

Effects of a fully 3D atmospheric structure on exoplanet transmission spectra: retrieval biases due to day–night temperature gradients

A. Caldas¹, J. Leconte¹, F. Selsis¹, I. P. Waldmann², P. Bordé¹, M. Rocchetto², and B. Charnay³

¹ Laboratoire d’astrophysique de Bordeaux, Université de Bordeaux, CNRS, B18N, allée Geoffroy Saint-Hilaire, 33615 Pessac, France
e-mail: jeremy.leconte@u-bordeaux.fr

² Department of Physics and Astronomy, University College London (UCL), UK

³ LESIA, Observatoire de Paris, PSL Research University, CNRS, Sorbonne Université, Université Diderot Paris, Sorbonne Paris Cité, 92195 Meudon, France

Received 5 October 2018 / Accepted 23 January 2019

ABSTRACT

Transmission spectroscopy provides us with information on the atmospheric properties at the limb, which is often intuitively assumed to be a narrow annulus around the planet. Consequently, studies have focused on the effect of atmospheric horizontal heterogeneities along the limb. Here we demonstrate that the region probed in transmission – the limb – actually extends significantly towards the day and night sides of the planet. We show that the strong day–night thermal and compositional gradients expected on synchronous exoplanets create sufficient heterogeneities across the limb that result in important systematic effects on the spectrum and bias its interpretation. To quantify these effects, we developed a 3D radiative-transfer model able to generate transmission spectra of atmospheres based on 3D atmospheric structures. We first apply this tool to a simulation of the atmosphere of GJ 1214 b to produce synthetic JWST observations and show that producing a spectrum using only atmospheric columns at the terminator results in errors greater than expected noise. This demonstrates the necessity for a real 3D approach to model data for such precise observatories. Secondly, we investigate how day–night temperature gradients cause a systematic bias in retrieval analysis performed with 1D forward models. For that purpose we synthesise a large set of forward spectra for prototypical HD 209458 b- and GJ 1214 b-type planets varying the temperatures of the day and night sides as well as the width of the transition region. We then perform typical retrieval analyses and compare the retrieved parameters to the ground truth of the input model. This study reveals systematic biases on the retrieved temperature (found to be higher than the terminator temperature) and abundances. This is due to the fact that the hotter dayside is more extended vertically and screens the nightside – a result of the non-linear properties of atmospheric transmission. These biases will be difficult to detect as the 1D profiles used in the retrieval procedure are found to provide an excellent match to the observed spectra based on standard fitting criteria. This must be kept in mind when interpreting current and future data.

Key words. planets and satellites: general – planets and satellites: atmospheres – radiative transfer – techniques: spectroscopic

1. Biases in the analysis of transmission spectra of tridimensional planets

With the first spectroscopic observations of exoplanets, we are now able to study planetary atmospheres beyond our solar system. In recent years, spectroscopic observations have seen tremendous developments, and the coming years are even more promising, particularly due to the launch of the *James Webb* Space Telescope (JWST; Beichman et al. 2014) and the dedicated ARIEL mission (Tinetti et al. 2017).

One of the most common methods to interpret atmospheric spectra is based on inverse atmospheric retrieval modelling (Madhusudhan 2018). However, because of the complex thermal structure of the atmosphere and the numerous gases to retrieve (with their possibly complex spatial distribution) the number of parameters to handle can render the inversion computational cost prohibitive, that is if there are enough data to have a well-constrained problem to start with. As a result, retrieval algorithms are required to make drastic assumptions on the forward model to render the problem tractable.

A problem arises when these assumptions do not hold to a sufficient degree in the observed atmosphere, which creates a

systematic bias that can lead the retrieval algorithm far from meaningful solutions. Identifying and alleviating these biases is therefore a crucial goal to prepare for the next generation of precision observatories, and there have been several attempts in this direction. For example, the often-made assumption of uniform mixing ratio in the atmosphere led Evans et al. (2017) to retrieve a 100–1000 times solar VO/H₂O ratio in the atmosphere of WASP-121 b. However, Parmentier et al. (2018) showed that accounting for the chemical dissociation of some species at the hottest altitudes allowed them to understand the data with solar abundances. Rocchetto et al. (2016) also thoroughly quantified the impact of assuming a vertically isothermal atmosphere.

Nevertheless, all current retrieval algorithms are still fundamentally limited by the assumption of a spherically symmetric atmosphere: they use 1D forward models to constraint spectra and atmospheric parameters of 3D objects for which we expect heterogeneities. Such an approach is bound to create counter-intuitive biases that we need to quantify. With that in mind, Feng et al. (2016) investigated how the non-uniform flux emitted by the planet could actually create a false-positive signal for methane in emission. In the same vein, Blečić et al. (2017) used the dayside emission spectrum computed with a post-processed 3D Global

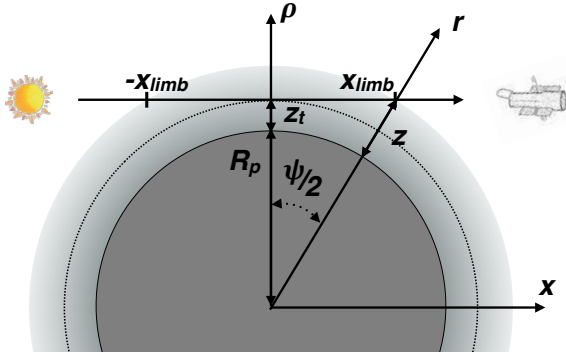


Fig. 1. Schematic of the geometry of a light ray crossing the atmosphere. The inner circle is the arbitrary reference surface of the planet of radius R_p . The lighter-grey region is the atmosphere. The distance from the centre of the planet is $r = R_p + z$. A light ray is defined by its distance of closest approach to the centre of the planet, ρ , and the corresponding tangential altitude $z_t = \rho - R_p$. The direction of the light ray defines the x direction. As we further discuss the extent of the limb along the ray, we introduce x_{limb} so that the absorption outside the $[-x_{\text{limb}}, x_{\text{limb}}]$ segment is negligible in determining the transit radius, and the corresponding limb opening angle ψ .

Climate Model (GCM) to identify the region effectively probed by the retrieval of secondary eclipse data.

On the transmission spectroscopy side, the study of the horizontal atmospheric heterogeneities have focused on the effect of clouds, with [Line & Parmentier \(2016\)](#) showing that the presence of clouds on only parts of the limb could mimic a high mean molecular weight atmosphere. Yet, these latter authors produced their forward spectra by simply averaging two 1D models so that only a limited kind of heterogeneity could be investigated. To go further [Charnay et al. \(2015\)](#), [Way et al. \(2017\)](#), [Parmentier et al. \(2018\)](#), and [Lines et al. \(2018\)](#), among others, produced transit spectra from 3D atmospheric simulations. However, because of the difficult geometry, they still relied on a 1D radiative-transfer transmission code that is either fed an average limb profile from a 3D simulation or that performs spectra of all the columns at the terminator of the model – that they assume to be equivalent to the limb plane – before averaging. Even if the second approach – that we hereafter refer to as limb-averaged or (1+1)D method – does capture the spatial variations of the atmosphere along the terminator, it completely neglects horizontal variations across it. Indeed, as the ray goes from the dayside to the nightside before coming to the observer, it crosses one of the most steeply changing regions of the atmosphere: the transition from day to night side. The effect of such a thermo-compositional transition within the limb on retrieved parameters is unknown at present. Indeed, although various authors have also developed a fully consistent transmission model able to predict such effects, these authors have not tried to retrieve physical parameters from their forward spectra. [Fortney et al. \(2010\)](#), [Burrows et al. \(2010\)](#), and [Dobbs-Dixon et al. \(2012\)](#), for example, focused on the potential differences between the east and west limbs, while [Miller-Ricci Kempton & Rauscher \(2012\)](#) and [Showman et al. \(2013\)](#) mainly looked at the effect of the doppler shifting by winds on high-resolution spectra.

2. Simple estimate of the maximum width of the limb and goals of the study

The day–night transition would not be a problem if the region probed in transmission were infinitely thin. We would just see a slice of the atmosphere. But in fact, and quite counter-intuitively,

the width of this region – which is our definition of the limb¹ – is much larger on some planets than generally expected. Therefore, the transit spectrum encodes a much wider diversity of temperatures and compositions.

Although the effect of this larger extent of the limb is demonstrated below a posteriori by the results of our 3D transit model, let us here try to give simple arguments to estimate how different planets can be affected. In other words, how wide can we expect the limb to be on any given planet?

Of course, the problem in providing such a simple estimate is that the region that contributes to the transit spectrum does not only depend on the global parameters of the planet, but also on the precise chemical-physical conditions in the atmosphere and how they vary spatially, as is demonstrated below. There is therefore a certain degree of arbitrariness if one wants to come up with a simple general estimate. For this reason, we first use a simple geometrical argument. The advantage of this is that it allows us to identify the key dimensionless parameter controlling the limb width. In [Appendix A](#) we derive a model of a more specific case of chemical inhomogeneity and show that the two approaches indeed yield similar results.

Let us consider a light ray passing through the limb as shown in [Fig. 1](#). Estimating the width of the limb comes down to the computation of the maximum distance from the limb plane, x_{limb} , at which the atmosphere still measurably affects the optical depth along all the observed rays, and especially the deepest one. The choice we have to make here is the highest pressure probed in transit (P_{bot} , that we assume to define the planetary radius R_p) and the lowest pressure at which the atmosphere is still able to significantly affect the transmission of a given ray (P_{top}). Then the width of the limb is given by

$$\psi \equiv 2 \arccos \left(\frac{R_p}{R_p + z(P_{\text{top}})} \right). \quad (1)$$

Due to the fact that, in an isothermal atmosphere with an atmospheric scale height at the surface equal to H and a varying gravity, the hypsometric relation writes

$$z(P_{\text{top}}) - z(P_{\text{bot}}) = H \ln \left(\frac{P_{\text{bot}}}{P_{\text{top}}} \right) \left(\frac{1}{1 - \frac{H}{R_p} \ln \left(\frac{P_{\text{bot}}}{P_{\text{top}}} \right)} \right), \quad (2)$$

we get

$$\psi \equiv 2 \arccos \left(1 - \frac{H}{R_p} \ln \left(\frac{P_{\text{bot}}}{P_{\text{top}}} \right) \right). \quad (3)$$

The first important result is that we see that the important dimensionless quantity in our problem is the ratio of the scale height to the planetary radius. The higher this parameter, the larger the curvature effects in the atmosphere. This parameter appears often in the following analyses. We also directly see that the larger the pressure range probed, the wider the limb. Many models predict that the lowest levels probed in transit are around 100 mb in the visible/near infrared. On the other hand, [Kreidberg et al. \(2014\)](#) showed that in order to explain the flat spectrum of GJ 1214 b, an opaque aerosol deck is needed as high as 10^{-3} – 10^{-2} mb, showing that absorbers at such altitudes can indeed still affect the transit spectrum.

¹ The limb, as defined here, should not be confused with either (i) the limb plane, which is the plane perpendicular to the observer’s line of sight passing through the planet centre, or (ii) the terminator, which is also a plane passing through the planet centre, but which is perpendicular to the star–planet axis. The latter two are confounded only when the star, planet, and observer are aligned.

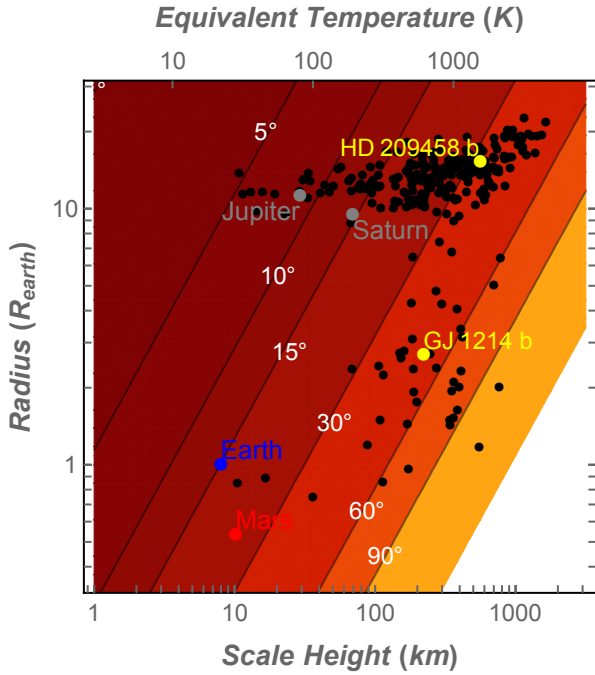


Fig. 2. Simple estimate of the opening angle of the region around the terminator that affects the transmission spectrum (i.e. the limb). The contours give the opening angle in degrees as a function of the planetary radius and the atmospheric scale height of the atmosphere at 10 mb (the variation of gravity with height is accounted for). The atmosphere is assumed to become transparent above the 1 Pa pressure level. Black dots are known planets for which the radius and surface gravity are measured. Some key objects are identified. The equivalent temperature is derived from the scale height assuming an atmosphere of hydrogen and helium (solar abundances) and a surface gravity of 10 m s^{-2} . For hot, low-gravity planets like GJ 1214 b, the limb can cover almost half the surface of the planet.

To get a quantitative idea of the possible width of the limb, Fig. 2 shows the result of Eq. (3) as a function of the planetary radius and scale height of the atmosphere. To obtain a relatively conservative estimate, we restrained the pressure range to $(P_{\text{bot}}, P_{\text{top}}) = (10 \text{ mb}, 10^{-2} \text{ mb})$. One can see that even for the Earth, the limb region already spans more than ten degrees. For a warm sub-Neptune planet like GJ 1214 b, this can be as large as 45° – 50° . In numerical terms, this means that within a 3D atmosphere model of this planet with a typical resolution, say 128 grid points in longitude for sake of concreteness, a single ray would interact with about 24 consecutive horizontal cells before leaving the planet. From these numbers, it becomes evident that the 1+1D approach, by picking only one out of those 24 cells as representative of the terminator, is a crude approximation to a real transmission spectrum.

Our goal is to identify the various biases of retrieval methods created by thermal and compositional – including clouds – inhomogeneities in the atmosphere in transmission. To that purpose, we need a transmission-spectrum generator able to match the complexity of a real 3D planet. We therefore developed a tool able to compute transmission spectra using a parametrized 3D atmospheric structure or the outputs of a 3D atmospheric simulation by a global climate model – namely the LMD Generic model (Wordsworth et al. 2011; Leconte et al. 2013; Charnay et al. 2015). This tool, Pytmosph3R, and its architecture are described in Sect. 3. We then show an extensive validation in Sect. 4. Subsequently, Sect. 5 presents a first application of this

tool to a simulation of the atmosphere of GJ 1214 b where we demonstrate the necessity of a real 3D approach to model data for such precise observatories. Finally, we investigate how day–night temperature gradients expected for exoplanets cause a systematic bias in retrieval analysis of real data performed with 1D forward models (Sect. 6).

3. Presentation of Pytmosph3R

Pytmosph3R is designed to compute transmission spectra based on 3D atmospheric simulations performed with the LMDZ generic global climate model. It produces transmittance maps of the atmospheric limb at all wavelengths that can then be spatially integrated to yield the transmission spectrum. The code is entirely written in Python. In this section we present the various modules of the code:

- the geometrical framework used to map the atmospheric structure from the spherical coordinates used by the GCM onto cylindrical coordinates that are more suitable for following photons crossing the atmosphere;
- the two algorithms that can be used for the calculation of the slant optical path – a discretised and an integral method;
- the various sources of opacity included in our radiative-transfer model;
- the spatial integration to produce spectra.

3.1. Definition of coordinate systems

3.1.1. The spherical grid used in atmospheric simulations

Typical 3D atmospheric simulations – LMDZ included – provide state variables such as temperature and mixing ratios of various absorbers/scatterers on a longitude/latitude/pressure grid. Although pressure is a convenient variable to compute atmospheric motions, transit spectroscopy is fundamentally about knowing the physical area of the opaque region of the atmosphere. We must therefore first interpolate the outputs of the climate model on a spherical (λ, φ, z) grid, where λ is the longitude, φ is the latitude, and z the altitude. When needed, we also refer to $r \equiv R_p + z$, the distance to the planet centre, or α the colatitude.

The longitude/latitude grid is evenly spaced and follows the native grid of the atmospheric model. The altitude grid is also evenly spaced. However, as is discussed in detail in Sect. 4, the resolution of this grid can, and usually should, be higher than the native resolution of the input simulation as it sets the precision of the output spectrum. We find that a good compromise between computation time and accuracy is reached for a vertical resolution of about a tenth of the scale height. We set n_z , n_λ , and n_φ to be the number of grid cells in each of the three dimensions.

The top should also be chosen to be high enough for the atmosphere to be transparent there. This is quantified hereafter. If this altitude is above the top of the input simulation, the code extrapolates the atmosphere above this top assuming hydrostatic equilibrium and a fixed temperature (or a profile that the user needs to define).

Subsequently, we integrate the hydrostatic equilibrium equation within each column of the model to compute the values of all the necessary variables from the climate model (e.g. temperature and mixing ratios) on this new altitude grid. During this integration, the variation of gravity with altitude is taken into account (see Eq. (7)), an effect that proved to be crucial to reach the precision needed.

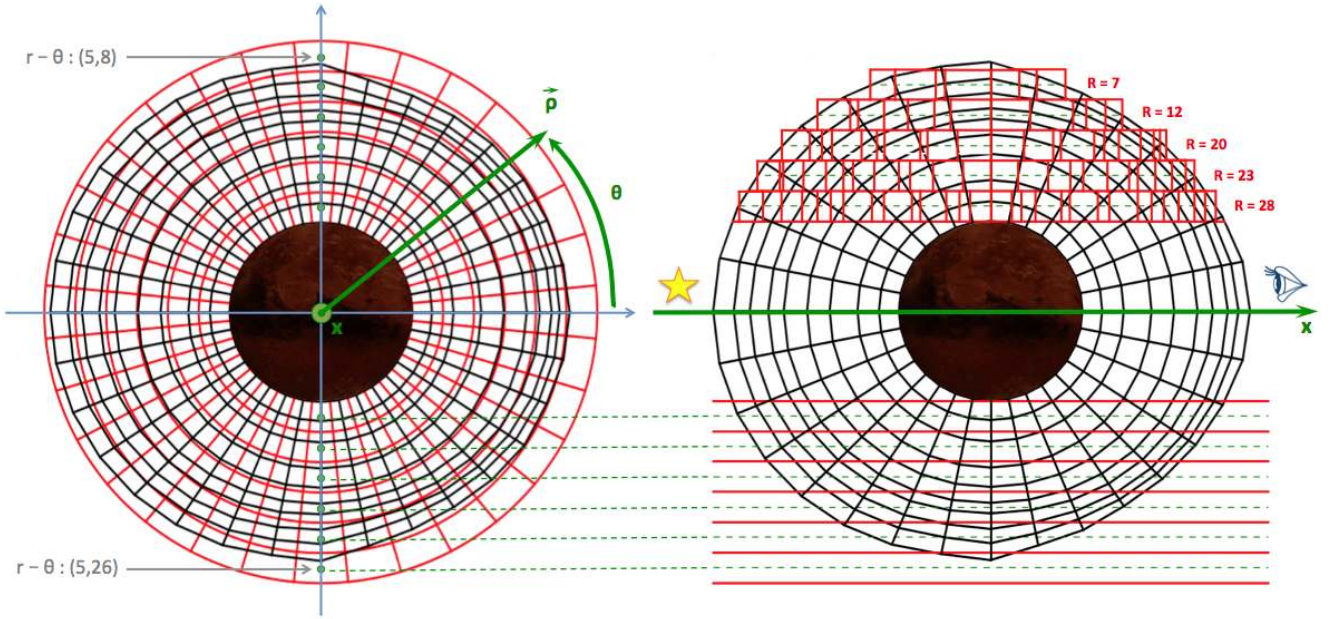


Fig. 3. Schematic showing the cylindrical grid (in green). *Left panel:* grid as seen from the observer. The pink dots are examples of (ρ, θ) -rays (to follow the Python convention used by the code, the indices given start at 0). The native GCM grid at the terminator of an example simulation is shown in black to illustrate its non-uniform vertical spacing. *Right panel:* the side where we can follow light rays (dashed lines) as they go horizontally from left to right. The red boxes illustrate how the x -discretisation is computed for each ray to follow closely where it goes from one cell of the spherical grid to another.

3.1.2. The cylindrical grid used for the radiative transfer

As long as the light from the star propagates in straight lines, cylindrical coordinates provide the most natural set to treat transit geometry. Indeed, due to the great distance between the observer and the planet, light rays follow parallel paths so that the radiative transfer can be solved within a hollow cylinder that is tangent to the planetary surface at the bottom and stops at the top of the modelled atmosphere.

We therefore define a cylindrical grid whose coordinates are denoted (ρ, θ, x) where the x -axis is the line between the centre of the planet and the observer (the line of sight), with x increasing toward the observer (see Fig. 3). The $x=0$ plane corresponds to the plane perpendicular to the line of sight passing through the centre of the planet (hereafter referred to as the plane of the sky). The factor ρ is the distance from the x -axis in the plane of the sky, and θ is the azimuth, defined here as the angle on the limb plane measured from a reference direction (see below).

Hereafter, we refer to the ray of light that crosses the limb plane at those coordinates as a “ (ρ, θ) -ray”. The transmittance map of the atmosphere is simply constructed by computing the chord optical depth of the ray along the x direction for every (ρ, θ) . From this, and assuming a luminosity/spectral distribution (limb-darkening, spots) over the stellar disc as well as a transit trajectory, we can produce spectral transit light curves.

The resolution of the cylindrical grid is based on that of the spherical one, which has a finer altitude resolution compared to the GCM: $\Delta\rho = \Delta z$ and $\Delta\theta = \Delta\varphi$. There is no benefit in increasing the resolution in θ , GCM cells being considered horizontally uniform. To test the impact of shooting our rays through the middle of our layers or at their interfaces, the ρ grid can be shifted relative to the r grid using the ω parameter ($0 \leq \omega < 1$) so that $\rho = r + \omega \Delta z$. To speed up computation, (ρ, θ) -rays are eventually divided irregularly along the x direction in order to closely follow their path from one spherical cell to another, as explained in Sect. 3.2 (see also Fig. 3).

3.1.3. Orientation of the planet and correspondence between coordinate systems

The cylindrical coordinate system needs to be properly oriented with respect to the spherical grid. For this, we simply require knowledge of the longitude and latitude of the observer in the spherical coordinates, $(\lambda_{\text{obs}}, \varphi_{\text{obs}})$, at the time of observation (or alternatively the colatitude α_{obs})². The unit vector pointing toward the observer, $\hat{\mathbf{u}}_{\text{obs}}$, then defines the direction of the x -axis of our cylindrical coordinates.

The last thing that we need to define is the arbitrary reference direction for the azimuth. For this we choose the projection of the rotation axis of the planet onto the plane of the sky.

With these two definitions, there is a unique relationship between the spherical and cylindrical coordinates for a given point. The translation from one system to the other with an arbitrary orientation however requires a set of non-linear equations to be solved; detailed in Appendix B.

3.2. Dividing (ρ, θ) -rays into subpaths

Along each (ρ, θ) -ray (identified by the indices i_ρ and i_θ) we locate all the intersections with relevant interfaces of the spherical grid and divide the ray into segments of irregular lengths, each of them belonging to a different cell. All quantities used to calculate the optical path – pressure and density excepted – are considered constant within each segment/cell. The pressure/density can be either kept constant within a segment/cell or assumed to follow hydrostatic equilibrium.

In practice we first divide each (ρ, θ) -ray with a constant Δx step, calculate the spherical coordinates of the resulting discrete

² In the simple case of a planet with a null obliquity, $\varphi_{\text{obs}} = |\pi/2 - i|$, where i is the orbital inclination with respect to plane of the sky. The sign depends on the convention for i . Furthermore, at conjunction, $\lambda_{\text{obs}} = \lambda_\star + \pi$, where λ_\star is the substellar longitude.

points and give them the three indices $(i_r, i_\lambda, i_\varphi)$ of the spherical cell they belong to (see Appendix B). The step Δx is chosen to be small enough ($< \Delta z$) that two successive points can only belong to either the same cell or to two adjacent cells (i.e. cells separated by a facet, an edge, or a corner). One, two, or all three indices $(i_r, i_\lambda, i_\varphi)$ can be incremented between two successive points.

When a change of index occurs between two points, the code analytically determines the position of the intersection(s) between the (ρ, θ) -ray and the surface(s) separating cells. This comes down to solving for the unknown position x_{int} along the ray knowing ρ , θ , and the equation of the surface(s) crossed. The equations to be solved for the three types of intersections (depending on the varying index) are detailed in Appendix C.

From these positions, the length of the subpaths belonging to individual cells can be measured. When more than one index is incremented between two points (near an edge or a corner, which implies that a third cell has been crossed) and once the intersections have been located, their x-position are sorted in increasing order so that subpaths can be measured and attributed to specific cells.

3.3. Optical depth

At this point, all the (ρ, θ) -rays are now subdivided into $N_x(\rho, \theta)$ segments of length $\{\Delta x_i(\rho, \theta)\}_{i=1, N_x}$. The number and length of these segments of course changes for each (ρ, θ) -ray depending on the number of intersections found in the previous step. Each of these segments has been assigned to a given cell of the spherical grid so that we know all the quantities describing the physical state of the atmosphere in the i th segment: temperature (T_i), volume mixing ratio of the j th of the N_{spe} species ($\chi_{j,i}$), and the mass mixing ratio of the k th of the N_{con} species of condensed particles ($q_{k,i}$).

The goal is now to compute the optical depth (hence the transmittance) of the atmosphere for each (ρ, θ) which is given by

$$\tau(\rho, \theta) = \sum_i^{N_x} \Delta\tau_i(\rho, \theta), \quad (4)$$

where $\Delta\tau_i$ is the optical depth of a given segment.

Pytmosph3R can calculate optical depth in two ways. Pressure (and density) can be either considered constant within a cell, hereafter referred to as the discretised method, or it can be integrated along the optical path assuming a hydrostatic vertical structure within the cell, the so-called semi-integral method. In both cases, we assume that cross-sections are constant along the given sub-path to spare considerable CPU time. This approximation yields negligible errors as long as the simulation is sufficiently resolved in the vertical.

3.3.1. Discretised calculation of the optical depth

With the discretised method, pressure and density are considered constant so that the optical depth of a segment simply reduces to

$$\Delta\tau_i = \left(\frac{P_i}{k_B T_i} \sum_j^{N_{\text{spe}}} \chi_{j,i} (\sigma_{\text{mol},j} + \sigma_{\text{sca},j} + \sigma_{\text{con},j}) + \sum_k^{N_{\text{con}}} k_{\text{mie},k} \right) \Delta x_i, \quad (5)$$

where σ_{mol} , σ_{sca} , and σ_{con} are the cross-sections for the molecular, Rayleigh scattering, and continuum absorptions, respectively, and k_{mie} is the absorption coefficient associated to the Mie scattering by aerosols. The parametrization used for these absorptions is discussed below.

3.3.2. Integral method

With the integral method, we now assume that the pressure follows the hydrostatic law within a segment, thus varying with altitude. The optical depth of a segment for any contribution is then given by

$$\Delta\tau_i = \int_{x_i}^{x_i+\Delta x_i} \frac{P(x)}{k_B T_i} \chi \sigma dx = \frac{\chi \sigma}{k_B T_i} \int_{x_i}^{x_i+\Delta x_i} P(x) dx, \quad (6)$$

where x_i is the position of the beginning of the segment. We use z_i to refer to the corresponding altitude. The simplification comes from the fact that the altitude profile of the pressure within an isothermal atmosphere is analytical even if the gravity varies with height. It is given by

$$\begin{aligned} \ln \left(\frac{P(z_{i+1})}{P(z_i)} \right) &= \int_{z_i}^{z_{i+1}} \frac{-Mg(z_i)}{RT} \frac{dz}{\left(1 + \frac{z-z_i}{R_p+z_i}\right)^2} \\ &= -\frac{Mg(z_i)}{RT} \left(\frac{z_{i+1}-z_i}{1 + \frac{z_{i+1}-z_i}{R_p+z_i}} \right), \end{aligned} \quad (7)$$

where M is the molar mass of the gas, R the universal gas constant, and $g(z)$ the gravity at a given altitude. This entails

$$\begin{aligned} \int_{x_i}^{x_i+\Delta x_i} P(z(x)) dx &= \int_{z_i}^{z_i+\Delta z_i} P(z_i) \exp \left[-\frac{Mg(z_i)}{RT_i} \left(\frac{z-z_i}{1 + \frac{z-z_i}{R_p+z_i}} \right) \right] dx \\ &= \int_{z_i}^{z_i+\Delta z_i} P(z_i) \exp \left[-\frac{Mg(z_i)}{RT_i} \left(\frac{z-z_i}{1 + \frac{z-z_i}{R_p+z_i}} \right) \right] \\ &\quad \times \frac{(R_p+z)}{\sqrt{(R_p+z)^2 - \rho^2}} dz, \end{aligned} \quad (8)$$

where the last integration is carried out numerically between the lowest and the highest point of the segment. The increased accuracy of this method is due to two factors: (i) the variation of gravity with height is built in, and (ii) more importantly the exponential variation of pressure is fully accounted for. This explains why a much lower number of vertical layers are needed with this method to reach numerical convergence.

3.4. Sources of opacity

3.4.1. Molecular lines

Pytmosph3R deals with molecular absorptions in two possible ways: tables of monochromatic cross-sections or correlated- k coefficients (also called the k -distribution method; Fu & Liou 1992).

Using k -distributions considerably reduces the computing time but new k -tables must be pre-computed each time one wants to change the atmospheric composition or the resolution of the output spectrum.

Because our code is designed to work with the TauREx retrieval code (Waldmann et al. 2015), we use the same set of high-resolution cross-sections produced by the ExoMol project. This data set is precomputed on a $T - \log P$ grid going from 200 to 2800 K every 100 K and from 10^{-3} to 10 bar every 0.3 dex. Pytmosph3R users can either choose linear or optimal interpolation in temperature. For the optimal interpolation we follow Hill et al. (2013) who prescribes

$$\sigma_{i,\lambda}(T) = a_{i,\lambda} \exp \left(-\frac{b_{i,\lambda}}{T} \right), \quad (9)$$

where i is the molecular/atomic species index, λ the wavelength, and T the temperature of the cell. The (a, b) scaling factors are given by

$$b_{i,\lambda} = \left(\frac{1}{T_1} - \frac{1}{T_u} \right)^{-1} \ln \frac{\sigma_{i,\lambda}(T_u)}{\sigma_{i,\lambda}(T_1)}, \quad (10)$$

$$a_{i,\lambda} = a_{i,\lambda}(T_u) \exp\left(\frac{b_{i,\lambda}}{T_u}\right), \quad (11)$$

where T_u and T_1 are upper and lower temperatures, respectively. k -distributions can only be interpolated linearly in temperature. Along the pressure coordinate, the interpolation is log-linear.

The effect of using different interpolation schemes has been tested and we find that it does not introduce significant differences.

3.4.2. Continuum absorptions

Our principal source of continuum opacity is due to collision-induced absorptions. We account for this process for all the species for which such information is available in the HITRAN database and following the prescriptions of [Richard et al. \(2012\)](#). Furthermore, for specific species such as water vapor, we can add a continuum that accounts for the truncation of the far wings of the lines and the neglect of many weak lines in some of our sets of cross-sections or correlated- k tables. In such cases, the water continuum is added using the CKD model ([Clough et al. 1989](#)). We however note that care must be taken to ensure that the molecular opacities used must be computed consistently so as not to count some effects twice.

3.4.3. Rayleigh scattering

Multiple scattering is neglected. The contribution of Rayleigh scattering is therefore treated as a simple extinction. The cross-section of any single gas molecule is given by the common formula

$$\sigma_{\text{sca}}(\lambda) = \frac{24\pi^3}{N_{\text{std}}^2 \lambda^4} \left(\frac{n_\lambda^2 - 1}{n_\lambda^2 + 2} \right)^2 F_k(\lambda), \quad (12)$$

where λ is the wavelength (here in m), N_{std} is the number density of a gas under standard conditions, n_λ is wavelength-dependent, real refractive index of the gas, and $F_k(\lambda)$ is the King correction factor which accounts for the depolarization. The accuracy of this essential part of the radiative transfer mainly depends on the calculation of refraction indices. We used the most recent data available in the literature. For sake of completeness, we have reviewed the parametrization that we use for H_2 , He, H_2O , N_2 , CO, CO_2 , CH_4 , O_2 , and Ar in Appendix D and in Table D.1.

3.4.4. Mie scattering for aerosols

Transmission spectra of transiting exoplanets are affected by clouds and hazes. The LMDZ GCM can include cloud physics and provide the properties and 3D distribution of liquid/solid condensates and aerosols. Assuming spherical particles with a size similar to or larger than the considered wavelengths, we use Mie scattering formalism to compute their extinction factor Q_{ext} and resulting opacities, following the same method as in radiative-transfer modules of the GCM ([Madeleine et al. 2011](#)). We linearly interpolate the value Q_{ext} on effective radius and

wavelength using pre-calculated lookup tables. The absorption coefficient is estimated as

$$k_{\text{mie},k}(\lambda) = \frac{3}{4} \frac{Q_{\text{ext}}(r_{\text{eff}}, \lambda)}{\rho_{\text{con}} r_{\text{eff}}} q_k \rho_{\text{gas}}, \quad (13)$$

where Q_{ext} is the extinction coefficient for the wavelength λ and a given effective radius r_{eff} , q_k and ρ_{con} are the mass mixing ratio and the density of the species considered, respectively, and ρ_{gas} is the total gas density.

3.5. Generation of transmittance maps and spectra

To generate the global absorption spectrum of the planet, the wavelength-dependent (ρ, θ) map of optical depth is first converted into a transmittance map

$$\mathcal{T}(\rho, \theta, \lambda) = e^{-\tau(\rho, \theta, \lambda)}. \quad (14)$$

In the most general case, the in-transit flux should be computed by convolving this transmittance map with a given surface-brightness distribution for the star. However, in the most simple case of a homogeneous stellar disc, the effective area of the planet reads

$$\begin{aligned} A_p(\lambda) &= \pi R_p^2 + \sum_{\rho, \theta} (1 - \mathcal{T}(\rho, \theta, \lambda)) S_\rho \\ &= \pi R_p^2 + \sum_{\rho, \theta} (1 - e^{-\tau(\rho, \theta, \lambda)}) S_\rho, \end{aligned} \quad (15)$$

where R_p is the radius of the planet (at the bottom of our model atmosphere), $S_\rho = 2\pi(\rho + \Delta\rho/2)\Delta\rho/N_\theta$, N_θ being the number of θ points, and $\Delta\rho$ is the layer thickness. Eventually, the relative dimming of the stellar flux due to the planet is given by

$$\Delta F(\lambda) = \frac{A_p}{\pi R_\star^2} = \left(\frac{R_p}{R_\star} \right)^2 + \sum_{\rho, \theta} \frac{(1 - e^{-\tau(\rho, \theta, \lambda)}) S_\rho}{\pi R_\star^2}, \quad (16)$$

where R_\star is the stellar radius. For each monochromatic transit depth, we can determine an effective radius of the planet defined as

$$R_{\text{eff}}(\lambda) = R_\star \sqrt{\Delta F(\lambda)}. \quad (17)$$

One can simulate a realistic light curve by precisely locating the planet transmittance map over the weighted stellar disc at each time-step of the transit. During ingress and egress, where the stellar disc is hidden by only a fraction of the planet, only the elements in front of the stellar surface need to be considered. Astrophysical (stellar) noise can then be added by simulating a realistic spatial distribution of the star luminosity and its variability during a transit, for example by taking into account limb darkening, spots, and granulation ([Chiavassa et al. 2017](#)). This model complements other codes able to produce synthetic emission/reflection spectra and light curves from GCM simulations ([Selsis et al. 2011](#); [Turbet et al. 2016, 2018](#)).

4. Model validation

4.1. Validation approach

In the following subsections we validate the different modules of the code: calculation of the vertical structure, 3D geometry of the radiative transfer, integration scheme for the optical depth, interpolation scheme for the opacities, and so on. In the absence of another validated code able to produce synthetic transmission

spectra from 3D simulations, we decided to compare our code with a hierarchy of 1D models with increasing complexity – from a purely analytical model (Sect. 4.2) to the more realistic forward spectrum generator used in the TauREx retrieval tool (Waldmann et al. 2015; Sect. 4.3).

In each case, the comparison with a given 1D model proceeds as follows: on one hand we generate a spectrum from the 1D model with a given vertical temperature profile (possibly isothermal). On the other hand, we use the same vertical profile to generate a full spherically symmetric 3D structure on our spherical grid. This structure goes through our whole code to generate a spectrum that should ideally be the same as the 1D spectrum. Finally, we increase the resolution of our 1D/3D grids until convergence is reached.

Relatively early in our tests, it appeared that, to numerical precision, our transmittance maps were completely insensitive to the azimuth angle, θ , in any spherically symmetric configuration, as it should be. This not completely trivial result shows that our careful way of computing the path length of the ray in each cell indeed prevents singularities – in particular at the poles – in our initial longitude/latitude grid. This also allowed us to test the convergence of the algorithm of vertical integration at very high resolutions. Indeed, when a very large number of layers is used – above ~ 1000 , which is already much larger than what is needed in practice for convergence – the computing time for the full 3D code becomes prohibitive. For this reason, in some cases below, the results shown are derived from a sector of the limb only (i.e. a given θ). This approach is made possible by the fact that the numerical differences between the various sectors is negligible.

As expected from our analytical arguments and further demonstrated in the following sections, we find that the most, and in fact almost only, important parameter determining the resolution needed to model a given atmosphere is the ratio of the planetary radius to the scale height (R_p/H): the lower this parameter, the larger the vertical resolution needed. This stems from two factors. A lower ratio means that (i) gravity will vary more significantly in each vertical layer, and (ii) curvature effects will be more pronounced and rays will pass through more atmospheric columns along their path (see Fig. 2). As a result the validation cases we present hereafter focus on covering a wide variety of R_p/H . However, because the scale height is not set, but is determined from the composition and temperature of the atmosphere and from the surface gravity, fixing R_p/H entails varying one of these parameters to compensate. For each of the validation setups described in the following sections, the compensation parameter is therefore always stated. This can sometimes result in a physically inconsistent set of planetary parameters that we deem acceptable in the context of our validation. The values used for the other parameters are given in Table 1, unless otherwise stated.

As becomes clear in the following, our 3D model can reproduce our validation cases to numerical precision. We used these tests to further derive some guidelines on the number of layers and the model roof pressure to use in various cases to reach a satisfactory accuracy. To quantify what we call a “satisfactory accuracy”, we use the difference on the predicted transit depth between two models and ask that it be significantly smaller than the photon noise that can be expected for the given target in one transit. If one wants a more stringent condition, one can use the expected systematic noise floor for a given instrument (expected to be on the order of a few parts per million with JWST, for example).

Table 1. Numerical values for the parameters used in our fiducial validation case.

Planetary radius	R_p (m)	7×10^7
Surface gravity	g_0 ($\text{m}^2 \text{s}^{-1}$)	8.8
Surface pressure	P_s (Pa)	1×10^6
Top pressure	P_{top} (Pa)	1×10^{-4}
He mass fraction		0.17
Water volume mixing ratio	$\chi_{\text{H}_2\text{O}}$	0.05
Mean molar mass	M_a (kg mol^{-1})	2.7×10^{-3}
Opacity of gray absorber	κ ($\text{m}^2 \text{kg}^{-1}$)	1×10^{-3}

4.2. Validation with a monochromatic analytical model at constant gravity

To focus on the basics of radiative transfer we first compare the optical depths computed with our model with those given by an analytical solution. For an isothermal, horizontally uniform atmosphere with constant gravity and a scale height $H \ll R_p$, Guillot (2010) provides an analytical formula for the transmission optical depth as a function of the distance to the centre of the planet:

$$\tau_{\text{tr}} = \tau_{\perp} \left(\frac{2\pi R_p}{H} \right)^{\frac{1}{2}}, \quad (18)$$

where τ_{tr} is the chord optical depth and τ_{\perp} the vertical optical depth. As described in Appendix A, we use this formula to calculate the planet transit. We then compare the results with those of Pytmosph3R for a wide range of vertical resolutions: 50 to more than 10 000 levels. For 1000 levels and more, and as long as the thin atmosphere assumption is respected, the effective absorption radii from both models agree within ~ 1 cm (i.e. $\sim 10^{-9}$ relative accuracy).

4.3. Comparison to the TauREx forward model

4.3.1. Reasons for benchmarking against the forward model of an atmospheric retrieval tool

As one of our main goals is to perform a retrieval on a spectrum derived from a 3D model as if it were from a real planet, we decided to use the forward model of an existing, and validated, atmospheric retrieval code for our comparison.

The added benefit of this validation approach is that we make sure that in the relevant case of spherically symmetric atmosphere, both our 1D and 3D forward models can produce spectra without any numerical biases – meaning here that the differences between the spectra can always be reduced to be much smaller than the noise that will be prescribed in the retrieval step by choosing a sufficient vertical resolution.

As a result, we know that when we retrieve the properties of an atmosphere generated with our 3D model, the various potential biases in the retrieved parameters are entirely due to the heterogeneities of the atmosphere and not the differences in the numerics.

In our case, we are using TauREx, whose forward model is described in more detail in Waldmann et al. (2015). For the reasons mentioned above, we decided to use the same philosophy used in TauREx in the implementation of many physical processes, in particular concerning the molecular opacities and the Rayleigh scattering. Two notable exceptions are (i) the vertical

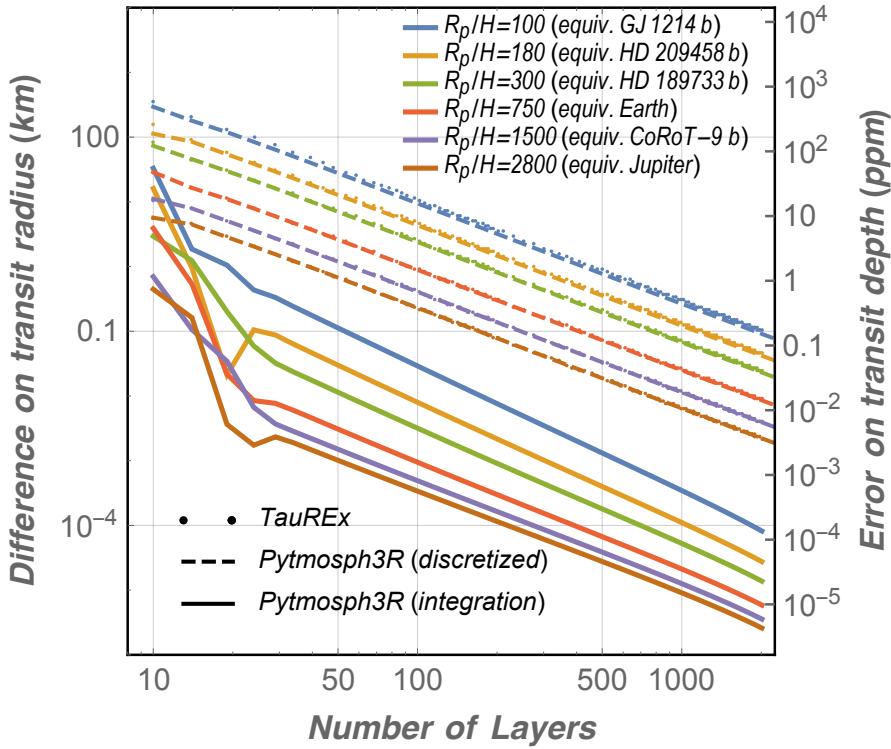


Fig. 4. Absolute difference between the transit radius computed for a given number of layers and model and our reference case (integral method, isothermal, grey opacity, 50 000 layers). The dots show the results of TauREx, the dashed lines represent Pytmosph3R in its discretised version, and the solid lines are for the integral version. Results are plotted for six values of R_p/H that increase from top to bottom for each set of curves. The names of planets representative of these values are labelled. For Pytmosph3R, we use the results of one sector only ($\theta=0^\circ$) to keep the computing time reasonable for the most resolved simulations. The axis on the right converts this radius difference into a transit depth difference in the case of GJ 1214 b (i.e. a planet of 2.6 earth radii in front of a star of 0.2 solar radii). To rescale this for any type of planet or star, we multiply by $(R_p/2.7 R_E)(0.2 R_\odot/R_\star)^2$.

grid and (ii) the optical path calculations that are discussed in greater detail in the following section. Let us just mention here that our requirement that our 3D model be as compatible with TauREx as possible is the reason for keeping two algorithms for the calculations of the optical depth in the code: the integral scheme which is the optimal scheme in terms of convergence and should be used in general, and the discretised one that always gives a result closer to TauREx when the same number of layers is used in both models.

4.3.2. Validation for an isothermal, grey atmosphere with varying gravity

We now assume an isothermal atmosphere with an altitude-dependent gravity. We consider a uniform composition of $H_2/He/H_2O$. In this first step, to validate only the geometrical part of the code, we assume that H_2 and He are transparent, and that water has a grey opacity that is independent of temperature and pressure. In this set of simulations, the scale height is set to the prescribed value by adjusting the temperature of the atmosphere, all other parameters being kept fixed to the value given in Table 1.

For each R_p/H ratio, we first compute the transit radius of a high-resolution reference model using Pytmosph3R with 50 000 layers. We then compute the difference between this reference and the transit radius given by our three models at various vertical resolutions. This is summarised in Fig. 4 where the results for the Pytmosph3R/discretised method are shown using dashed lines, for the Pytmosph3R/integral method using solid lines, and TauREx using dots. The various colours are for different R_p/H . We note that although we quote some planet names for each ratio to give an idea of what type of planet it describes, all calculations were performed for the same planetary radius ($1 R_J$).

The conclusions that can be drawn from this test are the following:

- all three models converge toward the same result as vertical resolution is increased so that the differences can be reduced to an arbitrarily low value;
- as advertised, atmospheres with a low R_p/H require less vertical layers to reach a given accuracy;
- for the same number of layers, the accuracy of our integral version of the code is orders of magnitude better than the discretised one. This is thus the preferred mode for most applications;
- as expected, the discretised mode, although less accurate, is always closer to TauREx and should probably be used if a retrieval analysis is to be performed. This indeed introduces as little bias as possible during the retrieval step.

In terms of the number of layers needed to reach convergence – the precision needed depends of course on the observations – the integral mode of Pytmosph3R requires as little as 20–30 layers in almost all practical cases of interest. This yields about two to three points per pressure decade or about one per scale height. For the discretised version as well as for TauREx and indeed most forward models using the same philosophy, 100 layers are generally enough, but this should be taken with caution. Hot and low-gravity objects in particular require a finer resolution and some published models have probably reached convergence only marginally.

Despite our efforts, it can be seen that for a given number of layers, the discrete version of Pytmosph3R is not exactly equivalent to TauREx. We find that this small discrepancy is fundamentally due to the fact that our vertical grid uses altitude rather than pressure levels as is usually the case in 1D models. When gravity varies with altitude, an iso-altitude grid is not equivalent to an iso-log pressure grid, hence the small difference.

4.3.3. Comparison with the full TauREx forward model

Here we compare spectra from Pytmosph3R and TauREx for isothermal atmospheres with the same composition as before.

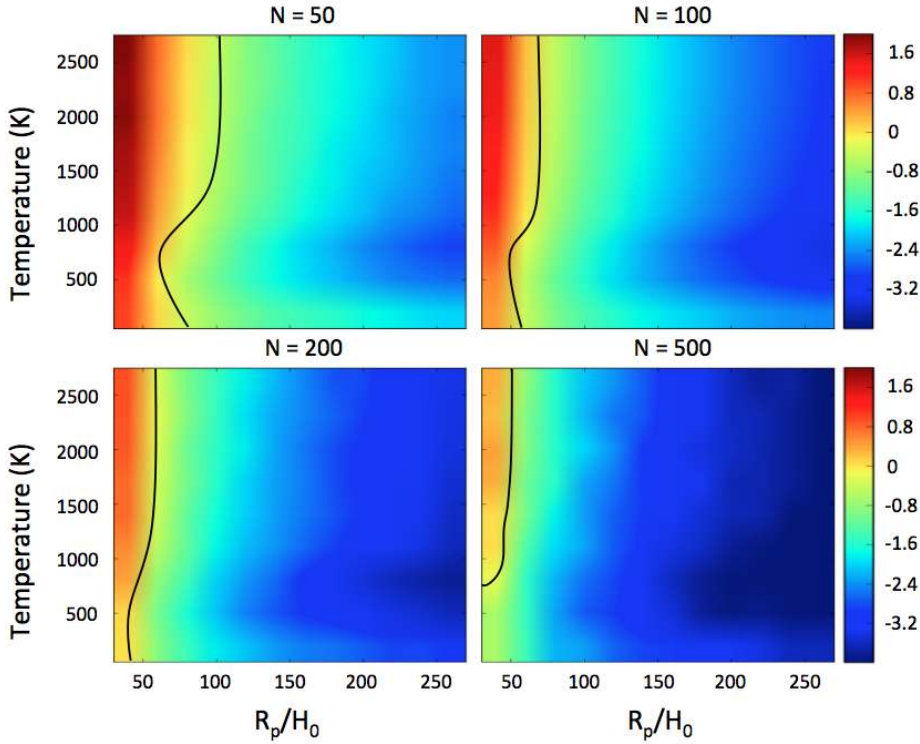


Fig. 5. Decimal logarithm of the reduced chi-square for the effective radius between our 3D model and the full forward model from TauREx over a range of temperature between 200 and 2600 K and of R_p/H between 40 and 260. As a comparison, all planets described in Fig. 4 have $R_p/H > 100$. The black line shows where the reduced χ^2 equals unity, i.e. where model differences become insignificant compared to the noise. We assumed an error calculated with a sun-like star at 100 ly, 1 h of integration and JWST instruments. The same work has been done for four vertical resolutions: 50, 100, 200 and 500 layers.

There are two differences however. First, we now compute a full spectrum with all our opacity sources varying with wavelength, temperature, and pressure. Second, we now fix the atmosphere temperature beforehand and adapt the surface gravity of the planet to get the desired R_p/H ratio. This stems from the fact that temperature will now impact the molecular features directly and not only through the scale height. This allows us to test our interpolations in the opacity databases as well.

As the model produces a full spectrum and not only a monochromatic transit radius, we quantify the differences between the two codes using a reduced chi-square calculated as follows:

$$\chi^2 = \sum_{\lambda} \frac{(\delta_{3D,\lambda} - \delta_{1D,\lambda})^2}{\sigma_{\text{obs},\lambda}^2} \frac{1}{N_{\lambda}}, \quad (19)$$

where $\delta_{3D,\lambda}$ and $\delta_{1D,\lambda}$ are the transit depths at the wavelength λ given by Pythmosph3R and TauREx, N_{λ} is the number of spectral bins, and $\sigma_{\text{obs},\lambda}$ is the stellar photon noise computed for JWST observations. The latter is given by

$$\sigma_{\text{obs},\lambda} = \frac{1}{\sqrt{N_{\text{obs},\lambda}}}, \quad (20)$$

where

$$N_{\text{obs},\lambda} = \frac{\pi \tau_{\lambda} c \Delta t R_*^2 A}{2d^2} \int_{\lambda_{\text{wl}}}^{\lambda_{\text{wl}+1}} \frac{d\lambda}{\lambda^4 \left(\exp\left(\frac{h_{\text{pc}}}{k_{\text{B}} T_s \lambda} - 1\right) \right)}. \quad (21)$$

Here, τ_{λ} is the system throughput, R_* and T_s are the radius and temperature of the host star, respectively, A is the collecting area of the telescope (here 25 m²), Δt is the integration time, and d is the distance of the target. In our example, we considered a Sun-like star, a one-hour integration, and a system at 100 ly. Increasing the noise tends towards making Pytmosph3R and TauREx spectra indistinguishable.

We computed the logarithm of this reduced chi-squared as a function of the atmospheric temperature and R_p/H for four different vertical resolutions. Results are shown in Fig. 5.

As already discussed, we can see that the agreement between the two codes at a given vertical resolution generally increases with R_p/H . At a given R_p/H ratio and in this specific example, the agreement also slightly depends on temperature. This is due to the overall increase of H₂O opacity with temperature that pushes the opaque region upward and increases the limb opening angle (see Fig. 2). With high enough vertical resolution the two models can however agree well within the noise budget. As the lowest encountered R_p/H is about 100 (for a hot-Neptune like GJ 1214 b), we confirm that 50 layers should be sufficient in most cases. These maps can however be used as a guide to chose the resolution needed if a more stringent case is found.

4.4. Effect of model top altitude

When computing transmission spectra, if the model does not reach a height in the atmosphere that is transparent enough at all wavelengths, the transit radius of the planet may be underestimated in opaque parts of the spectrum. The choice of the model top pressure therefore generally results from a trade-off between computation time and convergence.

This point is even more crucial here because several other technical reasons can limit the maximum altitude of a 3D climate model to a few tenths of a pascal or more (short radiative timescales, absence of non-LTE radiative treatment or conductive heat transfer, etc.).

To illustrate these limitations, Fig. 6 shows two spectra computed by Pytmosph3R for a simulation of GJ 1214 b (see Sect. 5). The black spectrum directly uses the outputs of the global climate model and stops around a pressure level of 0.5 Pa. The blue spectrum is obtained by extending the model upward to 10⁻⁴ Pa assuming that the atmosphere remains isothermal above the GCM model top. For this particular case, this is necessary

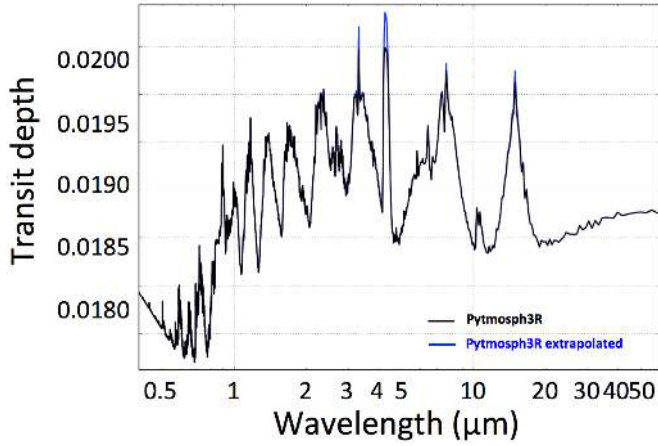


Fig. 6. Transit spectrum generated by Pytmosph3R from the outputs of a GCM simulation of GJ 1214 b (Charnay et al. 2015). The black curve represents the effective radius obtained when accounting only for the part of the atmosphere explicitly modelled by the GCM (down to ~ 0.5 Pa). The blue curve shows the result when the model is extended by an isothermal atmosphere down to 10^{-4} Pa. Further lowering the pressure of the model top does not alter the spectrum.

and sufficient to reach a degree of convergence commensurate with the future precision of the data (~ 1 – 10 ppm). Extending the model to even lower pressures does not significantly change the resulting spectrum.

5. The case of GJ 1214b as an illustration of Pytmosph3R capabilities

In this section, we apply Pytmosph3R to GCM simulations of the atmosphere of GJ 1214 b to illustrate the possibilities of the code. This also allows us to show examples of horizontal inhomogeneities and some of their effects on transmission spectra. Despite the flat spectra currently obtained with WFC3/HST, possibly due to high-altitude clouds/aerosols (Kreidberg et al. 2014), GJ 1214 b is one of the rare known targets that is not a gas giant but still offers a favourable configuration for transmission spectroscopy: the vicinity to the Sun (13 pc), a low-density implying the presence of an atmosphere, a short period (1.58 d), a red dwarf host, and an expectedly high scale-height-to-radius ratio (see Fig. 2). Modelling the formation, distribution and spectral signature of aerosols motivated the use of a 3D atmospheric model of the atmosphere (Charnay et al. 2015) to compute transmission maps and spectra with Pytmosph3R.

5.1. Input 3D atmospheric model

We use the “100 \times solar metallicity” simulation of Charnay et al. (2015) made with the LMDZ generic global climate model. The simulation has a 64×48 horizontal resolution with 50 layers equally spaced in log pressure, spanning 80 bars to ~ 0.5 Pa. An important assumption in this simulation is the local chemical equilibrium: in each cell, the composition is imposed by the local temperature, pressure, and elemental composition. As the simulation assumes a circular orbit with null obliquity and a synchronized rotation, the state of the atmosphere after convergence is relatively stable, exhibiting only stochastic variations. We use an arbitrary time-step to produce synthetic spectra. The thermal structure is shown in Fig. 7 and the absorber/cloud distribution

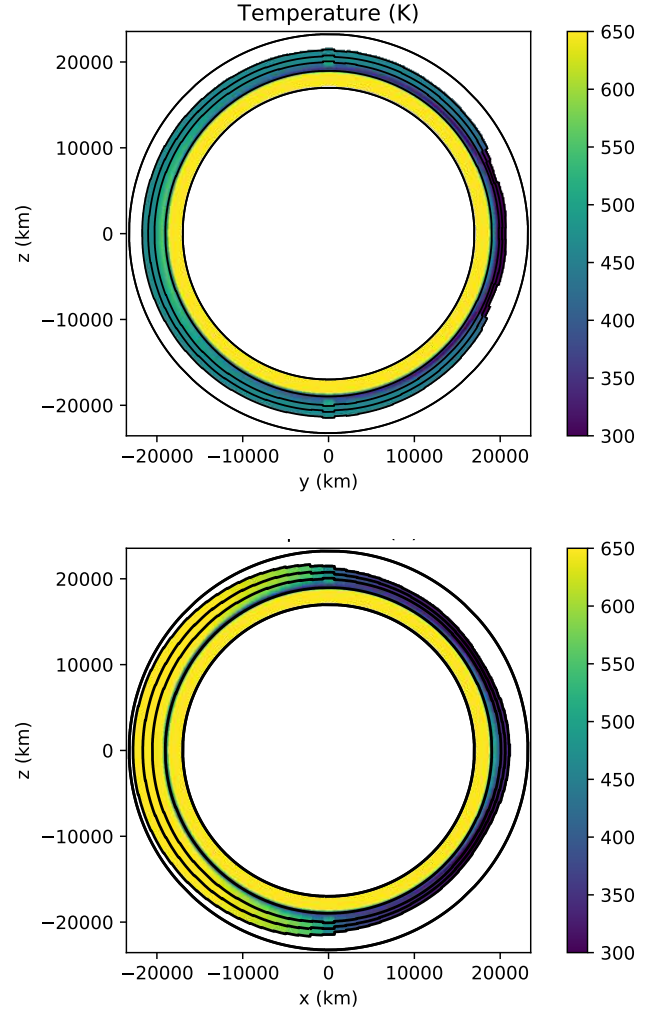


Fig. 7. Colour maps showing the temperature distribution in the model atmosphere of the 100 \times solar metallicity simulation from Charnay et al. (2015). The inner white disc represents the inner part of the planet with a radius assumed to be equal to that of GJ 1214 b (17.600 km). *Top panel:* temperature at the terminator. The planet is seen from the observer during transit, the poles being at the top and bottom. *Bottom panel:* the star is on the left and the observer on the right on the $z = 0$ axis. The poles are on the $x = 0$ axis. From centre outward, the five solid lines are respectively the 10^6 , 10^3 , 1, 10^{-2} , and 10^{-4} Pa pressure levels. The outer circle is there as an eye guide to highlight the non-sphericity of the planet. Temperature is well homogenised below the $\sim 10^3$ Pa level. Maps are to scale and show that the dayside is noticeably more vertically extended than the nightside.

in Fig. 8. We redirect the reader to Charnay et al. (2015) for more details on the model.

5.2. Transmittance maps

In order to understand the global transmission spectrum Pytmosph3R offers the possibility to draw a transmittance map in any spectral bin of the spectrum. Viewing the azimuthal and vertical inhomogeneities provides some important information for the interpretation of spectral features.

Figure 10 shows transmittance maps for a simulated atmosphere of GJ 1214 b, with and without the effect of clouds. Temperature is directly or indirectly at the origin of the inhomogeneities. Colder regions at the western terminator are less extended vertically, and locally induce a smaller absorption

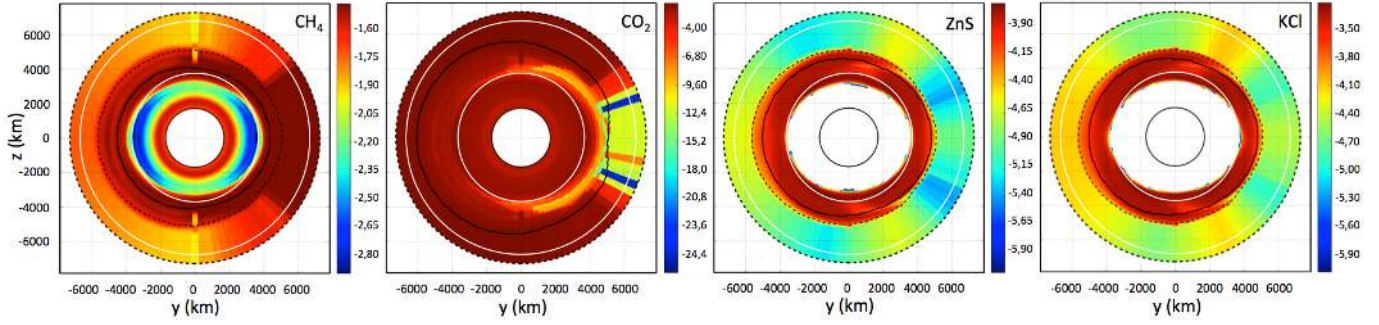


Fig. 8. Distribution of the logarithm of the mass mixing ratio of some absorbing species at the terminator in a simulation of the atmosphere of GJ 1214 b by Charnay et al. (2015). For KCl and ZnS, only the molecules in the condensed phase are considered. The dashed (solid) black curves are the contour of unit optical depth with (without) cloud opacities at a wavelength relevant to the molecule considered ($1.17\ \mu\text{m}$ for CH_4 ; $4.30\ \mu\text{m}$ for CO_2 ; $1.08\ \mu\text{m}$ for KCl et ZnS). The size of the inner white disc delimiting the base of the atmosphere is reduced for clarity. The whole simulated atmosphere is shown. The two white circles delimit the region enlarged in Fig. 10.

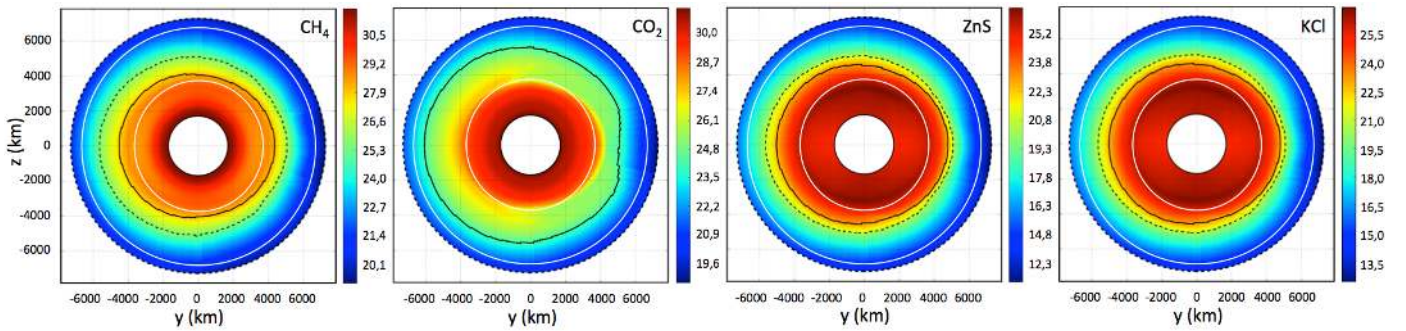


Fig. 9. Decimal logarithm of the column density of some molecules in the star-observer direction for a simulation of the atmosphere of GJ 1214 b by Charnay et al. (2015). See the caption of Fig. 8 for details and notations.

radius. As the simulation imposes a local chemical equilibrium, colder regions are also poorer in CO_2 and richer in CH_4 , affecting the transmittance distribution in the displayed CO_2 (14.9 and $4.3\ \mu\text{m}$) and CH_4 ($1.17\ \mu\text{m}$) bands. In the $14.9\ \mu\text{m}$ CO_2 bands another thermal effect comes from the strong temperature dependency of the absorption cross-section, which strongly enhances the transmittance inhomogeneities compared with the $4.3\ \mu\text{m}$ band, and is much less sensitive to temperature. The effect of clouds is the strongest at $0.95\ \mu\text{m}$ although they mask some chemical inhomogeneities at high pressure.

5.3. Comparison with averaging methods based on 1D models

In the absence of a code like Pytmosph3R, producing a transmission spectrum from a 3D simulation using a radiative-transfer model based on a horizontally homogeneous atmospheric profile implies an average of some kind.

5.3.1. Mean profile

One method consists in averaging first the atmospheric quantities and then computing a transmission spectrum. One single atmospheric profile (temperature, pressure, chemical abundances and cloud properties as a function of altitude) is obtained by averaging the $2N - 2$ profiles found on the terminator, weighted by the fraction of azimuth covered by each cell (N being the number of latitude points on the simulation grid). The transmission spectrum is then computed assuming that this atmospheric profile covers the whole planet. Figure 11 shows the comparison between spectra resulting from this mean profile method and those generated by Pytmosph3R.

The difference is mostly due to the lower-temperature patch at the morning terminator near the equator (west of the substellar point) which is due to the equatorial jet bringing cold air from the nightside there (see Fig. 7). This difference is considerably larger than the expected photon noise with JWST over most of the spectrum making this method clearly inadequate. This, in a sense, already highlights why 1D retrieval methods may be biased when interpreting the spectra of real atmospheres.

5.3.2. Limb integration

A better technique that is often applied (e.g. by Charnay et al. 2015; Way et al. 2017; Parmentier et al. 2018; Lines et al. 2018), involves computing in a first step $2N - 2$ transmission spectra assuming a horizontally homogeneous atmosphere, one for each atmospheric column found on the terminator. The total transmission spectrum is then calculated as an average of these intermediate spectra, weighted by the fraction of azimuth covered by each column. Figure 12 compares spectra obtained with this limb integration technique and shows the comparison between spectra resulting from this approach (red line) and from Pytmosph3R (red line).

A quick comparison of Figs. 11 and 12 clearly shows that limb integration performs much better than the mean profile approach. However, significant discrepancies remain throughout the spectrum and especially in some molecular bands. By definition, these differences only come from the atmospheric inhomogeneities along the path of the ray. These are due to the effect of the following factors.

- Day–night temperature gradient: as the dayside is hotter than the night side, the vertical extent of the atmosphere changes along the ray. As shown in Sect. 6, this causes a net increase

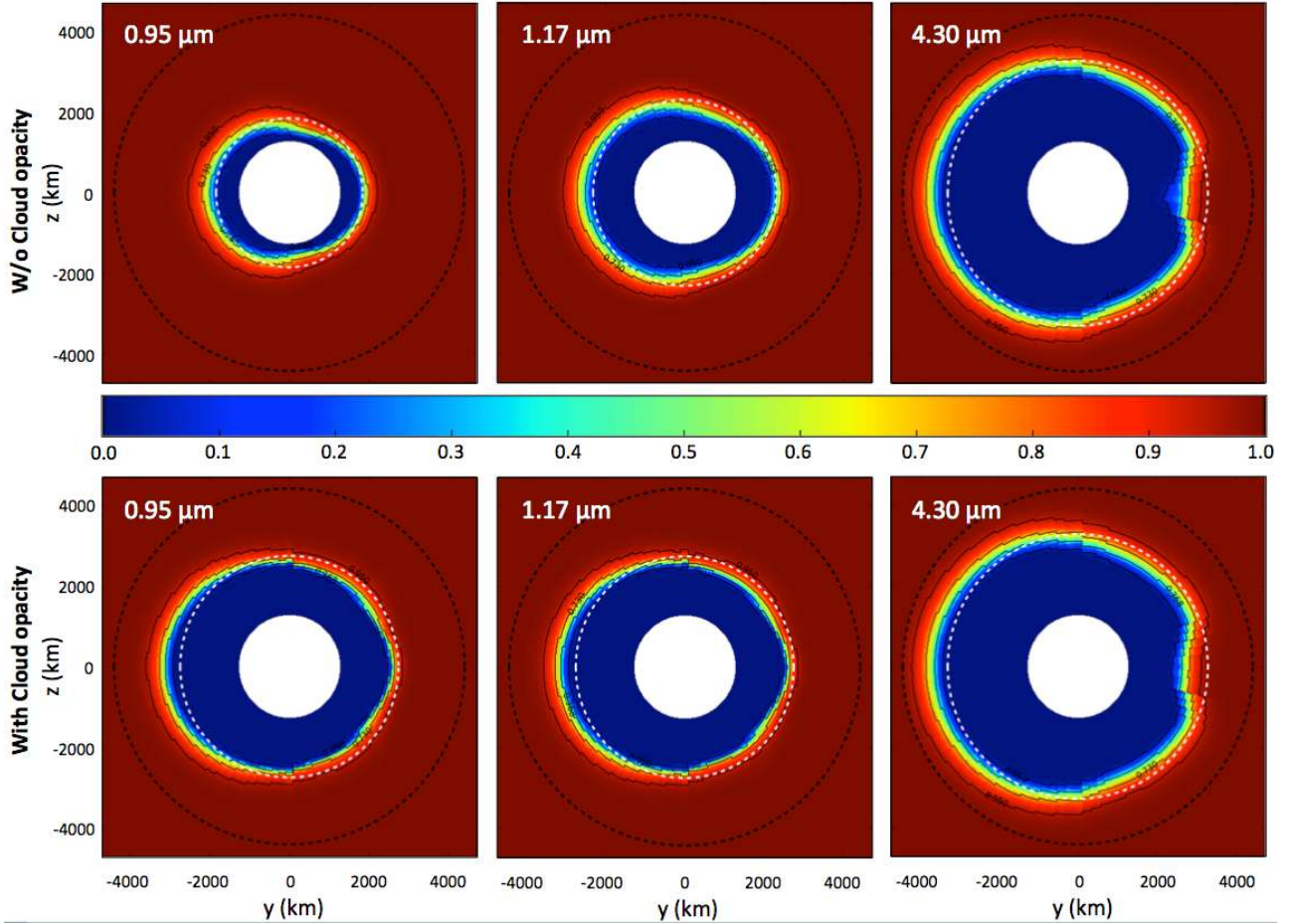


Fig. 10. Transmittance maps obtained with the simulation of the atmosphere of GJ 1214 b depicted in Fig. 8. The region shown between the inner white disc and the outer black dotted circle is an enlargement of the region delimited by white circles in Fig. 8. The white dotted circle corresponds to the effective radius: the radius of the opaque disc resulting in the same overall absorption as the simulation over a homogeneous stellar disc. The maps in the upper row are obtained with Mie scattering turned off in order to show the effect of clouds. Clouds dominate near $0.95 \mu\text{m}$, methane is the main gaseous absorber at $1.17 \mu\text{m}$, and carbon dioxide at $4.30 \mu\text{m}$.

in absorption visible in the water bands (see lower panels of Fig. 12). Although this effect is on the order of the photon noise for a single transit for the (relatively) cold atmosphere of GJ1214 b, it can strongly affect the retrieval of the properties of hotter planets as demonstrated in the following section.

- Day–night compositional gradient: for the absorption bands due to absorbers with a heterogeneous distribution (like CO_2 in our case, which absorbs prominently at 4.5 and $15 \mu\text{m}$) a change of composition along the line of sight creates a signal that is much greater than the expected noise and that cannot be modelled by the limb integration method. Below, we discuss the possible causes of such a compositional gradient. Although it is the most prominent effect in our GJ 1214 b model, the parameter space to be covered in order to fully quantify it is large and is deferred to a future study.
- Day–night asymmetry of the cloud distribution: the temperature change at the terminator, in turn, allows us to expect changes in the properties of the clouds there (Lee et al. 2016). This will also be looked at in a future study.

5.4. Comments on the possible causes of compositional heterogeneities

In the model of GJ 1214 b we use, the concentrations of CO , CO_2 , and CH_4 are computed assuming chemical equilibrium. For such

atmospheres and cooler ones, the 3D variations of these species are therefore strongly overestimated. While the hottest regions of the atmosphere (above ~ 1000 K) are expected to reach equilibrium faster than typical dynamic timescales, this is not the case in the coldest layers (below ~ 700 K) probed by transmission, where endothermic reaction becomes extremely slow. As a consequence, thermochemistry is expected to produce a more homogeneous composition controlled by the hottest and deepest regions.

However, more irradiated planets must have hotter atmospheres overall, spanning a larger range of temperatures (due to shorter radiative timescales) with very different equilibrium compositions and kinetics fast enough to maintain local equilibrium (Agúndez et al. 2014). These atmospheres are expected to exhibit the strongest horizontal variations of temperature and composition. Parmentier et al. (2018) recently showed that water itself may be dissociated on the dayside of some hot Jupiters and recombine close to their terminator. In addition, UV-driven photochemistry may create additional heterogeneities by allowing some reactions to take place on the dayside of the planet only.

We can thus expect hot planets to exhibit strong day–night compositional gradients that may become a dominant issue in retrieving atmospheric properties through transmission spectroscopy. Quantifying the stellar irradiation at which these effects become significant will require further modelling.

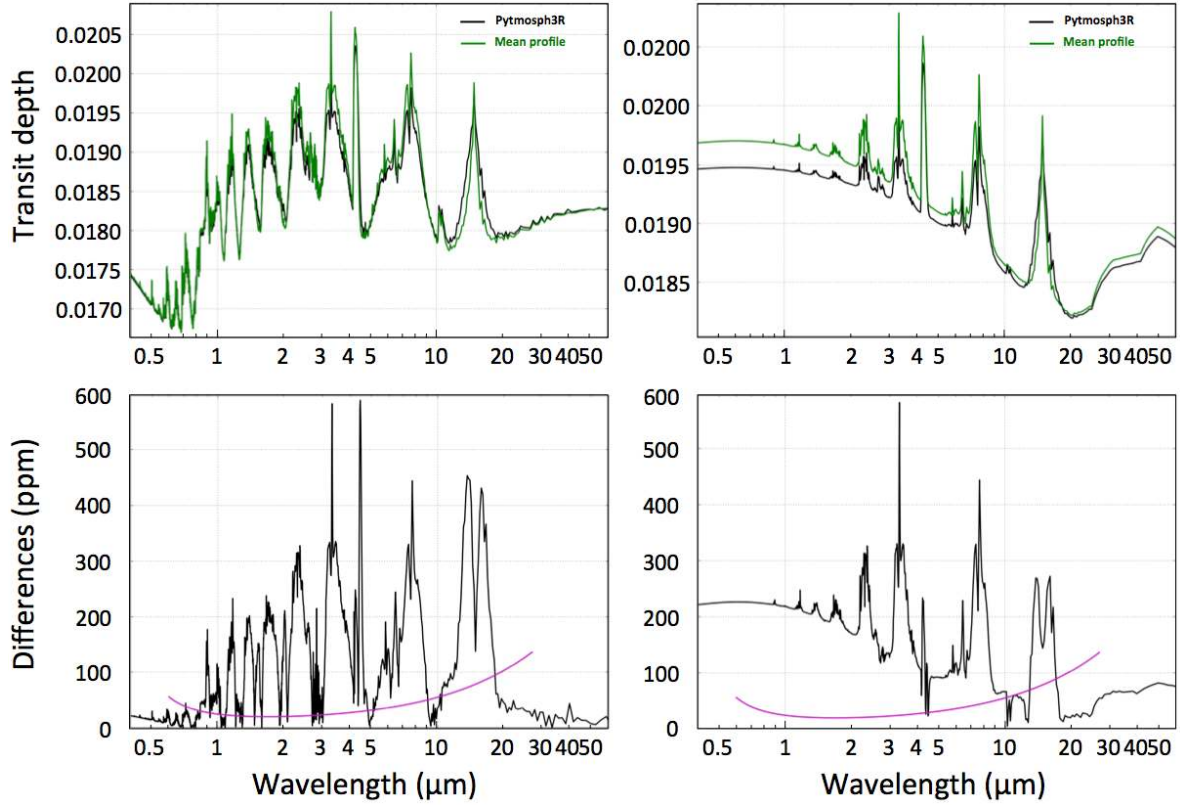


Fig. 11. Transit spectra computed from a 3D simulation of GJ 1214 b atmosphere by Charnay et al. (2015) with Pytmosph3R (black) and a mean profile approximation (see text). Spectra are shown with (*right panels*) and without (*left panels*) the radiative effects from the KCl and ZnS clouds. In both cases the atmosphere is extrapolated vertically beyond the top of the simulation. Cloud particle radius is fixed to $0.5 \mu\text{m}$. The plots on the lower part show the difference between the two methods. The purple line indicates the photon noise with a JWST aperture, an exposure time of twice the transit duration and a (low) resolution of $R = 100$.

6. Effect of dayside–nightside temperature differences on retrieval

As visible in Fig. 2, the region probed in primary transit is much larger than is usually acknowledged, especially on hot and/or low-gravity objects. Therefore, during transit, we are not only probing a thin plane – the so-called terminator – but an area that can extend significantly on both the day and night sides. Because these two parts of the planets are expected to exhibit quite different temperatures (as visible in Fig. 7), it seems important to quantify the extent of the imprint of this temperature inhomogeneity on the transit spectrum of the planet, and how it will affect any attempt to retrieve the temperature at the terminator.

To answer these questions, we conduct a simple experiment. For two prototypical planets (based on GJ 1214 b and HD 209458 b), we build idealised 3D atmospheric structures that are symmetric about the star–planet axis but that continuously go from a high temperature T_{day} on the dayside to a lower temperature, T_{night} , on the nightside. The atmosphere is assumed to have a uniform composition to enable us to concentrate on thermal effects. We then simulate the transit spectrum and try to invert it. Finally, the retrieved temperature is compared to the input one.

6.1. Parametrization of the atmospheric structure

The structure that we chose for the atmosphere is inspired by the 3D simulations of Charnay et al. (2015) for GJ 1214 b whose temperature distribution is shown in Fig. 7. Its most salient feature is the continuous transition in temperature between the day and

night sides. For simplicity, we assume that this transition occurs linearly over a region that is parametrized by its opening angle – hereafter referred to as the transition angle (β).

Moreover, as is well known and further exemplified by Fig. 7, the atmosphere of gaseous exoplanets is usually well mixed at depth. The temperature is therefore assumed to be uniform below a pressure level P_{iso} (taken to be 10 mb as in the simulation, although some models predict inhomogeneities to persist at deeper levels).

In summary, for a given location identified by its pressure, P , and the (possibly negative) local solar elevation angle, α_* , the temperature is given by

$$\begin{cases} P > P_{\text{iso}} & \begin{cases} 2\alpha_* > \beta & T = T_{\text{day}} \\ \beta > 2\alpha_* > -\beta & T = T_{\text{day}} \\ -\beta > 2\alpha_* & T = T_{\text{night}} + (T_{\text{day}} - T_{\text{night}}) \frac{\alpha_* + \beta/2}{\beta} \\ & T = T_{\text{night}} \end{cases} \end{cases} \quad (22)$$

Examples of such idealized atmospheric structures are shown in Fig. 13 for three different day–night transition angles. We chose to show the structures that are most representative of the real GJ 1214 b case to allow for a direct comparison with Fig. 7. For brevity we also refer hereafter to the “uniform case”: this refers to a case where the whole upper atmosphere has a uniform temperature equal to that at the terminator that serves as a comparison. Therefore, in this case,

$$\begin{cases} P > P_{\text{iso}} & T = T_{\text{day}} \\ P < P_{\text{iso}} & T = (T_{\text{day}} + T_{\text{night}})/2 \end{cases} \quad (23)$$

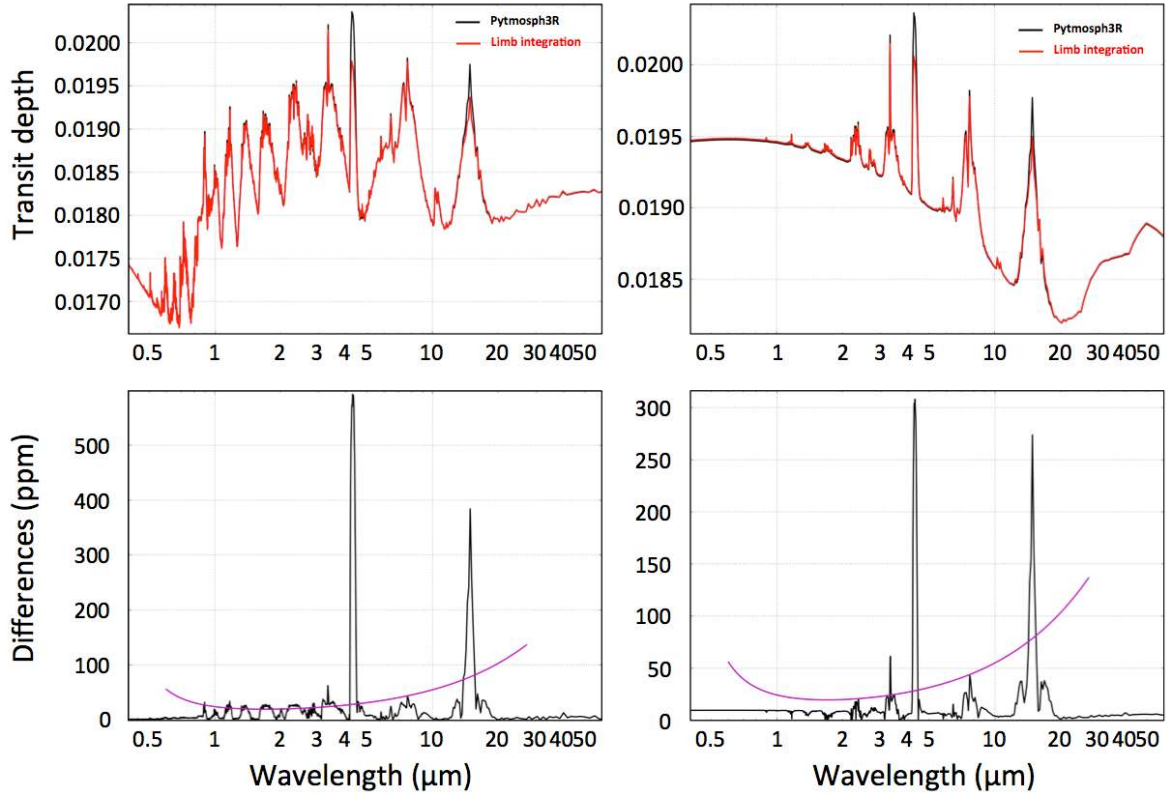


Fig. 12. As in Fig. 11 but the Pytmosph3R spectra (black) are compared here with those obtained with the limb integration approximation (red).

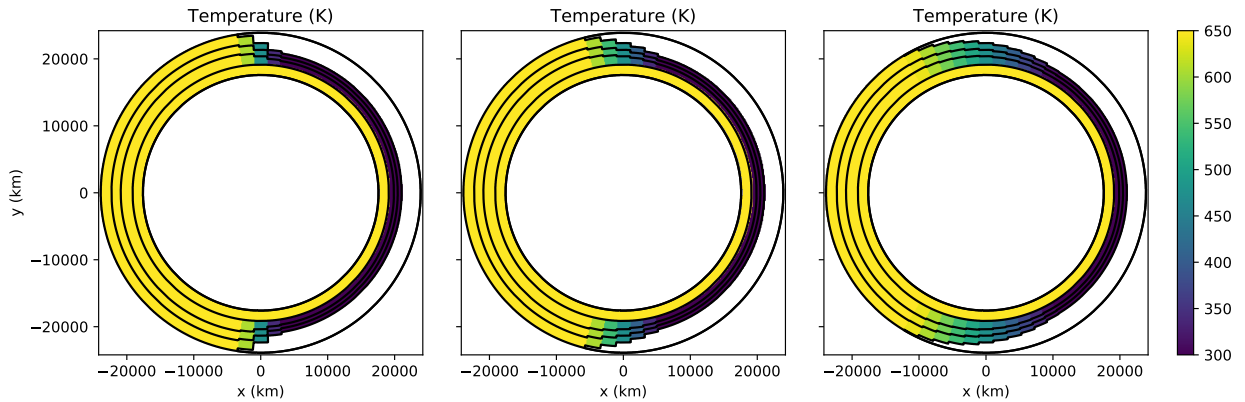


Fig. 13. Colour maps showing the temperature distribution of our model atmospheres in a plane perpendicular to the terminator for three different transition angles between a 650 K dayside and a 300 K nightside (*from left to right*, the day–night transition angle is $\beta = 15, 30,$ and 60°). The inner white circle represents the inner part of the planet with a radius assumed to be equal to that of GJ 1214 b (17,600 km). The star is on the left and the observer on the right on the $y=0$ line. From the centre outward, the five solid lines are respectively the $10^6, 10^3, 1, 10^{-2},$ and 10^{-4} Pa pressure levels. Below the 10^3 Pa level, the atmosphere is assumed to efficiently redistribute heat and is horizontally isothermal. These maps are to scale and show that the dayside is noticeably more extended than the nightside.

Once our four parameters – $T_{\text{day}}, T_{\text{night}}, \beta,$ and P_{iso} – have been chosen, the 3D atmospheric structure is integrated from the surface of the planet that is assumed to be the 10 bar level. Hereafter, we refer to the radius of this isobar as R_p , and assume that it contains most of the mass of the object, meaning that the gravity above this level only depends on the altitude.

The reason we are performing such a test on idealised structures and not more realistic ones from a GCM is that we want to isolate the effect of the day–night heterogeneities. In a GCM simulation, there would also be vertical and equator–pole temperature variations that would preclude the identification of a single effect.

6.2. Effect of the day–night temperature difference on the transmission spectrum

One might naively expect that because of the symmetry of our temperature distribution, the contributions of the hot and cold sides should cancel out. This is however not the case, as can be seen in Fig. 14.

Indeed, by comparing the transmittance map for a given planet in a uniform case, or with a day–night temperature gradient, we directly see that the opaque region extends higher up in the latter case. The greater scale height on the dayside is not compensated by the lower one on the night side.

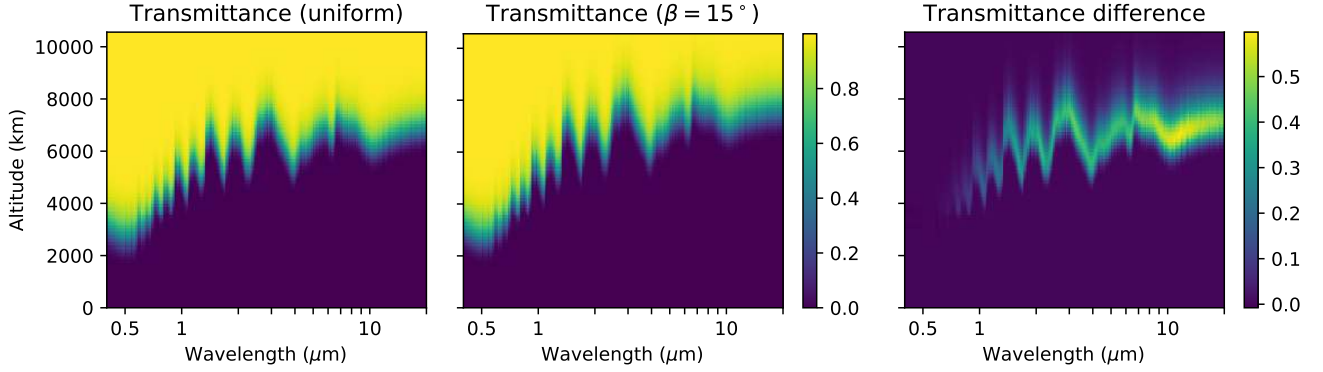


Fig. 14. Maps of the spectral transmittance of clear atmosphere models of HD 209458 b as a function of wavelength and altitude. The parameters are $T_{\text{day}} = 1800$ K, $T_{\text{night}} = 1000$ K, and $P_{\text{iso}} = 10$ mb, which are representative of the real planet (Parmentier et al. 2013). A low transmittance – blue shades – is representative of opaque regions deep down and a transmittance near unity – yellow – is representative of the transparent upper atmosphere. *Left panel:* case with a uniform upper temperature of 1400 K. *Middle panel:* case with a $\beta = 15^\circ$ transition region between the day and night sides. The higher opacity of the middle case is further highlighted by the *right panel* that shows the map of the transmittance difference between the left and the middle cases. The altitude of the top of the isothermal region is ~ 3800 km.

This can be understood easily using a slightly modified version of the analytical model of Guillot (2010) or Vahidinia et al. (2014) where we separate the atmosphere into two hemispheres that differ only through their temperatures – and thus atmospheric scale heights (H_{day} and H_{night}) – above the pressure level P_{iso} which is located at an altitude z_{iso} above the reference radius of the planet R_p . Below z_{iso} , the scale height is the same everywhere (H_{iso}). In essence, this corresponds to the case described above in the limit where $\beta \rightarrow 0$.

We follow the notations in Fig. 1 and the formalism in Appendix A. The difference here is that the scale-height variations mean that the number density at a given altitude is

$$n(z) = \begin{cases} n_0 e^{-\frac{z}{H_{\text{iso}}}} & z < z_{\text{iso}} \\ n_0 e^{-\frac{z_{\text{iso}}}{H_{\text{iso}}}} e^{-\frac{z-z_{\text{iso}}}{H_i}} & z > z_{\text{iso}} \end{cases}, \quad (24)$$

where i is either day or night depending on the hemisphere.

In the limit where all the altitudes in the atmosphere are small compared to R_p , the slant optical depth is the sum of the dayside and nightside contributions, yielding

$$\begin{aligned} \tau_{\text{tr}}|_{z_t > z_{\text{iso}}} &= \int_{-\infty}^0 n \sigma_{\text{mol}} dx + \int_0^{\infty} n \sigma_{\text{mol}} dx \\ &= \sigma_{\text{mol}} n_0 e^{-\frac{z_{\text{iso}}}{H_{\text{iso}}}} \left(\sum_{i=\text{day}}^{\text{night}} e^{-\frac{z_t - z_{\text{iso}}}{H_i}} \int_0^{\infty} e^{-\frac{x^2}{2R_p H_i}} dx \right) \\ &= \sigma_{\text{mol}} n_0 e^{-\frac{z_{\text{iso}}}{H_{\text{iso}}}} \left(e^{-\frac{z_t - z_{\text{iso}}}{H_{\text{day}}}} \sqrt{\frac{\pi R_p H_{\text{day}}}{2}} + e^{-\frac{z_t - z_{\text{iso}}}{H_{\text{night}}}} \sqrt{\frac{\pi R_p H_{\text{night}}}{2}} \right), \end{aligned} \quad (25)$$

if $z_t > z_{\text{iso}}$, and

$$\begin{aligned} \tau_{\text{tr}}|_{z_t < z_{\text{iso}}} &= \sigma_{\text{mol}} n_0 \left(\int_{-\sqrt{2(z_{\text{iso}} - z_t)R_p}}^{\sqrt{2(z_{\text{iso}} - z_t)R_p}} e^{-\frac{z_t}{H_{\text{iso}}}} e^{-\frac{x^2}{2R_p H_{\text{iso}}}} dx \right. \\ &\quad \left. + e^{-\frac{z_{\text{iso}}}{H_{\text{iso}}}} \sum_{i=\text{day}}^{\text{night}} \int_{\sqrt{2(z_{\text{iso}} - z_t)R_p}}^{\infty} e^{-\frac{z_t - z_{\text{iso}}}{H_i}} e^{-\frac{x^2}{2R_p H_i}} dx \right) \end{aligned}$$

$$\begin{aligned} &= \sigma_{\text{mol}} n_0 \left(e^{-\frac{z_t}{H_{\text{iso}}}} \sqrt{2\pi R_p H_{\text{iso}}} \operatorname{erf} \left[\sqrt{\frac{z_{\text{iso}} - z_t}{H_{\text{iso}}}} \right] \right. \\ &\quad \left. + e^{-\frac{z_{\text{iso}}}{H_{\text{iso}}}} \sum_{i=\text{day}}^{\text{night}} e^{-\frac{z_t - z_{\text{iso}}}{H_i}} \sqrt{\frac{\pi R_p H_i}{2}} \left(1 - \operatorname{erf} \left[\sqrt{\frac{z_{\text{iso}} - z_t}{H_i}} \right] \right) \right), \end{aligned} \quad (26)$$

if $z_t < z_{\text{iso}}$. We note that in the above equations, σ_{mol} is the mean cross section of the gas normalized to the total density.

Now, to see the increase in optical depth, let us divide the result above by the optical depth in the uniform case (τ_{uni}) given by Eq. (A.3). This yields

$$\frac{\tau_{\text{tr}}|_{z_t > z_{\text{iso}}}}{\tau_{\text{uni}}} = \frac{1}{2} \left(\sqrt{1 + \Delta H} e^{\frac{\Delta H \Delta z}{1 + \Delta H}} + \sqrt{1 - \Delta H} e^{-\frac{\Delta H \Delta z}{1 - \Delta H}} \right), \quad (27)$$

where $\Delta H \equiv (H_{\text{day}} - H_{\text{night}})/2$, meaning that $(H_{\text{day}}, H_{\text{night}}) = H_{\text{iso}}(1 \pm \Delta H)$, and $\Delta z = (z_t - z_{\text{iso}})/H_{\text{iso}}$. The expansion in ΔH gives

$$\frac{\tau_{\text{tr}}|_{z_t > z_{\text{iso}}}}{\tau_{\text{uni}}} \approx 1 + \frac{1}{8} \left(-1 - 4\Delta z + 4\Delta z^2 \right) \Delta H^2 + O(\Delta H^4). \quad (28)$$

Firstly, we see that because of the symmetry of the setup, the first-order term disappears. Secondly, this readily results in a larger optical depth in the heterogeneous case than in the uniform case for all the rays with a tangent altitude that is $(1 + \sqrt{2})/2 \approx 1.2$ scale heights greater than the altitude of the isothermal region. This qualitatively explains why there is little difference in the transmittance below the altitude of the isothermal region (~ 3800 km) in Fig. 14.

Finally, when these transmittance maps are integrated vertically, we get the transmission spectra shown in Fig. 15. As expected, the effective altitude at which the atmosphere becomes opaque systematically increases when the horizontal thermal gradient is increased near the terminator.

An important point is that the spectrum for the non-uniform case – although different from the spectrum that is obtained for a uniform atmosphere with the temperature of the terminator – can be similar to the spectrum that would be obtained for a uniform but globally hotter atmosphere. It is therefore not surprising that a retrieval algorithm would be biased and would retrieve a hotter atmosphere.

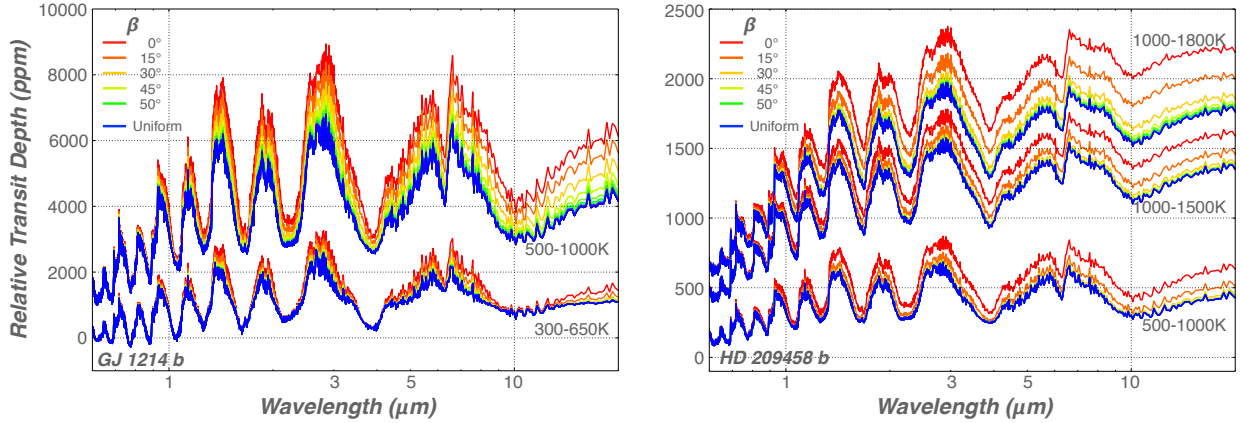


Fig. 15. Spatially integrated spectra of the relative transit depth (in ppm) expected for GJ 1214 b (*left panel*) and HD 209458 b (*right panel*) as a function of the day–night transition angle (β). The colour goes from red to green when β goes from 0 to 50° every 15°. The blue curve is the reference uniform case. Spectra with larger β are undistinguishable from the uniform case. The different labelled groups of curves are for our various sets of day–night temperatures. For each ($T_{\text{day}} - T_{\text{night}}$) group, the temperature at the terminator of all the models – including the uniform one – is $(T_{\text{day}} + T_{\text{night}})/2$. Transit depth can be seen to increase monotonically when β decreases.

6.3. Bias in retrieved temperatures

The last step of our analysis is to actually run a 1D retrieval procedure on the transmittance spectra obtained with our parametrized 3D atmospheric structures to assess the biases entailed by such an approach. To do so, we use the JWST simulation tool PandExo (Batalha et al. 2017) to simulate the expected uncertainties over a wavelength range of 1.0–10 μm for the duration of a single transit. For this, we combined the simulated spectra of the NIRISS/SOSS, NIRSpec/G395M, and MIRI/LRS instruments. Observed uncertainties above 10 μm were too large to have a significant impact on retrieval results and were discarded. Finally, we binned the simulated observations to a resolution of 100 constant in wavelength. Given that we are investigating biases due to the model rather than observational noise, we here only consider the uncertainties calculated by PandExo but do not add additional noise to the mean of our simulated observations. Significant biases in the posterior distributions of retrieved parameters can occur when considering single random noise instances of the data. To correctly alleviate such noise-induced biases, one would need to combine posterior distributions of multiple noise-instance retrievals. Fortunately, Sect. 5.2 of Feng et al. (2018) shows that the combination of multiple noise-instance retrievals converges to the noise-free retrieval posterior distributions, as expected from the central limit theorem. We did not include non-Gaussian noise due to instrument or stellar systematics, as these are not currently known or are data-set dependent, or both. We therefore note that the retrieved parameter uncertainties presented here are theoretical lower limits.

For each model of our grid in temperature and day–night transition angle, we ran the spectrum through the TauREx retrieval software (Waldmann et al. 2015). Here we considered water as the only trace gas with absorption cross sections computed using the Barber et al. (2006) line list. We include Rayleigh scattering and collision-induced absorption of H₂–H₂ and H₂–He (Borysow et al. 2001; Borysow 2002; Rothman et al. 2013) and assumed the atmosphere to be cloud-free. The vertical temperature–pressure profile was modelled to be isothermal. A typical posterior distribution for the retrieved parameters resulting from this procedure is shown in Fig. 16, along with the simulated input spectrum and the fitted one.

The retrieved temperatures and water abundances as a function of the transition angle (β) for our GJ 1214 b case are shown in Fig. 17. Figure 18 shows the temperature result for HD 209458 b. We tested different dayside and nightside temperatures to see how planets with various irradiances would behave. To be able to compare the retrieved temperature (T_{ret}) in these different cases, we use the relative retrieved temperature

$$\Theta \equiv \frac{T_{\text{ret}} - T_{\text{night}}}{T_{\text{day}} - T_{\text{night}}}, \quad (29)$$

which should therefore be equal to 0.5 if we were to retrieve the terminator temperature. Although there are some quantitative differences among these cases, some robust trends emerge:

- The retrieved temperature is systematically biased towards a higher temperature than that of the terminator ($\Theta \geq 0.5$).
- There are two regimes separated by a critical day–night transition angle that depends on the characteristics of the planet and is roughly consistent with our estimate of the opening angle of the part of the planet that is probed in transit, that is, the limb (ψ , denoted by a vertical dashed line; see Fig. 2).
- For $\beta < \psi$, the retrieved temperature decreases roughly linearly with the transition angle. As the transition angle between the dayside and nightside goes to zero (very sharp transition expected for the hottest planets), the retrieved temperature approaches the dayside temperature.
- For $\beta > \psi$, the temperature structure within the limb, hence the retrieved temperature, does not vary significantly. Whether the actual retrieved temperature is equal to the temperature at the terminator depends on the case in question (see below).
- Despite the uniform composition in our models, the retrieved abundance is always significantly biased – in the sense that the real abundance is outside the formal error bars of the retrieval – although the magnitude and direction depends on the specific case. It is sensible to assume that other more-complex biases will arise if chemical gradients are present as well.

If in the HD 209458 b case (see Fig. 18), the retrieved temperature converges toward the temperature at the terminator when the latter becomes more uniform ($\beta \rightarrow 180^\circ$), this is not necessarily the case for GJ 1214 b. We find that this absence of convergence at large angles always occurs when the hot, deep

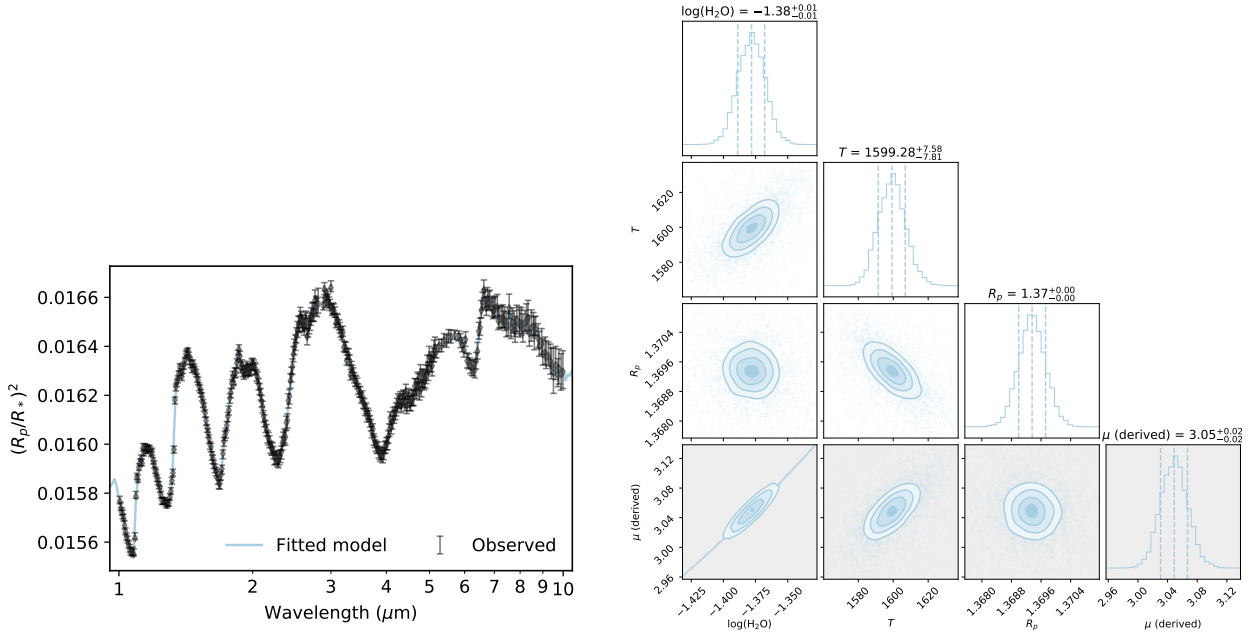


Fig. 16. Typical result of the retrieval procedure. This specific case is HD 209458 b with $\beta = 15^\circ$ and temperatures of 1800 and 1000 K on the dayside and nightside, respectively. *Left panel:* best-fit 1D spectrum (blue curve) along with the spectrum produced by our 3D tool used as input for the retrieval (black points with error bars being the 1σ uncertainty computed with PandExo). *Right panel:* posterior distribution for the retrieved parameters. This shows that the retrieval finds an acceptable fit, which results in relatively peaked posterior distribution and small error bars on the retrieved parameters. These values are however biased as the actual terminator temperature (1400 K) and atmospheric water abundance ($\log_{10}[\text{H}_2\text{O}] = -1.3$) are outside the range of values shown.

atmosphere below the $P_{1\text{so}}$ level is probed by the transit spectrum: the retrieval is biased by the vertical temperature gradient. This does not happen for our Hot Jupiter case because of the larger radius that pushes the transit photosphere at lower pressures. Although an important bias in itself, it has already been studied by Rocchetto et al. (2016), and will not be further discussed here.

6.4. Could we see that something is wrong?

Could an observer, having performed the retrieval, detect that the retrieved quantities are biased by the day–night temperature gradient? This is indeed a crucial point.

Unfortunately, this seems precluded, even with the exquisite precision of JWST. As can be seen in Fig. 16, the best-fit 1D, isothermal spectrum does not miss any feature of the input 3D spectrum. In fact, Fig. 19 shows that the reduced χ^2 of the optimal retrieved models are always near or below unity³. Based on this metric, the 1D isothermal atmosphere model therefore provides an acceptable fit to the data, at least in the low-resolution mode that we have explored here. In fact, it even seems that the fit is better when the bias is the strongest (low day–night transition angle). This counter intuitive result comes from the fact that when the temperature transition is sharp, we probe almost exclusively the dayside, and the atmosphere therefore appears more homogeneous.

A procedure that would use a radiative-transfer code similar to Pythmosph3R to retrieve a 3D structure/composition (which would imply formidable computing resources) would admittedly

³ Finding reduced χ^2 smaller than unity is allowed by the fact that, although the noise-induced uncertainty on the input spectrum is used as an input of the retrieval procedure, no noise instance has been added to the input spectrum itself (see Sect. 6.3). Contrary to the case of a comparison with real data, low χ^2 here are not a sign of over-fitting.

reveal the issue as its posterior distributions would expose the full extent of the degeneracies and result in larger, more reliable error bars on the retrieved quantities. Nevertheless, such a sophisticated tool may not be able to achieve a better retrieval. Indeed, even if a better χ^2 may be found with a forward model using a 3D thermal and compositional structure, the 1D model already provides an excellent match and the improvement, if any, would be achieved at the expense of adding so many parameters that parsimony criteria may favour the most simple model.

However, a much higher resolution may change this state of affairs, especially if we start to be sensitive to the line shape of individual lines. It is also possible that an east–west asymmetry of the terminator could add some signal that would be distinguishable from any 1D profile. This will have to be assessed in a future study.

7. Which atmospheres are affected?

Since the first detections of the thermal emission of a planet, the existence of a strong day–night temperature gradient on hot extrasolar planets has been well established (Cowan & Agol 2011). A clear trend has even emerged that the hotter the planet, the greater the thermal contrast (Komacek et al. 2017; Keating & Cowan 2018) – a thermal contrast that can reach more than a thousand degrees. The bias on the retrieved limb temperatures on real planets is thus potentially huge. The rather observationally unconstrained parameter, however, is the width of the day–night transition region (here β) and how it compares to the width of the limb that is effectively probed in transmission (ψ). Since phase-curve observations do not yet have the resolution necessary to precisely measure the width of the transition, we turn to published models.

For Neptune-like planets such as GJ 1214 b, our predictions based on a GCM tell us that the thermal-only effect of the

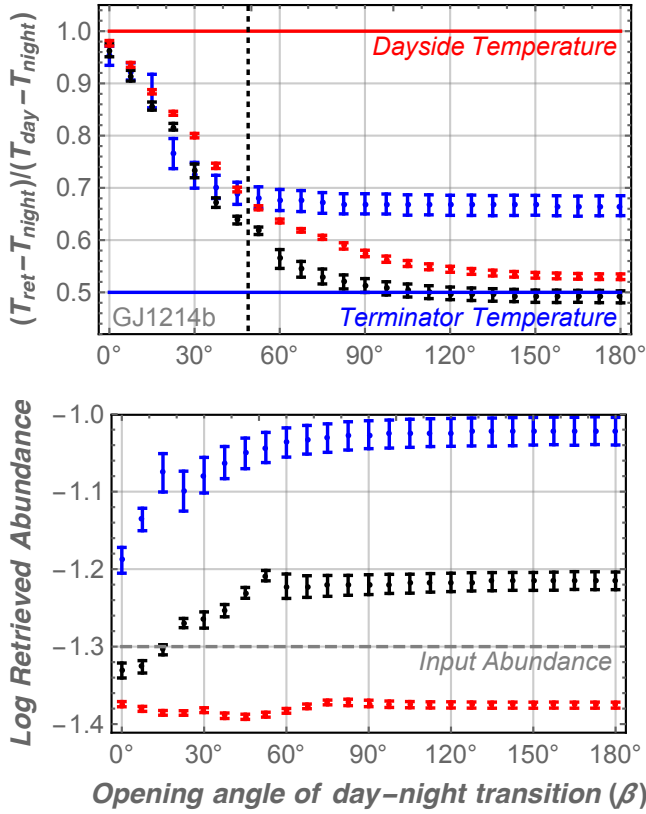


Fig. 17. Retrieved temperature (*top panel*) and water abundance (*bottom panel*) as a function of the transition angle between the day and nightside of a planet with the radius and surface gravity of GJ 1214 b. The colours correspond to the temperature range (blue: 300–650 K; black: 500–1000 K; red: 1000–1500 K). To put all cases on a single diagram we show the relative retrieved temperature, $\Theta \equiv (T_{\text{ret}} - T_{\text{night}})/(T_{\text{day}} - T_{\text{night}})$, so that in all cases $\Theta = 0.5$ at the terminator. The retrieved temperature is systematically biased toward the dayside temperature ($\Theta \geq 0.5$), especially when the transition is sharp. For comparison the estimate of the width of the limb given by Eq. (3) is shown by the vertical dotted line.

day–night temperature gradient is on the order of 50–100 ppm in the water bands, and clearly detectable by JWST (see Fig. 12). It is thus sensible to assume that any hotter Neptune-like planets should be increasingly affected because of their higher H/R_p ratio and their stronger thermal day–night contrast (see Fig. 15). On the contrary, the stratosphere of colder planets being more uniform, how cold a planet needs to be before such effects are undetectable remains to be elucidated.

However, we reiterate that only the direct thermal effect is discussed here, leaving out possible chemical heterogeneities. The strong signature of day–night gradients in the CO_2 and CH_4 bands of the spectrum in Fig. 12, which can reach 600–1000 ppm, are believed to be of even greater importance and could possibly strongly hamper the ability of conventional retrieval algorithms to retrieve meaningful molecular abundances in the case of heterogeneous atmospheres.

Such considerations are even more important for hot Jupiters. Despite their higher gravity, they can be much hotter, which means that the day–night contrasts are expected to be much stronger, both thermally and compositionally because chemical timescales are expected to be short compared to the advection timescales. A good example is the so-called ultra-hot Jupiters, where it is predicted that some very abundant molecules on the

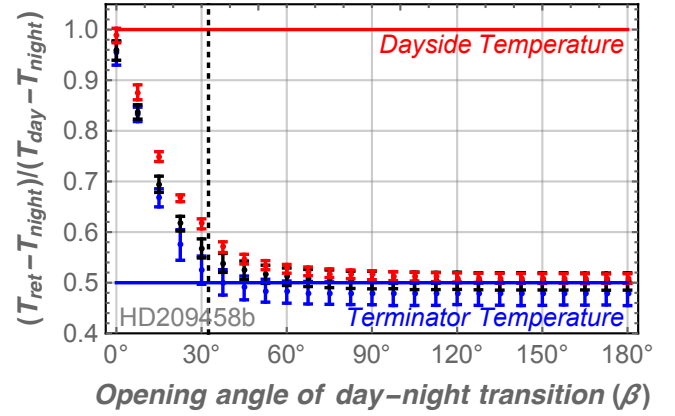


Fig. 18. Dimensionless retrieved temperature as a function of the transition angle between the dayside and nightside of a planet with the radius and surface gravity of HD 209458 b. The colours correspond to the temperature range (blue: 500–1000 K; black: 1000–1500 K; red: 1000–1800 K). See Fig. 17 for other details.

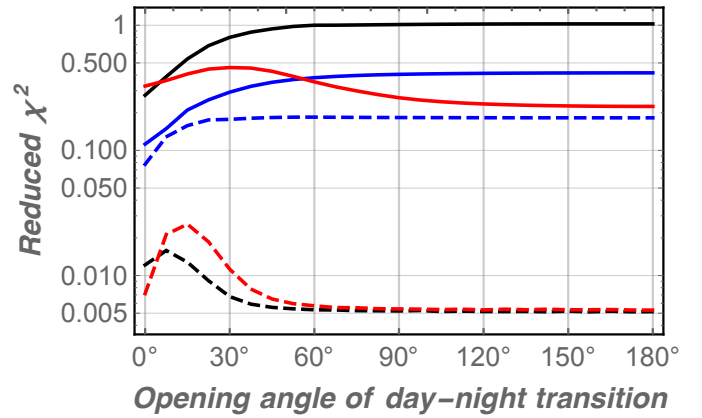


Fig. 19. Reduced χ^2 for the fit of the optimal retrieved models from Figs. 17 and 18. The colour-coding is the same as for those figures. Solid curves are for GJ 1214 b and dashed are for HD 209458 b. As the reduced χ^2 is always close to or smaller than 1, the fit would always be considered satisfactory.

nightside, such as water, could be almost completely absent on the dayside due to thermal dissociation (Parmentier et al. 2018).

8. Conclusion

Overall, our most important conclusion is that the region of the atmosphere probed by transit spectroscopy, that is, the limb, is not confined to a narrow annulus around the planet as often implied, but can indeed extend relatively far throughout its two hemispheres. This is especially true for hot and/or low-gravity objects, the most significant metric being the ratio of the atmospheric scale height to the radius of the planet.

As a result, in addition to the variations of atmospheric properties of the atmosphere along the terminator, the transit spectrum is also affected by their variations across the limb, i.e. along the path of the light rays.

To investigate all these effects, we have developed Pytmosph3R, a transit spectrum generator that can work with a 3D atmospheric structure, whether it is the output of a global circulation model or a more idealised one. Using this tool along with a 3D atmospheric model of GJ 1214 b, we have recovered

previous results demonstrating that the temperature and compositional variations along the terminator significantly affect the transit spectrum and will have to be accounted for in future studies. These effects can in principle be partially accounted for by using a (1+1)D, or limb-integrated, approach where one 1D spectrum is generated for each part of the limb before it is weighted and added to the others to generate the global one.

However, our fully 3D framework has shown that at the precision that will be reached by future observatories, the limb integrated approach is insufficient. Indeed, we have shown that for temperature gradients realistically expected for observable exoplanets, the transit spectrum is significantly affected by the structure of the atmosphere across the limb, that is, the thermal and compositional gradients between the dayside and nightside of the planet. We further demonstrated that this effect systematically biases 1D retrieval methods towards the temperature of the dayside. The extent of this bias, however, depends on the strength of the temperature contrast as well as its sharpness around the terminator, the latter being the most difficult to predict.

In other words, one should be aware of the fact that the temperature (or its profile) retrieved from transmission spectra may not apply to the terminator itself, and that temperatures at the terminator are in fact significantly smaller. This will of course be a routine problem for future high-precision observatories, but we have demonstrated that the effect on the spectrum – which can reach hundreds if not thousands of parts per million in some cases (see Fig. 15) – is well above the precision of 50–100 ppm that has been achieved with HST and Spitzer (Cowan et al. 2015). Current observations of some Hot Jupiter are therefore possibly already affected, but to what extent remains to be elucidated.

Acknowledgements. This project has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement no. 679030/WHIPLASH, no. 758892/ExoAI, no. 776403/Exoplanets A and under the European Union’s Seventh Framework Programme (FP7/2007-2013)/ERC grant agreement no. 617119/Exo-Lights. I.P.W. furthermore acknowledge funding by the Science and Technology Funding Council (STFC) grants: ST/K502406/1 and ST/P000282/1).

References

- Agúndez, M., Parmentier, V., Venot, O., Hersant, F., & Selsis, F. 2014, *A&A*, **564**, A73
- Barber, R. J., Tennyson, J., Harris, G. J., & Tolchenov, R. N. 2006, *MNRAS*, **368**, 1087
- Batalha, N. E., Mandell, A., Pontoppidan, K., et al. 2017, *PASP*, **129**, 064501
- Bates, R., Zarana, K. J., Tolbert, M. A., & Volkamer, R. 1984, *Planet. Space Sci.*, **32**, 785
- Beichman, C., Benneke, B., Knutson, H., et al. 2014, *PASP*, **126**, 1134
- Blecic, J., Dobbs-Dixon, I., & Greene, T. 2017, *ApJ*, **848**, 127
- Borysow, A. 2002, *A&A*, **390**, 779
- Borysow, A., Jorgensen, U. G., & Fu, Y. 2001, *J. Quant. Spectr. Rad. Transf.*, **68**, 235
- Burrows, A., Rauscher, E., Spiegel, D. S., & Menou, K. 2010, *ApJ*, **719**, 341
- Charnay, B., Meadows, V., Misra, A., Leconte, J., & Arney, G. 2015, *ApJ*, **813**, L1
- Chiavassa, A., Caldas, A., Selsis, F., et al. 2017, *A&A*, **597**, A94
- Clough, S., Kneizys, F., & Davies, R. 1989, *Atm. Res.*, **23**, 229
- Cowan, N. B., & Agol, E. 2011, *ApJ*, **729**, 54
- Cowan, N. B., Greene, T., Angerhausen, D., et al. 2015, *PASP*, **127**, 311
- Dalgarno, A., & Williams, D. A. 1962, *ApJ*, **136**, 690
- Dobbs-Dixon, I., Agol, E., & Burrows, A. 2012, *ApJ*, **751**, 87
- Evans, T. M., Sing, D. K., Kataria, T., et al. 2017, *Nature*, **548**, 58
- Feng, Y. K., Line, M. R., Fortney, J. J., et al. 2016, *ApJ*, **829**, 52
- Feng, Y. K., Robinson, T. D., Fortney, J. J., et al. 2018, *AJ*, **155**, 200
- Fortney, J. J., Shabram, M., Showman, A. P., et al. 2010, *ApJ*, **709**, 1396
- Fu, Q., & Liou, K. N. 1992, *J. Atm. Sci.*, **49**, 2139
- Guillot, T. 2010, *A&A*, **520**, A27
- Hill, C., Yurchenko, S. N., & Tennyson, J. 2013, *Icarus*, **226**, 1673
- Keating, D., & Cowan, N. B. 2018, ArXiv e-prints [arXiv:1809.00002]
- Komacek, T. D., Showman, A. P., & Tan, X. 2017, *ApJ*, **835**, 198
- Kreidberg, L., Bean, J. L., Désert, J.-M., et al. 2014, *Nature*, **505**, 69
- Leconte, J., Forget, F., Charnay, B., Wordsworth, R., & Pottier, A. 2013, *Nature*, **504**, 268
- Lee, G., Dobbs-Dixon, I., Helling, C., Bognar, K., & Woitke, P. 2016, *A&A*, **594**, A48
- Line, M. R., & Parmentier, V. 2016, *ApJ*, **820**, 78
- Lines, S., Manners, J., Mayne, N. J., et al. 2018, *MNRAS*, **481**, 194
- Lodders, K. 2002, *ApJ*, **577**, 974
- Madeleine, J.-B., Forget, F., Millour, E., Montabone, L., & Wolff, M. J. 2011, *J. Geophys. Res. Planets*, **116**, E11010
- Madhusudhan, N. 2018, *Atmospheric Retrieval of Exoplanets* (Cham: Springer Publishing Company), 104
- Miller-Ricci Kempton, E., & Rauscher, E. 2012, *ApJ*, **751**, 117
- Parmentier, V., Showman, A. P., & Lian, Y. 2013, *A&A*, **558**, A91
- Parmentier, V., Line, M. R., Bean, J. L., et al. 2018, *A&A*, **617**, A110
- Richard, C., Gordon, I., Rothman, L., et al. 2012, *J. Quant. Spectr. Rad. Transf.*, **113**, 1276
- Rocchetto, M., Waldmann, I. P., Venot, O., Lagage, P.-O., & Tinetti, G. 2016, *ApJ*, **833**, 120
- Rothman, L. S., Gordon, I. E., Babikov, Y., et al. 2013, *J. Quant. Spectr. Rad. Transf.*, **130**, 4
- Selsis, F., Wordsworth, R. D., & Forget, F. 2011, *A&A*, **532**, A1
- Showman, A. P., Fortney, J. J., Lewis, N. K., & Shabram, M. 2013, *ApJ*, **762**, 24
- Sneep, M., & Ubachs, W. 2004, *J. Quant. Spectr. Rad. Transf.*, **92**, 293
- Thalman, R., Zarzana, K. J., Tolbert, M. A., & Volkamer, R. 2014, *J. Quant. Spectr. Rad. Transf.*, **147**, 171
- Tinetti, G., Drossart, P., Eccleston, P., et al. 2017, *European Planetary Science Congress, 11*, EPSC2017
- Turbet, M., Leconte, J., Selsis, F., et al. 2016, *A&A*, **596**, A112
- Turbet, M., Bolmont, E., Leconte, J., et al. 2018, *A&A*, **612**, A86
- Vahidinia, S., Cuzzi, J. N., Marley, M., & Fortney, J. 2014, *ApJ*, **789**, L11
- Waldmann, I. P., Tinetti, G., Rocchetto, M., et al. 2015, *ApJ*, **802**, 107
- Way, M. J., Aleinov, I., Amundsen, D. S., et al. 2017, *ApJS*, **231**, 12
- Wordsworth, R. D., Forget, F., Selsis, F., et al. 2011, *ApJ*, **733**, L48

Appendix A: Slant optical depth and transit radius of an isothermal, grey atmosphere with constant gravity

A.1. Homogeneous case

We first start by computing the optical depth along a ray crossing the atmosphere of a planet where temperature, gravity, and composition are constant. Here, we follow the notations of [Vahidinia et al. \(2014\)](#), but the reader is referred to [Guillot \(2010\)](#) for a more in depth discussion. Notations are summarised in Fig. 1.

The constant atmospheric scale height, H , entails that the gas number density is given by

$$n = n_0 e^{-z/H}, \quad (\text{A.1})$$

where n_0 is the density at the reference radius (R_p). For a ray with a given tangent altitude z_t , the altitude in the atmosphere at a distance x from the limb plane is given by

$$z = z_t + \frac{x^2}{2R_p}, \quad (\text{A.2})$$

to first order in z_t/R_p . The optical depth along a ray from the star to a given position x due to a given species is therefore

$$\begin{aligned} \tau_{\text{tr}}(z_t) &= \int_{-\infty}^x \sigma_{\text{mol}} \chi n_0 e^{-z_t/H} e^{-x'^2/(2R_p H)} dx' \\ &= \sigma_{\text{mol}} \chi n_0 e^{-z_t/H} \int_{-\infty}^x e^{-x'^2/(2R_p H)} dx' \\ &= \sqrt{2\pi R_p H} \sigma_{\text{mol}} \chi n_0 e^{-z_t/H} \left(\frac{1}{2} + \frac{1}{2} \operatorname{erf} \left(\frac{x}{\sqrt{2R_p H}} \right) \right) \\ &\xrightarrow{x \rightarrow \infty} \sqrt{2\pi R_p H} \sigma_{\text{mol}} \chi n_0 e^{-z_t/H}, \end{aligned} \quad (\text{A.3})$$

where χ is the volume mixing ration of the considered species, and σ_{mol} its cross-section at the wavelength considered. Given that the vertical optical depth is given by

$$\tau_{\perp}(z_t) = H \sigma_{\text{mol}} \chi n_0 e^{-z_t/H}, \quad (\text{A.4})$$

we retrieve the result from [Guillot \(2010\)](#) that

$$\tau_{\text{tr}} = \left(\frac{2\pi R_p}{H} \right)^{\frac{1}{2}} \tau_{\perp}. \quad (\text{A.5})$$

Following Sect. 3.5, to first order, the transit depth is given by

$$\delta = R_{\star}^{-2} \left(R_p^2 + 2 \int_{R_p}^{\infty} (1 - e^{-\tau_{\text{tr}}(\rho)}) \rho d\rho \right), \quad (\text{A.6})$$

where $\rho = R_p + z_t$. Along with Eq. (A.3), Eq. (A.6) is used to validate our model in Sect. 4.2.

A.2. Heterogeneous composition

In this section, we slightly modify the model above to answer the following question: how far from the limb plane can an increase in the abundance of a given species still affect the transit in the relevant bands? To answer this, we assume that the mixing ratio of the considered species is χ_{day} along the line ray for $x < x_{\text{limb}}$ (where x_{limb} is negative if the transition is on the dayside) and χ_{night} beyond that. This is supposed to mimic a situation where an absorber, such as for example TiO, becomes less abundant at the

terminator and on the nightside because it condenses at cooler temperatures. Because of the symmetry of the problem, this also treats the situation where an absorber becomes more abundant on the nightside. Following Fig. 1, x_{limb} is also parametrized by the limb angle $\tan(\psi/2) = x_{\text{limb}}/(R_p + z_t)$.

If the composition were uniform, the optical depth at a given tangent altitude would be

$$\tau_{\text{tr}} = \sqrt{2\pi R_p H} \sigma_{\text{mol}} \chi_{\text{night}} n_0 e^{-z_t/H}. \quad (\text{A.7})$$

Using $\tau_{\text{tr}} \sim 1$ as a criterion for the effective altitude of absorption of the atmosphere, we get an effective altitude for the uniform case that is

$$z_{\text{uni}} = H \ln \left(\sqrt{2\pi R_p H} \sigma_{\text{mol}} \chi_{\text{night}} n_0 \right). \quad (\text{A.8})$$

Now, in the heterogeneous case, using the penultimate line of Eq. (A.3) yields

$$\begin{aligned} \tau_{\text{tr}} &= \sqrt{2\pi R_p H} \sigma_{\text{mol}} n_0 e^{-z_t/H} \\ &\times \left(\frac{\chi_{\text{day}} + \chi_{\text{night}}}{2} + \frac{\chi_{\text{day}} - \chi_{\text{night}}}{2} \operatorname{erf} \left(\frac{x_{\text{limb}}}{\sqrt{2R_p H}} \right) \right), \end{aligned} \quad (\text{A.9})$$

which is equal to the uniform case in the $x_{\text{limb}} \rightarrow -\infty$ limit, as expected. The effective altitude is then

$$\begin{aligned} z_{\text{het}} &= z_{\text{uni}} - H \ln 2 \\ &+ H \ln \left(\left(\frac{\chi_{\text{day}}}{\chi_{\text{night}}} + 1 \right) + \left(\frac{\chi_{\text{day}}}{\chi_{\text{night}}} - 1 \right) \operatorname{erf} \left(\frac{x_{\text{limb}}}{\sqrt{2R_p H}} \right) \right). \end{aligned} \quad (\text{A.10})$$

At what angular distance from the limb plane can we still see such an increase of a given species in the transit spectrum? To answer this question, one needs to quantify the angle ψ at which the resulting change in effective transit altitude of the atmosphere due to the heterogeneity becomes measurable. This writes

$$\delta_{\text{het}} - \delta_{\text{uni}} > \sigma_{\text{obs}} \Leftrightarrow z_{\text{het}}(\psi) - z_{\text{uni}} > \frac{R_{\star}^2}{2R_p} \sigma_{\text{obs}}, \quad (\text{A.11})$$

where σ_{obs} is the relative precision level of the observations. Using the parameters for HD 209458 b and assuming a noise floor of 10 ppm – which is probably conservative for JWST – we see that an increase of the TiO abundance on the dayside of only a factor 100 is visible as far as 15° from the limb plane. This increase is also conservative as the abundance of TiO at temperature below 1600 K is less than 10^{-10} ([Lodders 2002](#)). This results in a limb width $\sim 30^\circ$ which is consistent with our other estimate (see Fig. 2).

Appendix B: Finding the spherical coordinates of a cell in the cylindrical grid

To link our two coordinate systems, we use a cartesian reference frame centred around the centre of the planet and whose orthonormal reference axes are $\{\hat{\mathbf{X}}, \hat{\mathbf{Y}}, \hat{\mathbf{Z}}\}$. $\hat{\mathbf{Z}}$ is the unit vector along the rotation axis of the planet (pointing toward the north pole). $\hat{\mathbf{X}}$ points toward a reference point at the equator which will be the origin of longitudes. $\hat{\mathbf{Y}}$ is chosen to have a direct basis. The coordinates of any point in this system are $\mathbf{u} = (X, Y, Z)$.

Given the position of the observer

$$\hat{\mathbf{u}}_{\text{obs}} \equiv \begin{pmatrix} X_{\text{obs}} \\ Y_{\text{obs}} \\ Z_{\text{obs}} \end{pmatrix}_{\hat{\mathbf{X}}, \hat{\mathbf{Y}}, \hat{\mathbf{Z}}} = \begin{pmatrix} \sin \alpha_{\text{obs}} \cos \lambda_{\text{obs}} \\ \sin \alpha_{\text{obs}} \sin \lambda_{\text{obs}} \\ \cos \alpha_{\text{obs}} \end{pmatrix}, \quad (\text{B.1})$$

we need to know the physical conditions in the atmosphere at any given point \mathbf{u} determined by its cylindrical coordinates (ρ, θ, x) . For this we need to find the set of spherical coordinates (r, λ, α) corresponding to \mathbf{u} .

This can be done by first noticing that with our definitions (see Sect. 3.1.3), \mathbf{u} can be decomposed into a component in the plane of the sky and one along the line of sight

$$\mathbf{u}(\rho, \theta, x) = \mathbf{u}_{\text{ray}}(\rho, \theta) + x \hat{\mathbf{u}}_{\text{obs}}, \quad (\text{B.2})$$

where $\mathbf{u}_{\text{ray}} = (X_{\text{ray}}, Y_{\text{ray}}, Z_{\text{ray}})$ is the intersection between a (ρ, θ) -ray and the plane of the sky.

The first step is to compute X_{ray} , Y_{ray} , and Z_{ray} . These are uniquely determined thanks to the three following definitions that can be combined into one degree-two equation:

- since θ is the angle between the projection of the planetary rotation vector onto the plane of the sky and \mathbf{u}_{ray} , it can be shown that $Z_{\text{ray}} = \rho \sin \alpha_{\text{obs}} \cos \theta$;
- \mathbf{u}_{ray} is in the plane of the sky so that $\mathbf{u}_{\text{ray}} \cdot \hat{\mathbf{u}}_{\text{obs}} = 0 = X_{\text{ray}} X_{\text{obs}} + Y_{\text{ray}} Y_{\text{obs}} + Z_{\text{ray}} Z_{\text{obs}}$;
- by definition, $\rho \equiv |\mathbf{u}_{\text{ray}}| = \sqrt{X_{\text{ray}}^2 + Y_{\text{ray}}^2 + Z_{\text{ray}}^2}$.

Once these three components are known for each (ρ, θ) , the spherical coordinates of a point are given by solving

$$r^2 = \rho^2 + x^2, \quad (\text{B.3})$$

$$\lambda = \arctan\left(\frac{Y_{\text{ray}}(\rho, \theta) + x Y_{\text{obs}}}{X_{\text{ray}}(\rho, \theta) + x X_{\text{obs}}}\right), \quad (\text{B.4})$$

$$\alpha = \arccos\left(\frac{Z_{\text{ray}}(\rho, \theta) + x Z_{\text{obs}}}{r}\right). \quad (\text{B.5})$$

Finally, for numerical reasons, we determine the set of indices $(i_r, i_\lambda, i_\alpha)$ representing this specific cell in the spherical grid.

To give a concrete example, for the simple case of a synchronous planet for which the origin of longitudes is chosen at the substellar point and observed when the star, planet, and observer are perfectly aligned, we have $\hat{\mathbf{u}}_{\text{obs}} = (-1, 0, 0)$. Subsequently, solving the equations above yields $\mathbf{u}_{\text{ray}}(\rho, \theta) = (0, \rho \sin \theta, \rho \cos \theta)$, and the corresponding relationship is

$$r^2 = \rho^2 + x^2, \quad (\text{B.6})$$

$$\lambda = \arctan\left(\frac{\rho \sin \theta}{-x}\right), \quad (\text{B.7})$$

$$\alpha = \arccos\left(\frac{\rho \cos \theta}{r}\right). \quad (\text{B.8})$$

Appendix C: Computing the position of the intersection of a ray with a given interface of the spherical grid

In a spherical grid, the separation between cells is done by three types of surfaces: spheres, planes of constant longitude (meridians), and cones of constant (co)latitude. Our goal is to compute the location along a ray (x_{int}) of the intersection of this ray with any given of those surfaces. Because we know the indices of the cells before and after the intersection, we always know what type of surface we are crossing, and the value of the constant radius/longitude/colatitude identifying this surface (respectively r_{int} , λ_{int} , and α_{int}).

Because we also know the (ρ, θ) -ray we are dealing with, and the location of the observer, we must bear in mind that both $\mathbf{u}_{\text{ray}} = (X_{\text{ray}}, Y_{\text{ray}}, Z_{\text{ray}})$ and $\hat{\mathbf{u}}_{\text{obs}} = (X_{\text{obs}}, Y_{\text{obs}}, Z_{\text{obs}})$ are known.

The equations to be solved are the following.

- Intersection with a sphere:

$$x_{\text{int}} = \pm \sqrt{r_{\text{int}}^2 - \rho^2}. \quad (\text{C.1})$$

The value is positive between the limb plane and the observer and negative otherwise.

- Intersection with a meridian: using Eq. (B.4), we can solve for the intersection, which yields

$$x_{\text{int}} = \frac{Y_{\text{ray}}(\rho, \theta) - X_{\text{ray}}(\rho, \theta) \tan \lambda_{\text{int}}}{X_{\text{obs}} \tan \lambda_{\text{int}} - Y_{\text{obs}}}. \quad (\text{C.2})$$

- Intersection with a cone of constant colatitude: combining Eqs. (B.3) and (B.5) we get a second degree equation in x_{int}/ρ

$$0 = \left(\cos^2 \alpha_{\text{int}} - \cos^2 \alpha_{\text{obs}}\right) \left(\frac{x_{\text{int}}}{\rho}\right)^2 - (2 \cos \alpha_{\text{obs}} \sin \alpha_{\text{obs}} \cos \theta) \left(\frac{x_{\text{int}}}{\rho}\right) + \left(\cos^2 \alpha_{\text{int}} - \sin^2 \alpha_{\text{obs}} \cos^2 \theta\right). \quad (\text{C.3})$$

As can be seen from the equation above, the equation is the same for α_{int} and $\pi - \alpha_{\text{int}}$ (i.e. $\pm \varphi_{\text{int}}$). This means that one cannot know a priori whether the solutions found are in the northern or southern hemisphere. In fact, when two solutions exist, the ray either intersects the cone twice in the same hemisphere (when $|\varphi_{\text{obs}}| < |\varphi_{\text{int}}|$), or once on each side of the equator. To remove this degeneracy, one can compute the position of the intersection of the ray with the equator x_{equ} . Therefore, if $x_{\text{int}} > x_{\text{equ}}$ the intersection is in the same hemisphere as the observer and vice versa.

Appendix D: Rayleigh-scattering data

Typically, refractive indices follow the generic expression

$$(n - 1) \times 10^8 = A + \frac{B}{C - \lambda^{-2}}, \quad (\text{D.1})$$

with terms and their values as described in Table D.1, along with the corresponding King correction factor equations. This term is taken as unity for mono-atomic gases and is calculated ab initio as described in Bates et al. (1984) for diatomic gases. The wavelength dependency of variables is also specified in Table D.1. We note that for all the formulae in this section, the wavelength is expressed in μm .

Some molecules have non-standard parametrizations. For CO_2 (Sneep & Ubachs 2004)

$$\left(\frac{n_\lambda^2 - 1}{n_\lambda^2 + 2}\right) \times 10^8 = 1.1427 \times 10^{14} \times \left(\frac{5.79925 \times 10^{-5}}{5.0821 \times 10^2 - 1/\lambda^2} + \frac{1.2005 \times 10^{-6}}{7.9608 \times 10^1 - 1/\lambda^2} + \frac{5.3334 \times 10^{-8}}{5.6306 \times 10^1 - 1/\lambda^2} + \frac{4.3244 \times 10^{-8}}{4.619 \times 10^1 - 1/\lambda^2} + \frac{1.2181 \times 10^{-13}}{5.8474 \times 10^{-2} - 1/\lambda^2} \right) \quad (\text{D.2})$$

and

$$F_k = 1.1364 + \frac{2.53 \times 10^{-3}}{\lambda^2}. \quad (\text{D.3})$$

Table D.1. Values of the parameters used in Eq. (6).

Gas	Wavelength (μm)	$A/10^4$	$B/10^6(\mu\text{m}^{-2})$	$C/10^2(\mu\text{m}^{-2})$	King factor F_k	Reference
He	All	0.2283	0.2283	1.532	1	Thalman et al. (2014)
N ₂	$\lambda > 0.460$	0.6498	3.0740	1.44	$1.034 + 3.17 \times 10^{-4}/\lambda^2$	Thalman et al. (2014)
	$0.460 > \lambda > 0.254$	0.6677	3.1882	1.44		Bates et al. (1984)
O ₂	$\lambda < 0.254$	0.6999	3.2336	1.44	$1,096 + 1.385 \times 10^{-3}/\lambda^2 + 1.448 \times 10^{-4}/\lambda^4$	Thalman et al. (2014)
	$\lambda > 0.546$	2.1351	0.218567	0.409		Bates et al. (1984)
	$0.546 < \lambda < 0.288$	2.0564	0.248090	0.409		Sneep & Ubachs (2004)
	$0.288 > \lambda < 0.221$	2.21204	0.203187	0.409		
	$\lambda < 0.221$	2.37967	0.268988	0.409		
CO	All	2.2851	0.0456	0.51018	1.016	Sneep & Ubachs (2004)
Ar	All	0.6432135	0.028606	1.44	1	Thalman et al. (2014)

For CH₄ (Sneep & Ubachs 2004)

$$(n_\lambda - 1) = 4.6662 \cdot 10^{-4} + \frac{4.02 \times 10^{-6}}{\lambda^2}. \quad (\text{D.4})$$

For H₂O, if $\lambda > 0.23 \mu\text{m}$

$$(n_\lambda - 1) = \frac{4.92303 \times 10^{-2}}{2.380185 \cdot 10^2 - 1/\lambda^2} + \frac{1.42723 \times 10^{-3}}{5.73262 \times 10^1 - 1/\lambda^2} \quad (\text{D.5})$$

and if $\lambda < 0.23 \mu\text{m}$

$$(n_\lambda - 1) = 6.85143 \times 10^{-2} + \frac{2.10884 \times 10^{-2}}{1.32274 \times 10^2 - 1/\lambda^2} + \frac{1.4837 \times 10^{-3}}{3.932957 \times 10^1 - 1/\lambda^2}. \quad (\text{D.6})$$

For molecular hydrogen we calculate the cross-section as described in Dalgarno & Williams (1962), if $\lambda > 30 \mu\text{m}$

$$\sigma_s(\nu) = \frac{8.49 \times 10^{-33}}{\lambda^4} \quad (\text{D.7})$$

and if $\lambda < 30 \mu\text{m}$

$$\sigma_s(\nu) = \frac{8.14 \times 10^{-33}}{\lambda^4} + \frac{1.28 \times 10^{-34}}{\lambda^6} + \frac{1.61 \times 10^{-35}}{\lambda^8}. \quad (\text{D.8})$$