correction for these lines shown in Fig. 8 is exclusively due to the difference between the LTE and NLTE TP profiles.

In the far-ultraviolet (Fig. D.1, top), the Si III resonance line at 1206 Å is the one showing the strongest deviation from LTE (≈27%). Several other far-ultraviolet features (e.g. Lyα, C II 1335 Å, Si II 1530 Å, C IV 1545 Å) present deviations from LTE above 15%. Accounting for NLTE effects, we predict that several features in the planetary transmission spectrum will reach the 2–5% absorption level, thus presenting transit light curves 2 to 6 times deeper than the continuum level, which would be most likely detectable with HST. In particular, even though silicon is a species that is roughly as abundant as iron (assuming solar composition), it has not yet been detected in the transmission spectrum of KELT-9b. The strong features at ≈1530 Å belong to Si III resonance lines and would be ideal targets for detecting Si in the planetary atmosphere.

The mid-ultraviolet spectral region (Fig. D.1, bottom) is characterised by a forest of rather strong metal lines for which the NLTE correction is on average about 10%. However, there is a strong feature at ≈1670 Å that is a blend composed of two strong Fe II lines rising from low energy levels (i.e. 0.232 and 0.352 eV) and a strong Al II resonance line. This strong blend presents a NLTE correction larger than 15%.

In the near-ultraviolet (Fig. D.2) the strongest features are the Fe II absorption bands in the 2400–2800 Å wavelength range and the Mg II h&k resonance lines at ≈2800 Å. These are strong features in both the LTE and NLTE transmission spectra, but they appear to be significantly enhanced when accounting for NLTE effects, with deviations from LTE consistently above 15%. Interestingly, the NLTE transmission spectrum indicates that both Fe I...
and Fe i near-ultraviolet features are going to be particularly strong. This gives the opportunity to employ near-ultraviolet transmission spectroscopy to detect both ions, and thus observationally constrain the atmospheric Fe ionisation fraction, which would give additional precious constraints to the shape of the TP profile. The NLTE correction for the Mg i h&k resonance lines reaches 30%. This large correction is due to the difference in the temperature structure between the LTE and NLTE TP profiles. Because Mg i ionises in the lower atmosphere, between the $10^{-2}$ and $10^{-3}$ bar level, the Mg i resonance line at 2853 Å is not a prominent feature in the planetary transmission spectrum, similarly to what is observed for the ultra-hot Jupiter WASP-121b (Sing et al. 2019), and reaches a transit depth of about 1.5%.

The transmission spectrum in the blue part of the optical wavelength range (Fig. D.2, bottom; Fig. D.3, top) is characterised by several metal lines that only reach about the 1–1.5% absorption depth. There are, however, two features that peak significantly over the others reaching 2% transit depths. These are absorption depth. There are, however, two features that peak significantly over the others reaching 2% transit depths. These are the line is stronger in LTE than in NLTE). The most prominent of the order of 1.2%. This part of the transmission spectrum also shows the presence of a large number of lines belonging mostly to Fe-peak elements, which have been detected both directly (e.g. Cauley et al. 2019) and through cross-correlation (e.g. Hoeijmakers et al. 2018, 2019).

In the red part of the optical wavelength range (Fig. D.3, bottom), the Hα line is the strongest feature in the transmission spectrum, reaching a depth of almost 1.5%. Fossati et al. (2020) found that the LTE assumption would lead to overestimating the strength of the hydrogen Balmer lines, while Fig. 8 shows the opposite. This occurs because the LTE transmission spectrum presented here is based on the LTE TP profile, which is significantly cooler than the NLTE TP profile, thus leading to a lower concentration of excited hydrogen atoms, and thus to smaller Balmer features. Even though Na i is ionised in the lower atmosphere, there is enough Na i in the line forming region to produce quite prominent Na i D features that reach an absorption depth in transmission of about 1%. Interestingly, the feature rising to above the 1% level absorption depth at about 7780 Å is the O i near-infrared triplet that could be detectable in the already obtained high-resolution transmission observations.

Figure 8, and in particular the more detailed plots in Appendix D, show that the optical and near-infrared wavelength ranges host a number of lines with negative NLTE correction (i.e. the line is stronger in LTE than in NLTE). The most prominent of these lines with negative NLTE corrections ≥2.5% lie between 2200 and 2600 Å and at ≈4300, 4450, and 7900 Å. The lines with negative NLTE corrections in the near-ultraviolet wavelength range belong primarily to features of Fe-peak elements (mostly neutral) rising from underpopulated energy levels. The lines at ≈4300 Å and at wavelengths shorter than 4460 Å are all strong Ca i features rising from levels with an energy of about 1.88 eV. The features at ≈4480 and 7900 Å correspond to the Mg i triplets at 4481 and 7890 Å.

We leave the detailed analysis of the NLTE transmission spectrum and its comparison to the available observations, particularly of metal lines, to a future work. However, we check the quality of the LTE TP profile and transmission spectrum by comparing the synthetic profiles of the Hα and Hβ lines with the observations. This comparison is the focus of the next section.

Figure 8 and the above discussion demonstrate that NLTE effects play a fundamental role in shaping not only the planetary atmospheric TP profile, but also the transmission spectrum. The large difference between the predicted LTE and NLTE transmission spectra for several spectral features gives the further opportunity to observationally test whether the atmospheric temperature is indeed as high as predicted by our NLTE calculations, and thus to confirm or disprove the impact of NLTE effects in driving the atmospheric properties of ultra-hot Jupiters.

4.3. Comparison with observations

We compare here the synthetic LTE and NLTE transmission spectrum of the Hα and Hβ lines with the available observations. In particular, we consider the observations presented by Yan & Henning (2018, Hα), Cauley et al. (2019, Hα and Hβ), Turner et al. (2020, Hα), and Wyttenbach et al. (2020, Hα and Hβ). We refer to Sect. 2 of Fossati et al. (2020) for a thorough discussion of the similarities and differences, also in terms of data analysis, present among these observations.

Figure 10 shows the comparison between the observed and synthetic NLTE and LTE Hα and Hβ line profiles. The LTE transmission spectrum was computed employing the LTE TP profile (Fig. 3) and the procedure described in Sect. 6.3 of Fossati et al. (2020) because, by construction, Cloudy is not set to compute the $n = 1$ and $n = 2$ hydrogen level populations in LTE. For context, Fig. E.1 shows the H i departure coefficients for the first ten energy levels, showing that the lower three energy levels (responsible for the Lyman, Balmer, and Paschen series) are heavily overpopulated in the upper atmosphere.

The NLTE synthetic spectra of both lines are in general a good match to the data, while the LTE synthetic spectra are significantly weaker than the observations, mostly due to the cooler TP profile compared to the NLTE case. In the following, we consider only the NLTE synthetic profiles.

The main difference between the synthetic spectra and the observations is the line shape, where the observed line profiles have a more triangular shape in comparison to the rounder shape of the synthetic lines. This difference might arise because we consider only the day-side temperature profile to compute the
transmission spectra (see also Sect. 6.7 of Fossati et al. 2020). The line profiles forming on the planetary night side would be narrower because of the lower temperature, and thus they would contribute mostly to the shape of the line core. We quantify the fit of the NLTE synthetic spectrum by computing the $\chi^2$ and reduced $\chi^2$ ($\chi^2_{\text{red}}$) employing the non-rebinned observed spectra. The values are listed in Table 1.

The worst match is that with the H$\alpha$ profile of Yan & Henning (2018) because the NLTE synthetic profile is significantly weaker and narrower than the observation. The H$\alpha$ profile of Yan & Henning (2018) is by far the strongest of the four (i.e. its equivalent width is about 10$\sigma$ larger than those of Cauley et al. 2019 and Wyttenbach et al. 2020). The rather low $\chi^2$ and $\chi^2_{\text{red}}$ values obtained from the comparison with the H$\alpha$ profile of Turner et al. (2020) are driven by the large observational uncertainties, though visual inspection suggests that the synthetic spectrum is a good match to the data.

The profiles of Cauley et al. (2019) and Wyttenbach et al. (2020) are the most comparable in terms of uncertainties and line strength (Fossati et al. 2020, Table 1), though the profiles of Cauley et al. (2019) are slightly deeper and narrower than those of Wyttenbach et al. (2020). The $\chi^2$ and $\chi^2_{\text{red}}$ values listed in Table 1 indicate that the NLTE transmission spectrum is a good match to these observations. Therefore, accounting for metals and NLTE effects in the computation of the TP profile and of the transmission spectrum produces a good fit to the observed hydrogen Balmer lines.

We note that Cloudy is a static code, meaning that it does not account for hydrodynamic motions of the gas, but the model still leads to a good fit to the observed hydrogen Balmer lines. This is because the core of the H$\alpha$ and H$\beta$ lines probe the nbar and 10 nbar pressure levels$^8$, respectively, which are well below the sonic point (located between the $10^{-11}$ and $10^{-12}$ bar), where hydrodynamic motions become significant. This indicates that it is not possible to use the fit of these lines to directly constrain hydrodynamic motions of the planetary atmosphere and thus mass-loss rates. However, the TP profile gives information regarding the energy available in the atmosphere to drive mass loss. This is going to be the subject of a future work.

The TP profile presented here and obtained through forward modelling appears to be a better fit to the data compared to what had been found through a grid approach by Fossati et al. (2020). The likely reason for this is that the algorithm used to compute the parametric TP profiles considered by Fossati et al. (2020) did not provide enough flexibility to lead to a better match to the data compared to what had been obtained. In particular, none of the models in the grid of Fossati et al. (2020) reached the maximum temperature at pressures as high as $\approx 10^{-6}$ bar, remaining then isothermal at lower pressures. Remarkably, this pressure level is at the centre of the H$\alpha$ and H$\beta$ line formation region. In other words, the TP profiles employed by Fossati et al. (2020) presenting a temperature at the top of the atmosphere of 8000–9000 K were too cool around the main line formation region to lead to enough excited hydrogen atoms for matching the data. The TP profile Fossati et al. (2020) found as the best match to the data reached a temperature of 8000–8500 K, the highest temperature in the forward model, around the $10^{-6}$ bar level (i.e. the centre of the H$\alpha$ and H$\beta$ line formation region).

Finally, KELT-9b has been extensively observed at optical and near-infrared wavelengths, but ultraviolet transmission spectra are still not available. Furthermore, the ultraviolet is the wavelength range in which the stellar emission is highest. Therefore, to guide future ultraviolet observations, we convolved the NLTE transmission spectrum to the spectral resolution of the HST STIS E140M grating (red). Bottom: NLTE transmission spectrum in the far-ultraviolet wavelength region (grey) convolved to the spectral resolution of the HST STIS E230M grating (red) and to that of the spectograph on board the CUTE Small-Sat mission (blue). The considered CUTE spectral resolution accounts for spacecraft jitter.

---

**Table 1.** $\chi^2$ and reduced $\chi^2$ values obtained from the comparison between the synthetic NLTE transmission spectrum of the H$\alpha$ and H$\beta$ line profiles and the observed profiles obtained by Yan & Henning (2018), Cauley et al. (2019), Turner et al. (2020), and Wyttenbach et al. (2020).

| Source                        | H$\alpha$ | H$\beta$ |
|-------------------------------|-----------|---------|
| Yan & Henning (2018)          | $\chi^2$  | 497.46  |
|                               | $\chi^2_{\text{red}}$ | 7.54    |
| Cauley et al. (2019)          | $\chi^2$  | 159.78  |
|                               | $\chi^2_{\text{red}}$ | 1.11    |
| Turner et al. (2020)          | $\chi^2$  | 43.34   |
|                               | $\chi^2_{\text{red}}$ | 0.66    |
| Wyttenbach et al. (2020)      | $\chi^2$  | 184.63  |
|                               | $\chi^2_{\text{red}}$ | 0.93    |

---

$^8$ We arrived at this conclusion thanks to having solved the pressure-radius reference location degeneracy by matching the reference continuum pressure to the optical transit radius (Brogi & Line 2019; Fisher et al. 2020).
spectra collected with both instruments. The duration of the KELT-9b transit is about 3.9 h (Gaudi et al. 2017); therefore, it is possible to perform in-transit observations along two HST orbits, for a total of about 6000 s. Considering this exposure time, two transit observations, and the E140M and E230M gratings, we obtained average S/N values per pixel of about 85 and 310, respectively. Instead, for CUTE the S/N calculator, based on the CUTE data simulator (Sreejith et al. 2019), indicates that the average S/N value per pixel obtained for KELT-9b following a five-minute exposure (the baseline for CUTE observations; Fleming et al. 2018) is about 40. Assuming an observing efficiency of 65% (i.e. accounting for Earth occultations, passes within the South Atlantic Anomaly, and readout time), we obtained that CUTE would be able to collect roughly 30 exposures of five minutes each, hence finally obtaining spectra with an average S/N per pixel, integrating over the whole transit, of about 220. The baseline for CUTE is to collect ten transits for each target, thus reaching a final S/N per pixel of about 700. Figures F.1 and F.2 show the transmission spectra obtained considering instrument spectral resolution, wavelength sampling, binning (ten pixels in the case of the HST STIS observations and two pixels in the case of CUTE observations), and the S/N values given above. In the far-ultraviolet spectral range, all major features corresponding to resonance lines of C, S, O, Si, and C are detectable (i.e. consecutive data points at wavelengths corresponding to these lines lie above the noise level). In the near-ultraviolet, the Mg triplet resonance lines and the strongest Fe features are detectable employing both HST STIS and CUTE.

5. Conclusion

We employed the Cloudy NLTE radiative transfer code to self-consistently compute the upper atmospheric TP profile of the prototype ultra-hot Jupiter KELT-9b, accounting for NLTE effects. We obtained a TP profile about 2000 K hotter than predicted by other models assuming LTE. In particular, the profile displays a steep temperature rise in the 1–10−7 bar range, with the temperature increasing from ~4000 to ~8500 K, remaining roughly constant at lower pressures. We ran an additional Cloudy model for the upper atmosphere of KELT-9b, but assuming LTE, obtaining a TP profile comparable to those computed in LTE with other codes. Therefore, the stark difference in the upper atmospheric TP profile between our Cloudy NLTE model and those presented in the literature is due to NLTE effects.

We examined in detail the output of the LTE and NLTE Cloudy models, which indicate that Fe and Mg are responsible, respectively, for most of the upper atmospheric heating and cooling. We further explored this finding by computing new TP structure models with Cloudy in NLTE excluding Fe or Mg at a time. Following the removal of Fe from the atmospheric metal content, we obtained a TP profile that is significantly cooler than the original, and comparable to that computed in LTE. Instead, the removal of Mg from the atmospheric composition led to an upper atmospheric TP profile that is about 500 K hotter than the original. The departure coefficients computed by Cloudy for both Fe and Mg indicate that in the upper atmosphere, most Fe, Mg, MgI, and MgII levels are underpopulated, while most FeII levels are overpopulated. The MgII underpopulation leads to a lack of cooling compared to the LTE case. Instead, the FeII overpopulation leads to an increased heating, particularly because the majority of the strong FeII lines, rising from overpopulated energy levels, lie in the near-ultraviolet wavelength range, which is the spectral band in which the host star’s spectral energy distribution peaks.

We further employed Cloudy to compute LTE and NLTE transmission spectra of KELT-9b covering from the far-ultraviolet to the mid-infrared. We found that the strongest deviations form LTE (up to 30%) occur at ultraviolet wavelengths, though there are also features at longer wavelengths for which the NLTE correction is larger than 10%. We compared the NLTE synthetic transmission spectrum with the observed Hα and Hβ lines obtaining an excellent match, and thus validating our results. This positive outcome hints at the fact that it could be possible to use transmission spectra computed accounting for NLTE effects to guide the search for spectral features in the already available and future transit observations of KELT-9b. To this end, we employed the NLTE synthetic transmission spectrum to compute synthetic ultraviolet observations obtained with HST STIS after two transits, and with CUTE after ten transits. The results indicate that the majority of the strong features, such as the CII, SiII, and CIIV resonance lines in the far-ultraviolet and the Fett bands and the MgII h&k resonance lines in the near-ultraviolet would be detectable with both instruments. These are also the features presenting the strongest deviations from LTE, which would therefore provide a strong element for validating our results.

The striking difference between the LTE and NLTE TP profiles presented in this work suggests that, at least for KELT-9b, the LTE assumption might induce misinterpreting both observational and theoretical results. Furthermore, the strong temperature rise of about 2500 K across the main line formation region invalidates the isothermal assumption that is often taken to simplify atmospheric modelling to enable retrieval analyses. Therefore, any results obtained on the atmosphere of KELT-9b and based on an isothermal profile should be taken with caution, particularly if they have been obtained through comparisons with high-resolution observations.

Finally, it will be extremely important to test in the near future whether the importance of NLTE effects in controlling the atmospheric TP profile is a peculiarity of KELT-9b or is a feature common for less irradiated planets as well. Radiative transfer considerations suggest that NLTE effects might be even more relevant for cooler planets because of the decreased importance of collisions. If this is the case, it will be critical to identify the parameter space within which NLTE significantly affects TP profiles. This work is needed also to adequately interpret the many observations of ultra-hot Jupiters that have been and are being obtained with several facilities, both from space and from the ground. Furthermore, most current models of giant planet upper atmospheres and atmospheric escape do not include the metals that control the energy balance on ultra-hot Jupiters, such as KELT-9b. There is an urgent need to include the relevant species in these models to properly calculate mass-loss rates and to interpret observations of metal lines lying at ultraviolet wavelengths that probe the uppermost layers of the atmosphere.

Acknowledgements. M.E.Y. acknowledges funding from the OAW-Innovationsfonds IF_2017_03. D.S. acknowledges the support of the DFG priority program SPP-1992 “Exploring the Diversity of ExtrasolarPlanets” (DFG PR 36 24 602/41). 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. A.G.S. and L.F. acknowledge financial support from the Austrian Forschungsförderungsgesellschaft FFG project CONTROL P865968.
Appendix A: Cloudy TP profiles computed considering different reference radii

Fig. A.1. Cloudy TP profiles computed considering different reference radii. The adopted TP profile is shown in black. For visualisation purposes, the TP profiles are rigidly shifted horizontally by the value indicated in the legend.

Appendix B: Cloudy TP profiles computed considering different numbers of atmospheric layers

Fig. B.1. Cloudy TP profiles computed considering different numbers of atmospheric layers. The finally adopted TP profile is shown in red. For visualisation purposes, the TP profiles are rigidly shifted horizontally by the value indicated in the legend.

Appendix C: Iron and magnesium departure coefficients

Fig. C.1. Departure coefficients as a function of atmospheric pressure for the first 80 energy levels of FeI. The energy levels are numbered from 1 to 80, and are separated into groups of ten levels and two line styles (solid and dashed) for each panel. Within each group of ten energy levels, the order of the line colors corresponding to increasing energy is black, red, blue, dark green, magenta, yellow, brown, grey, bright green, and orange. Therefore, the black solid line shows $b$ for the ground state (i.e. $n = 1$ level), the red solid line shows $b$ for $n = 2$ level, the blue solid line shows $b$ for the $n = 3$ level, and the black dashed line shows $b$ for the $n = 11$ level, and so on. The dotted line (at 1.0) indicating $n_{\text{LTE}} = n_{\text{LTE}}$ is for reference. The departure coefficients appear to be diverging from 1.0 with increasing pressure to reach 1.0 in the lower atmosphere, the latest around the 0.2 bar level, at which point all departure coefficients are equal to one.
Fig. C.2. Same as Fig. C.1, but for Fe II.

Fig. C.3. Same as Fig. C.1, but for Fe III.
Fig. C.4. Same as Fig. C.1, but for Mg I.

Fig. C.5. Same as Fig. C.1, but for Mg II and up to the first 20 energy levels, which are those included in the 17.02 Cloudy distribution.
Appendix D: Ultraviolet and optical transmission spectra of KELT-9b in NLTE and LTE

Fig. D.1. Same as Fig. 8, but for the 1100–1550 Å (top) and 1500–2350 Å (bottom) wavelength ranges.
Fig. D.2. Same as Fig. 8, but for the 2300–3050 Å (top) and 3000–4050 Å (bottom) wavelength ranges.
Fig. D.3. Same as Fig. 8, but for the 4000–6100 Å (top) and 6000–11 000 Å (bottom) wavelength ranges.
Appendix E: Hydrogen departure coefficients

Fig. E.1. Departure coefficients as a function of atmospheric pressure for the first ten energy levels of H\textsc{i}. The order of the line colors corresponding to increasing energy is black, red, blue, dark green, magenta, yellow, brown, grey, bright green, and orange. Therefore, the black solid line shows $b$ for the ground state (i.e. $n = 1$ level), the red solid line shows $b$ for the $n = 2$ level, the blue solid line shows $b$ for the $n = 3$ level, and so on.
Appendix F: Simulated transmission spectra obtained following two HST STIS and ten CUTE transit observations

**Fig. F.1.** Simulated STIS ultraviolet transmission spectra. *Top:* simulated transmission spectrum (red line) obtained following two far-ultraviolet HST STIS transit observations collected employing the E140M grating. The transmission spectrum accounts for instrument spectral resolution and wavelength sampling, obtained by binning together ten pixels. The black line shows the transmission spectrum without noise and binning, but accounting for instrument spectral resolution and wavelength sampling. *Bottom:* same as top, but following two transit observations employing the E230M grating. The black line is underneath the red one for better visibility.
Fig. F.2. Simulated transmission spectrum (red line) obtained following ten CUTE transit observations. The transmission spectrum accounts for instrument spectral resolution and wavelength sampling, and it has been obtained binning together two pixels. The black line shows the transmission spectrum without noise and binning, but accounting for instrument spectral resolution and wavelength sampling.