Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY
MNRAS **469,** 1979–1996 (2017) doi:10.1093/mnras/stx804

# HD 209458b in new light: evidence of nitrogen chemistry, patchy clouds and sub-solar water

Ryan J. MacDonald★ and Nikku Madhusudhan★
*Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK*

Accepted 2017 March 30. Received 2017 March 29; in original form 2017 January 4

## ABSTRACT

Interpretations of exoplanetary transmission spectra have been undermined by apparent obscuration due to clouds/hazes. Debate rages on whether weak $H_2O$ features seen in exoplanet spectra are due to clouds or inherently depleted oxygen. Assertions of solar $H_2O$ abundances have relied on making a priori model assumptions, for example, chemical/radiative equilibrium. In this work, we attempt to address this problem with a new retrieval paradigm for transmission spectra. We introduce POSEIDON, a two-dimensional atmospheric retrieval algorithm including generalized inhomogeneous clouds. We demonstrate that this prescription allows one to break vital degeneracies between clouds and prominent molecular abundances. We apply POSEIDON to the best transmission spectrum presently available, for the hot Jupiter HD 209458b, uncovering new insights into its atmosphere at the day–night terminator. We extensively explore the parameter space with an unprecedented $10^8$ models, spanning the continuum from fully cloudy to cloud-free atmospheres, in a fully Bayesian retrieval framework. We report the first detection of nitrogen chemistry ($NH_3$ and/or HCN) in an exoplanet atmosphere at 3.7–7.7$\sigma$ confidence, non-uniform cloud coverage at 4.5–5.4$\sigma$, high-altitude hazes at $>3\sigma$ and sub-solar $H_2O$ at $\gtrsim$3–5$\sigma$, depending on the assumed cloud distribution. We detect $NH_3$ at 3.3$\sigma$, and 4.9$\sigma$ for fully cloudy and cloud-free scenarios, respectively. For the model with the highest Bayesian evidence, we constrain $H_2O$ at 5–15 ppm (0.01–0.03) $\times$ solar and $NH_3$ at 0.01–2.7 ppm, strongly suggesting disequilibrium chemistry and cautioning against equilibrium assumptions. Our results herald a new promise for retrieving cloudy atmospheres using high-precision *Hubble Space Telescope* and *James Webb Space Telescope* spectra.

**Key words:** methods: data analysis – techniques: spectroscopic – planets and satellites: atmospheres – planets and satellites: individual (HD 209458b).

## 1 INTRODUCTION

We stand at the precipice of a new age – one where the vision of characterizing exoplanets in exquisite detail is rapidly being realized. Tremendous progress has been made in recent years towards observing various aspects of exoplanetary atmospheres, including chemical signatures (Snellen et al. 2010; Deming et al. 2013), temperature profiles (Haynes et al. 2015; Line et al. 2016), circulation patterns (Stevenson et al. 2014), clouds/hazes (Sing et al. 2016) and escape processes (Ehrenreich et al. 2015); for recent reviews, see e.g. Madhusudhan et al. (2014a, 2016c) and Crossfield (2015). Using state-of-the-art atmospheric retrieval techniques, it is now also possible to use spectroscopic observations to extract precise constraints on the chemical abundances. Such constraints are just beginning to provide tantalizing clues to planetary formation and migration pathways (e.g. Öberg, Murray-Clay & Bergin 2011; Madhusudhan, Amin & Kennedy 2014c; Mordasini et al. 2016). While a variety of observations have been used to study exoplanetary atmospheres, the majority have focused on transiting hot Jupiters ($T \sim$ 800–3000 K) whose extended atmospheres and favourable geometry make them especially amenable to transit spectroscopy.

The most observed molecule in exoplanetary atmospheres to date is $H_2O$. In recent years, the *Hubble Space Telescope*'s (*HST*) Wide Field Camera 3 (WFC3) has enabled robust detections of $H_2O$ in numerous exoplanetary transmission spectra (e.g. Deming et al. 2013). However, in almost all cases, the amplitudes of $H_2O$ absorption features are significantly lower than those expected of a cloud-free solar-composition atmosphere – typically $\sim$2 scale height (Deming et al. 2013; Kreidberg et al. 2015; Sing et al. 2016) instead of $\sim$5–10 (Madhusudhan et al. 2014a). Taken at face value, this could imply a plethora of atmospheres inherently depleted in oxygen (Madhusudhan et al. 2014b). Alternatively, they may be explained by invoking a high-altitude ($P <$1 mbar) opaque cloud deck

★E-mail: r.macdonald@ast.cam.ac.uk (RJM); nmadhu@ast.cam.ac.uk (NM)

(Deming et al. 2013) or uniform-in-altitude grey opacity, such as haze particles (Pont et al. 2013). Given the increasing number of low-amplitude or even flat spectra observed (e.g Ehrenreich et al. 2014; Knutson et al. 2014a,b; Kreidberg et al. 2014; Sing et al. 2016) the consideration of clouds has been elevated to the forefront of transmission spectroscopy.

The fundamental issue with deriving chemical abundances in cloudy exoplanetary atmospheres lies in innate degeneracies between clouds and chemistry. A wide range of solutions exist, spanning high-altitude clouds with concealed solar abundances and low-altitude, or non-existent, clouds with sub-solar abundances. This naturally leads to extremely loose constraints consistent with the full range from sub-solar through supersolar abundances (e.g. Benneke 2015). It is thus clear that clouds pose an existential challenge to robustly estimating chemical abundances. Most efforts to retrieve atmospheric properties of cloudy atmospheres have employed one-dimensional models – i.e. homogeneous terminator cloud coverage. This is despite predictions from global circulation models (GCMs) that large temperature contrasts of many hundreds of K may fuel a prominence of *partially cloudy* terminators on tidally locked hot Jupiters (Parmentier et al. 2016). The effect of partial clouds on transmission spectra retrieval has recently been examined by Line & Parmentier (2016).

Here, we offer a potential solution to the problem of interpreting transmission spectra of cloudy exoplanets. We introduce POSEIDON, a new atmospheric retrieval algorithm that includes generalized two-dimensional inhomogeneous cloud distributions. By not assuming global cloud coverage across the terminator, regions without clouds are sampled during transmission – effectively allowing one to 'peer below' the clouds and break many of the degeneracies between clouds and chemical abundances. The method we propose enables the *simultaneous* retrieval of cloud/haze properties along with precise molecular abundance constraints.

We demonstrate our new retrieval paradigm on the best transmission spectrum available, namely that of the hot Jupiter HD 209458b. As the first transiting exoplanet (Charbonneau et al. 2000), HD 209458b ignited the fledgling field of exoplanetary atmospheres. Serving as the prototypical target for atmospheric characterization, it was the first exoplanet observed to contain Na (Charbonneau et al. 2002). Various carbon- and oxygen-rich molecular species have been claimed in its atmosphere, including $H_2O$, $CH_4$, CO and $CO_2$ (Barman 2007; Madhusudhan & Seager 2009; Swain et al. 2009; Snellen et al. 2010; Deming et al. 2013). Of these detections, $H_2O$ has been robustly verified by WFC3 spectroscopy (Deming et al. 2013) and CO by high-resolution Doppler spectroscopy (Snellen et al. 2010).

Constraining abundances of chemical species became possible with the invention of atmospheric retrieval for exoplanets (Madhusudhan & Seager 2009; Madhusudhan et al. 2011). Retrieval techniques allow atmospheric properties to be derived directly from observational data in a statistically robust manner. Retrieval is now a mature field, with a wide range of codes developed and deployed to analyse atmospheres in transmission (Benneke & Seager 2012, 2013; Benneke 2015; Waldmann et al. 2015a; Line & Parmentier 2016), thermal emission (Lee, Fletcher & Irwin 2012; Barstow et al. 2013; Line et al. 2013, 2016; Waldmann et al. 2015b) and reflected light (Lee, Heng & Irwin 2013; Lavie et al. 2016; Lupu et al. 2016). The most robust inferences of the molecular abundances to date have been derived from near-infrared spectra obtained by *HST* WFC3 in the ∼1.1–1.7 μm range that contains strong spectral features due to $H_2O$, $CH_4$, $NH_3$ and HCN.

The WFC3 transmission spectrum of HD 209458b (Deming et al. 2013) was first retrieved by Madhusudhan et al. (2014b). They reported a water abundance of ∼0.01–0.05 × the prediction for a solar abundance atmosphere ($\sim 5 \times 10^{-3}$), despite a supersolar stellar metallicity of [O/H] = 0.092 ± 0.036 (Brewer et al. 2016). Such low inferred $H_2O$ abundances may provide interesting constraints on planetary formation and migration (Madhusudhan et al. 2014c). However, this retrieval assumed a cloud-free atmosphere. Benneke (2015) revisited this spectrum with a retrieval including clouds, albeit with assumed radiative-convective equilibrium along with a priori C–N–O chemistry, and inferred a composition consistent with solar abundances (also see Section 5). This view was further reinforced by Sing et al. (2016), who used transmission spectra of 10 hot Jupiters to claim that clouds and hazes, not sub-solar $H_2O$, sufficiently explain the spectra. However, this work only considered a small grid of forward models with chemical equilibrium assumed, rather than a retrieval. In another effort, Tsiaras et al. (2016b) also detected $H_2O$ but were unable to robustly constrain its abundance. Suggestions of solar $H_2O$ abundances have since been called into question. Barstow et al. (2017) performed retrievals of the Sing et al. (2016) data sets and found that 9 of their 10 planets possess sub-solar water abundances once the assumption of chemical equilibrium is relaxed – with HD 209458b being the driest at (∼0.01–0.02) × solar. These competing lines of evidence leave the question of sub-solar water abundances unsettled.

In this work, we examine these differing conclusions through the application of our two-dimensional retrieval code, POSEIDON. Using a state-of-the-art nested sampling algorithm, we extensively explore the model parameter space in a fully Bayesian retrieval framework with $>10^8$ model spectra. Unlike the retrievals of Madhusudhan et al. (2014b) and Barstow et al. (2017), we additionally include the nitrogen-bearing molecules $NH_3$ and HCN. In what follows, we introduce our retrieval framework in Section 2. We validate POSEIDON using a simulated data set in Section 3. We retrieve the atmospheric properties of HD 209458b in Section 4. Finally, in Section 5, we summarize our results and discuss the implications.

## 2 ATMOSPHERIC RETRIEVAL WITH INHOMOGENEOUS CLOUDS

Here we introduce POSEIDION, our modelling and retrieval framework for transmission spectra. Extracting atmospheric properties from an observed spectrum involves two components: (i) a parametric *forward model*; and (ii) a *statistical retrieval* algorithm to sample the model parameter space. Typical forward models assume one-dimensional geometry, i.e. average temperature, composition and cloud properties across the terminator. In this work, we generalize the formulation to account for azimuthally inhomogeneous cloud properties.

### 2.1 Forward model

Our forward model computes the transmission spectrum of an exoplanet as it transits its host star. We model the day–night terminator of the atmosphere assuming hydrostatic equilibrium and terminator-averaged temperature structure and chemistry. The model allows for inhomogeneous azimuthal cloud and haze distributions. Line-by-line radiative transfer is evaluated under the geometry depicted in Fig. 1.

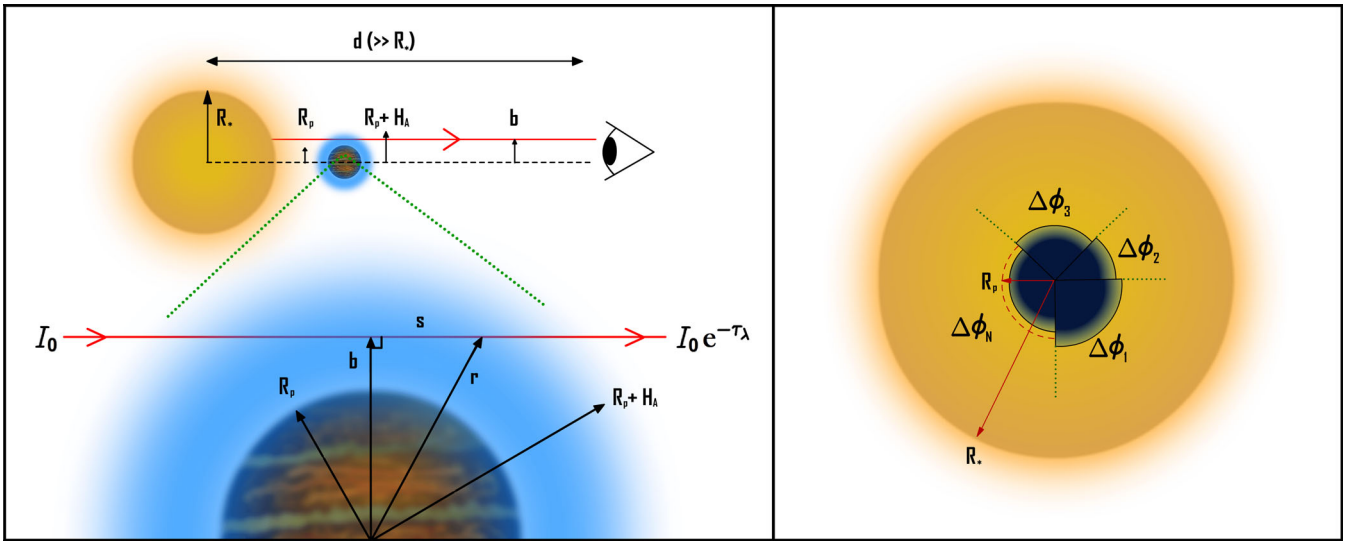



**Figure 1.** The geometry of a transiting exoplanet. Left-hand panel: Rays of stellar light of an initial intensity $I_0$ at impact parameter $b$ are attenuated due to passage through an atmosphere of height $H_A$. Right-hand panel: face-on projection, with the limb divided into $N$ regions of polar angular extent $\Delta\phi_n$, each of which may possess different cloud opacities as a function of $b$ (radial extent exaggerated). Note that the observed transit radius $R_p$ represents an average radius at which the atmosphere becomes opaque.

### 2.1.1 *Radiative transfer*

The transmission spectrum of a generalized two-dimensional atmosphere (derived in Appendix A) is represented by the wavelength-dependent transit depth, $\Delta_\lambda$, given by

$$\Delta_\lambda = \frac{1}{2\pi}\int_0^{2\pi} \delta_\lambda(\phi)\mathrm{d}\phi \approx \sum_{n=1}^{N} \bar{\phi}_n\, \delta_{\lambda,n}, \tag{1}$$

where $\delta_\lambda(\phi)$ is the transmission spectrum of an axially symmetric atmosphere with the same composition, temperature structure and cloud properties of the atmosphere at polar angle $\phi$ - i.e. the 'one-dimensional' transit depth. We approximate this general case by discretizing the atmosphere into $N$ sectors with different properties, specified by the reduced polar angular extent of the $n$th sector, $\bar{\phi}_n \equiv \Delta\phi_n/2\pi$, (such that $\sum_{n=1}^{N}\bar{\phi}_n \equiv 1$). Note that two sectors with different $\bar{\phi}$ but similar properties can be grouped into one. The equivalent one-dimensional transit depth of a sector, $\delta_{\lambda,n}$, is given by

$$\delta_{\lambda,n} = \frac{R_\mathrm{p}^2 + 2\int_{R_\mathrm{p}}^{R_\mathrm{p}+H_\mathrm{A}} b\left(1-\mathrm{e}^{-\tau_{\lambda,n}(b)}\right)\mathrm{d}b - 2\int_0^{R_\mathrm{p}} b\mathrm{e}^{-\tau_{\lambda,n}(b)}\mathrm{d}b}{R_*^2}. \tag{2}$$

Here, $R_\mathrm{p}$ and $R_\star$ are the observed radii of the planet and star, respectively, $H_\mathrm{A}$ is the maximal height of the atmosphere considered (corresponding to $10^{-6}$ bar), $b$ is the impact parameter and $\tau_\lambda(b)$ is the optical depth encountered by a ray at a given impact parameter. Equation (2) can be intuitively understood: The first term is achromatic absorption due to an opaque disc of radius $R_\mathrm{p}$, the second is absorption due to the annulus of the atmosphere and the third is a correction term accounting for rays that have sufficiently small $\tau_\lambda$ to transmit thorough the atmosphere below the observed planetary radius. Though the correction term for rays with $r < R_\mathrm{p}$ is often neglected, it must be considered for transmission spectra with two-dimensional clouds – the observed transit radius represents the average radius at which the planet becomes opaque, which can differ considerably from the local opaque radius in sectors with low opacity (see Fig. 1). Note that equation (1) reduces to equation (2) in the case $N = 1$. In this work, we consider only models with $N \leq 2$, leaving the more general case reserved for future work.

To evaluate equation (2), we require the optical depth

$$\tau_\lambda(b) = 2\int_{s=0}^{s_\mathrm{end}} \kappa_\lambda(s')\mathrm{d}s' = 2\int_b^{R_\mathrm{p}+H_\mathrm{A}} \kappa_\lambda(r)\left(\frac{r}{\sqrt{r^2-b^2}}\right)\mathrm{d}r, \tag{3}$$

where $\kappa$ is the *extinction* coefficient, $s$ is the ray path length through the atmosphere and we have suppressed the sector subscript $n$. The extinction is, in turn, a function of the chemistry and cloud properties in each layer:

$$\begin{aligned}
\kappa_\lambda(r) &= \kappa_\mathrm{chem}(r) + \kappa_\mathrm{cia}(r) + \kappa_\mathrm{cloud}(r),\\
\kappa_\mathrm{chem}(r) &= \sum_{m=1}^{N_\mathrm{spec}} n_m(r)\,\sigma_{m,\lambda}(r),\\
\kappa_\mathrm{cia}(r) &= \sum_{m=1}^{N_\mathrm{spec}}\sum_{l\geq m}^{N_\mathrm{spec}} n_m(r)\,n_l(r)\,\tilde{\alpha}_{ml,\lambda}(r),
\end{aligned} \tag{4}$$

where $n_m$ and $\sigma_{m,\lambda}$ are the number density and absorption cross-section of the $m_{th}$ species, $\tilde{\alpha}_{ml,\lambda}$ is the binary absorption coefficient due to collisionally induced absorption between species $m$ and $l$, and $\kappa_\mathrm{cloud}$ is the extinction due to clouds and hazes (defined in Section 2.1.4). Each of these quantities are written as functions of the radial coordinate $r$, which is short-hand for the pressure and temperature in each layer at which the cross-sections are computed. In order to evaluate the total extinction, we thus must specify the pressure–temperature ($P$–$T$) profile of the atmosphere.

### 2.1.2 *P–T profile*

We divide our model atmosphere into 100 layers spaced uniformly in log-pressure between $10^{-6}$ and $10^2$ bar. The temperature in each layer is computed via the parametric $P$–$T$ profile equations given in

Madhusudhan & Seager (2009):

$$P = P_0\, e^{\alpha_1 \sqrt{T - T_0}} \qquad (P_0 < P < P_1),$$
$$P = P_2\, e^{\alpha_2 \sqrt{T - T_2}} \qquad (P_1 < P < P_3),$$
$$T = T_3 \qquad (P > P_3), \tag{5}$$

where $P_0$ and $T_0$ are the pressure and temperature at the top of the atmosphere, $P_{1,3}$ and $T_{1,3}$ are specified at layer boundaries, and $P_2$ and $T_2$ specify the conditions at the (potential) temperature inversion point. The temperature slopes are controlled by $\alpha_1$ and $\alpha_2$. This profile is generic and can include thermal inversions. However, in transmission, there is no source function (i.e. thermal emission), which means that inversions are unphysical. We account for this taking $P_2 \leq P_1$.

Once $P$ and $T$ are specified in each layer, the total number density, $n_{\rm tot}(r)$, and $r$ are determined by the ideal gas law and hydrostatic equilibrium. This requires a reference pressure, $P_{\rm ref}$, to be specified, which we take to be the pressure at which $r = R_{\rm p}$. The gravitational field follows an inverse square law over radial distance with $g(R_{\rm p}) = 9.192\,{\rm m\,s^{-1}}$ and $R_{\rm p} = 1.359 R_{\rm J}$ for HD 209458b. We further assume an $H_2$–He dominated atmosphere with mean molecular weight per particle $\mu = 2.3$. Taking into account continuity at boundaries, and considering only average terminator profiles, our parametric $P$–$T$ profile is specified by seven parameters: $(\alpha_1, \alpha_2, T_0, P_1, P_2, P_3, P_{\rm ref})$.

### 2.1.3 Chemistry

We consider the main chemical species regarded as spectrally active between 0.3 and 1.7μm in hot Jupiter atmospheres: $H_2$, He, $H_2O$, $CH_4$, $NH_3$, HCN, CO, $CO_2$, Na and K (Seager & Sasselov 2000; Madhusudhan et al. 2016c). We quantify the abundance of each chemical species by its mixing ratio $X_i \equiv n_i/n_{\rm tot}$, which we assume to be uniform in the observable atmosphere. We ascribe a single mixing ratio for each species in the day–night terminator region of the atmosphere – effectively corresponding to the average limb abundance.

The molecular cross-sections are pre-computed line by line following the methodology outlined in Hedges & Madhusudhan (2016) and Gandhi & Madhusudhan (submitted). We use the latest theoretical and experimental line lists available, drawing from the HITEMP data base for $H_2O$, CO and $CO_2$ (Rothman et al. 2010) and EXOMOL for $CH_4$, HCN and $NH_3$ (Tennyson et al. 2016). The pre-computed cross-sections are binned down to a resolution of 1 $cm^{-1}$ on a predefined temperature and pressure grid ranging from $10^{-4}$ to $10^{2}$ bar and 300–3500 K. Given a $P$–$T$ profile, POSEIDON interpolates these cross-sections linearly in log-pressure and temperature to those in each atmospheric layer. Fig. 2 shows the resulting molecular cross-sections at a representative temperature and pressure of 1400 K and $10^{-3}$ bar, respectively. For Na and K, we use the cross-sections used in Christiansen et al. (2010), based on semi-analytic Lorentzian line profiles, which will be replaced with more accurate cross-sections in future work. For collisionally induced absorption, we consider $H_2$–$H_2$ and $H_2$–He absorption from the HITRAN data base (Richard et al. 2012).

### 2.1.4 A generalized two-dimensional cloud/haze prescription

Correctly modelling exoplanet transmission spectra requires the inclusion of clouds and hazes. We consider 'cloudy' regions of our atmosphere to consist of an opaque cloud deck at pressures

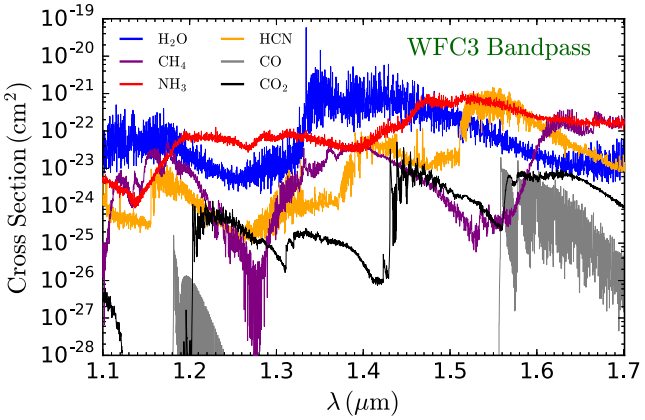



**Figure 2.** Cross-sections of spectrally active molecules in the *HST* WFC3 bandpass, evaluated at 1400 K and $10^{-3}$ bar – representative of the upper atmosphere of HD 209458b probed in transmission.

$P \geq P_{\rm cloud}$ and scattering due to hazes above the clouds (Lecavelier des Etangs et al. 2008). The extinction is given by

$$\kappa_{\rm cloud}(r) = \begin{cases} a\,\sigma_0(\lambda/\lambda_0)^{\gamma} & (P < P_{\rm cloud}) \\ \infty & (P \geq P_{\rm cloud}), \end{cases} \tag{6}$$

where $\lambda_0$ is a reference wavelength (here, 350 nm), $\sigma_0$ is the $H_2$-Rayleigh scattering cross-section at the reference wavelength ($5.31 \times 10^{-31}\,{\rm m^2}$), $a$ is the 'Rayleigh-enhancement factor' and $\gamma$ is the 'scattering slope'. The first term in equation (6) accounts for scattering due to high-altitude hazes, whilst the second term models the sharp cut-off in transmission due to the high optical depths encountered inside clouds in the transit geometry (Fortney 2005).

POSEIDON accounts for the possibility of generalized azimuthally dependant terminator cloud coverage (Fig. 1). In this work, we take $N = 2$ in equation (1) and consider region 1 to contain clouds/hazes according to equation (6) and region 2 to be clear. In this limit, we recover the 'patchy cloud' case given in Line & Parmentier (2016):

$$\Delta_\lambda = \bar{\phi}\,\delta_{\lambda,\rm cloudy} + (1 - \bar{\phi})\,\delta_{\lambda,\rm clear}, \tag{7}$$

where we have defined $\bar{\phi}_1 \equiv \bar{\phi}$, $\delta_{\lambda,1} \equiv \delta_{\lambda,\rm cloudy}$, $\delta_{\lambda,2} \equiv \delta_{\lambda,\rm clear}$ and used $\bar{\phi}_2 = 1 - \bar{\phi}_1$ to make the correspondence clear. The reduced polar angle $\bar{\phi}$ then encodes the total terminator cloud coverage (though the cloud need not be distributed continuously with a sharp boundary at the clear interface). In clear regions, we arbitrarily set $P_{\rm cloud} = 50$ bar (in order to ensure that it has no effect on the transmission spectrum) and consider scattering in the visible to be solely due to molecular $H_2$ – using the cross-section given in Dalgarno & Williams (1962).

Fig. 3 shows the effect of varying our cloud parameters ($a$, $\gamma$, $P_{\rm cloud}$, $\bar{\phi}$) on a default fiducial transmission spectrum (blue) generated by the POSEIDON forward model. The top panels demonstrate that $a$ and $\gamma$ encode the strength and slope, respectively, due to scattering, which tends to manifest most prominently at visible wavelengths $\lambda < 0.7$μm. In contrast, the altitude of the opaque cloud and the cloud coverage along the terminator strongly influence spectra across both the optical and infrared. Raising the cloud deck (lowering $P_{\rm cloud}$) results in an increased transit depth at all wavelengths, as the $\tau = 1$ surface is pushed to progressively higher altitudes. As $\bar{\phi} \to 1$, the base level of the spectrum becomes increasingly flat until the one-dimensional limit of a uniform cloud acting to shrink the transmitting annulus of the atmosphere is recovered.

With the combination of temperature structure, chemistry and cloud properties now all specified, equation (1) can be computed

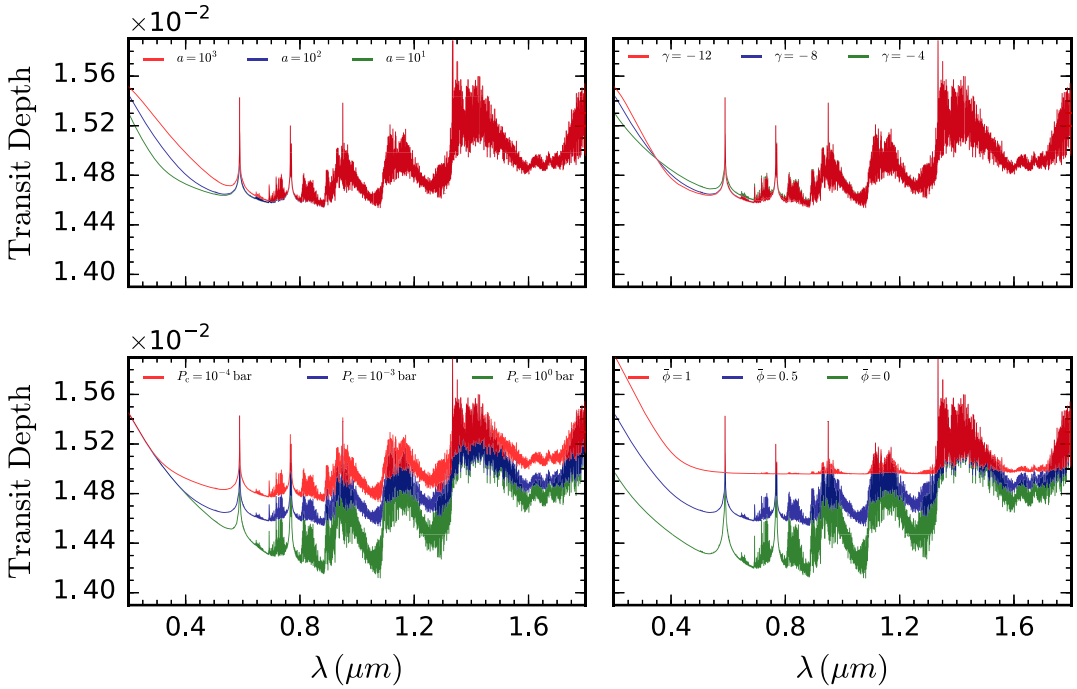



**Figure 3.** The variation of an $R \approx 10\,000$ transmission spectrum (blue) with cloud/haze properties. The top panels show progressively enhanced scattering in the optical due to hazes, whilst the bottom panels show the variation with the opaque cloud deck pressure (left-hand panel) and terminator cloud fraction (right-hand panel). For each spectrum here, the atmosphere has solar-composition at an isothermal temperature of 1400 K.

and the transmission spectra evaluated. By specifying this model in terms of parameters, a large ensemble of spectra can be evaluated for different atmospheric properties. We now turn to the essence of retrieval: the usage of a statical algorithm to extract atmospheric properties.

### 2.2 Statistical retrieval

Ultimately, we are interested in extracting the underlying properties of an exoplanet atmosphere – i.e. the values of the parameters underlying the forward model ($P$–$T$ profile, chemistry, clouds, etc.,) – from observed transmission spectra. An additional question one must assess is the suitability of the forward model itself in light of the data. These two tasks, parameter estimation and model comparison, can be accomplished within a Bayesian framework.

The basic architecture of our retrieval algorithm is depicted in Fig. 4. The atmosphere is encoded by a vector of underlying physical parameters. For a given parameter combination, the forward model outputs a spectrum that is convolved with relevant instrument point spread functions (PSFs) and/or integrated over the respective instrument functions to produce predicted data points. At each point in parameter space, these predicted data points, denoted by $y_{\rm mod}$, are compared with the observed data points, $y_{\rm obs}$, to compute the likelihood of the given set of parameters. The likelihood, in turn, informs the choice of the next set of parameters by the statistical retrieval algorithm, depicted on the right-hand side of Fig. 4. The algorithm allows thorough exploration of the entire multidimensional parameter space and computation of the Bayesian evidence – which quantifies the suitability of the model itself. POSEIDON employs a nested sampling algorithm to accomplish this purpose.

We proceed to summarize the parametrization of our forward model in Section 2.2.1, define the statistical aspects and terminology of atmospheric retrieval in Section 2.2.2 and, finally, describe

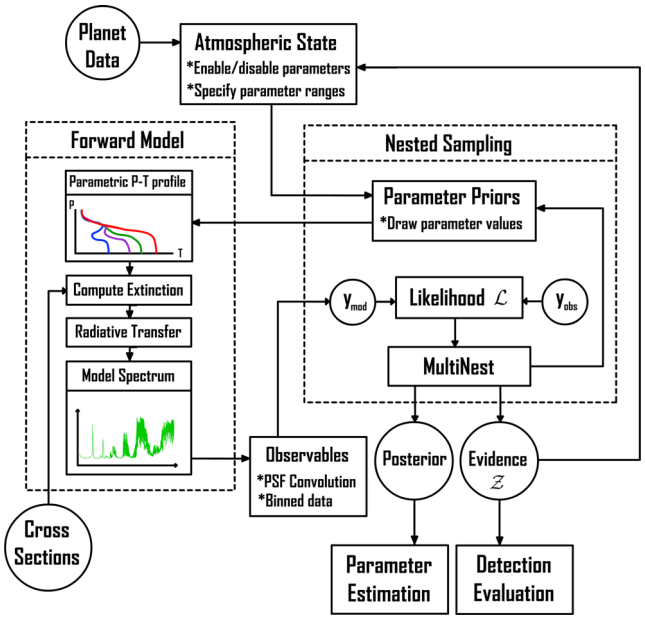



**Figure 4.** Architecture of the atmospheric retrieval algorithm POSEIDON. A forward model is repeatedly called to generate transmission spectra for different parameter inputs – each selected and guided by a nested sampling algorithm. The output is a set of posterior samples and the Bayesian evidence.

the nested sampling statistical algorithm employed by POSEIDON in Section 2.2.3.

#### 2.2.1 *Atmospheric parametrization*

In this implementation, the POSEIDON forward model is described by up to 16 parameters: seven for the terminator $P$–$T$ profile, five for

**Table 1.** Retrieval parameters and prior probabilities.

| Parameter | Prior | Range |
|---|---|---|
| $\alpha_{1,2}$ | Uniform | 0.02–1 $K^{-1/2}$ |
| $T_0$ | Uniform | 800–1600 K |
| $P_{1,2}$ | Log-uniform | $10^{-6}$–$10^{2}$ bar |
| $P_3$ | Log-uniform | $10^{-2}$–$10^{2}$ bar |
| $P_{\rm ref}$ | Log-uniform | $10^{-4}$–$10^{2}$ bar |
| $X_i$ | Log-uniform | $10^{-10}$–$10^{-2}$ |
| $a$ | Log-uniform | $10^{-4}$–$10^{8}$ |
| $\gamma$ | Uniform | −20–2 |
| $P_{\rm cloud}$ | Log-uniform | $10^{-6}$–$10^{2}$ bar |
| $\bar{\phi}$ | Uniform[a] | 0–1 |

[a]For additional sectors, a Dirichlet prior is more appropriate.

the terminator chemistry ($X_{H_2O}$, $X_{CH_4}$, $X_{NH_3}$, $X_{HCN}$, $X_{Na}$) and a four-parameter cloud/haze description. The allowed range and Bayesian prior used for each parameter are given in Table 1. Note that we have restricted our choice of molecules to those expected to dominate absorption in the WFC3 bandpass (see Fig. 2); thus, the retrieval results we present do not include $X_{CO}$ and $X_{CO_2}$. We have further elected not to describe the K abundance by an additional parameter, given the simplicity of our current alkali cross-section implementation, instead fixing the K/Na ratio to the solar value ($\approx 6.4 \times 10^{-2}$). Since we focus on hot Jupiters, we assume an $H_2$–He dominated atmosphere with a fixed He/$H_2$ ratio of 0.17.

We elect for generous 'uninformative' priors. A uniform prior probability distribution (Section 2.2.2) is ascribed to parameters expected to vary by less than two orders of magnitude, whilst a uniform-in-the-logarithm prior is used for parameters expected to vary over many orders of magnitude. We elect in this initial work to restrict ourselves to $N = 2$ atmospheric sectors, in which case only one parameter, $\bar{\phi}$, is required to describe two-dimensional effects. An additional subtlety in the choice of these priors must be made explicit: Since both the mixing ratios $X_i$ and the reduced polar extent $\bar{\phi}_i$ are subject to the constraint of addition to unity, in the most general case, a flat Dirchlet prior (uniform over a simplex subspace) is most appropriate. Here, we use the fact that we know a priori that the dominant component of hot Jupiter atmospheres is $H_2$–He to treat the remaining gases as trace species with log-uniform priors – this assumption must be relaxed for high mean molecular weight atmospheres, such as those of super-Earths (Benneke & Seager 2012). Similarly, for a two-sector atmosphere, we need to ascribe only a single uniform parameter $\bar{\phi}_1$, with the extent of the second sector automatically specified by the unity summation condition.

### 2.2.2 *Bayesian framework*

Consider a set of forward models $M_i$ described by an underlying set of physical parameters $\boldsymbol{\theta}$. Our a priori expectations on the values of the parameters are encoded in the *prior probability density function*: $p(\boldsymbol{\theta}|M_i)$. By obtaining a set of observations $\boldsymbol{y}_{\rm obs}$, we can formally update our knowledge on the values of these parameters via Bayes' theorem:

$$p\left(\boldsymbol{\theta}|\boldsymbol{y}_{\rm obs}, M_i\right) = \frac{p\left(\boldsymbol{y}_{\rm obs}|\boldsymbol{\theta}, M_i\right)\, p\left(\boldsymbol{\theta}|M_i\right)}{p\left(\boldsymbol{y}_{\rm obs}|M_i\right)} \equiv \frac{\mathcal{L}\left(\boldsymbol{y}_{\rm obs}|\boldsymbol{\theta}, M_i\right)\, \pi\left(\boldsymbol{\theta}|M_i\right)}{\mathcal{Z}\left(\boldsymbol{y}_{\rm obs}|M_i\right)}, \tag{8}$$

where in the equivalency we have defined the conventional notation for the *likelihood function* $\mathcal{L}$, prior $\pi$, and *Bayesian evidence* $\mathcal{Z}$. The quantity on the left-hand side is the *posterior probability density function*, which quantifies our knowledge on the values of the parameters of model $M_i$ following an observation. The priors for each parameter are given in Table 1, whilst the likelihood – a measure of the plausibility of the forward model producing the observed data for a choice of model parameters – is given, for observations with independently distributed Gaussian errors, by

$$\mathcal{L}\left(\boldsymbol{y}_{\rm obs}|\boldsymbol{\theta}, M_i\right) = \prod_{k=1}^{N_{\rm obs}} \frac{1}{\sqrt{2\pi}\sigma_k} \exp\left(-\frac{[y_{{\rm obs},k} - y_{{\rm mod},k}(\boldsymbol{\theta})]^2}{2\sigma_k^2}\right), \tag{9}$$

where $\boldsymbol{y}_{\rm mod}(\boldsymbol{\theta})$ is the set of (binned) model data points produced by convolving the output of the forward model with the relevant instrument PSFs and integrating over the corresponding instrument functions.

The Bayesian evidence is the key quantity employed in Bayesian model comparison. From equation (8), we see that it is simply the normalizing factor that ensures the integral of the posterior probability density over the entire parameter space evaluates to unity

$$\mathcal{Z}\left(\boldsymbol{y}_{\rm obs}|M_i\right) = \int_{\rm all\,\boldsymbol{\theta}} \mathcal{L}\left(\boldsymbol{y}_{\rm obs}|\boldsymbol{\theta}, M_i\right) \pi\left(\boldsymbol{\theta}|M_i\right) {\rm d}\boldsymbol{\theta}. \tag{10}$$

The value of the evidence can be qualitatively understood as a 'figure of merit' that is maximized by models with a high likelihood in a compact parameter space (Trotta 2008). To see this, consider the addition of a new parameter to a model. By extending the dimensionality of the parameter space, the value of the prior probability density $\pi$ will be diluted across this additional volume. The evidence for this more complex model will then increase only if the new volume contains previously unsampled regions of high likelihood. In contrast to simply fitting a spectrum, the Bayesian evidence provides an automatic implementation of Occam's Razor by penalizing models with unjustified complexity.

The quantitative utility of the evidence becomes clear when using Bayes' theorem to consider the relative probability of two competing models in light of the data:

$$\frac{p\left(M_i|\boldsymbol{y}_{\rm obs}\right)}{p\left(M_j|\boldsymbol{y}_{\rm obs}\right)} = \frac{\mathcal{Z}\left(\boldsymbol{y}_{\rm obs}|M_i\right)}{\mathcal{Z}\left(\boldsymbol{y}_{\rm obs}|M_j\right)} \frac{p\left(M_i\right)}{p\left(M_j\right)} \equiv \mathcal{B}_{ij} \frac{p\left(M_i\right)}{p\left(M_j\right)}, \tag{11}$$

where in the equivalency we have defined the *Bayes factor* for model *i* versus model *j*. We assume the final factor, expressing their a priori odds ratio, to be unity. Thus, the ratio of the evidences between two models allows the adequacy of the models themselves to be assessed. Values of at least $\mathcal{B}_{ij} = 3, 12, 150$ are often interpreted as 'weak', 'moderate' and 'strong' detections in favour of $M_i$ over $M_j$ (Trotta 2008; Benneke & Seager 2013). The Bayes factor can, in turn, be related to the commonly used frequentist measure of $\sigma$-significance (Selke, Bayarri & Berger 2001), which we also communicate when presenting Bayes factors during model comparison.

Once an adequate model is identified via Bayesian model comparison, one can then constrain the parameters within the chosen model. This is accomplished by drawing samples from the posterior (equation 8) and marginalizing (integrating) over the full range of the other parameters. The resulting probability density histograms for each parameter encode our knowledge of the underlying physics of the atmosphere. Strictly speaking, the Bayesian evidence is not required if parameter estimation is the only goal (as $\mathcal{Z}$ only normalizes the posterior), though the implicit assumption in this case is that the model itself is 'correct'.

### 2.2.3 *Nested sampling*

The statistical algorithms employed by retrieval codes are becoming increasingly sophisticated. Early retrievals used grid-based parameter space exploration (Madhusudhan & Seager 2009) or optimal estimation techniques (Lee et al. 2012; Line et al. 2012; Barstow et al. 2013) that do not allow for full marginalization of the posterior. Rigorous parameter estimation in a Bayesian framework was enabled by the integration of Markov chain Monte Carlo algorithms into retrieval codes (e.g. Madhusudhan et al. 2011; Benneke & Seager 2012; Line et al. 2013). However, these techniques do not permit a computationally efficient evaluation of the multidimensional integral in equation (10), and so renders Bayesian model comparison challenging. Atmospheric retrieval is currently undergoing a phase transition, with the technique of nested sampling (Skilling 2004) – which allows efficient computation of the Bayesian evidence, whilst also providing posterior samples for parameter estimation – now utilized in the majority of applications (Benneke & Seager 2013; Benneke 2015; Waldmann et al. 2015a; Lavie et al. 2016; Line & Parmentier 2016; Lupu et al. 2016).

POSEIDON implements the multimodal nested sampling algorithm MULTINEST (Feroz & Hobson 2008; Feroz, Hobson & Bridges 2009; Feroz et al. 2013) via the PYTHON wrapper PYMULTINEST (Buchner et al. 2014). MULTINEST computes the Bayesian evidence numerically by transforming equation (10) to a one-dimensional integral: evaluated by sweeping through progressively increasing iso-likelihood contours containing a set of 'live points' drawn from progressively shrinking ellipsoids. The full history of these samples serves to widely chart the posterior – see Benneke & Seager (2013) for a detailed discussion of the algorithm. A notable strength of MULTINEST in the context of atmospheric retrieval is its ability to navigate significantly non-Gaussian, degenerate and non-trivially curved posteriors, in addition to being fully parallelized for cluster computing.

By coupling the POSEIDON forward model to MULTINEST, we are able to assess the plausibility of a wide variety of physical phenomena (e.g. detections of chemical species, hazes, clear versus cloud-free atmospheres). For each model, we derive the posterior probability distributions of the underlying parameters, which typically requires the evaluation of many millions of spectra to obtain robust parameter estimates. Now that our retrieval framework has been established, we proceed to demonstrate its effectiveness.

## 3 VALIDATION

There are two key steps to validate a retrieval code: (i) verify that the forward model outputs correct spectra; and (ii) successfully retrieve the atmospheric state underlying a simulated data set produced by the forward model. Tests of this manner enable the accuracy and reliability of a retrieval code to be established, as well as bringing to light any potential biases or degeneracies in the results that could impact the interpretation of real spectra.

### 3.1 Forward model validation

Prior to the generation of a simulated data set, we undertook an internal comparison between the POSEIDON forward model and that utilized in Madhusudhan et al. (2014b). Though the radiative transfer schemes and parametric *P–T* profiles in both models are identical, to make a direct comparison, the cloud and haze parametrization built into POSEIDON was set to a cloud-free atmosphere (i.e. $\bar{\phi} = 0$), and we temporarily replaced our new cross-sections with those employed in Madhusudhan et al. (2014b). We tested our model over a range of parameters, finding only a minor offset ($\sim 10^{-5}$) in the transit depth due to the higher order numerical methods employed by POSEIDON.

### 3.2 Retrieval validation

We now proceed to demonstrate the typical results from an application of the POSEIDON retrieval code to a given data set. The goal here is to start with a synthetic data set, based on a known model spectrum, and use POSEIDON to retrieve the underlying model parameters. This allows us to test how well the parameters can be obtained.

We first generated a high-resolution fiducial solar-composition transmission spectrum, using the planetary properties of HD 209458b. The *P–T* profile parameters were chosen to produce a monotonically decreasing temperature with altitude, such that the temperature in the photosphere ($P \sim 10^{-2}$ bar) is around that of the planetary equilibrium temperature ($T_{\rm phot} \approx 1400$ K). We ascribed a cloud coverage of 40 per cent to the terminator, with cloudy regions consisting of a high-altitude opaque cloud deck at 0.1 mbar subsumed in a uniform-with-altitude haze. The values of the physical parameters used to produce the spectrum are given in the embedded table in Fig. 6.

With the spectrum generated, we produced a simulated data set at a precision commensurate with currently available observations. We first convolved the high-resolution spectrum with the relevant PSFs for the *HST* STIS (G430/G750) and WFC3 (G141) instruments and integrated over the corresponding instrument functions to produce a set of low-resolution binned spectral points at the same wavelength locations as the HD 209458b data given in Sing et al. (2016). We added Gaussian errors at the levels of 120 and 40 ppm in the visible and near-infrared respectively to produce simulated STIS and WFC3 data points. The combined data set served as the input to POSEIDON.

We initialized multiple nested sampling runs with 1000–8000 live points in order to confirm consistency. To efficiently compute the millions of models required to adequately explore the entire 16-dimensional parameter space, we evaluated transmission spectra at 2000 wavelengths uniformly spaced over the range 0.2–2.0 μm. We show the full posterior for the retrieval with 4000 live points in Fig. 6. We find that this number of live points offers an optimum trade-off between minimizing the error in the computed Bayesian evidence ($\Delta \ln \mathcal{Z} \approx 0.05$) and optimizing the overall time necessary to reach the convergence criteria.

The median-retrieved spectrum shows excellent agreement with the true spectrum. Fig. 5 shows that they are coincident throughout the spectral range 0.3–2.0 μm to <40 ppm precision – even where there is limited data coverage. The WFC3 bandpass is constrained even more tightly, typically to <20 ppm precision. The only region of significant deviation is in the UV below 0.3 μm, where there are no data points to inform the retrieval.

The true *P–T* profile agrees with median retrieved profile to 50 K. Indeed, the $1\sigma$ extent of the profile is tightly constrained to $\sim$100 K, with the contours matching the overall shape of the profile with altitude. As expected, the constraint becomes tighter around the photosphere ($P \sim 10^{-2}$ bar), where the information content is greatest, and expands at altitudes away from those probed in transmission. This demonstrates that the terminator *P–T* profile shape can be correctly inferred from high-precision transmission spectra.

Fig. 6 demonstrates that POSEIDON correctly retrieves all the parameters used to generate our simulated data set to $1\sigma$. The most tightly constrained parameters are the abundances of $H_2O$ and $CH_4$,

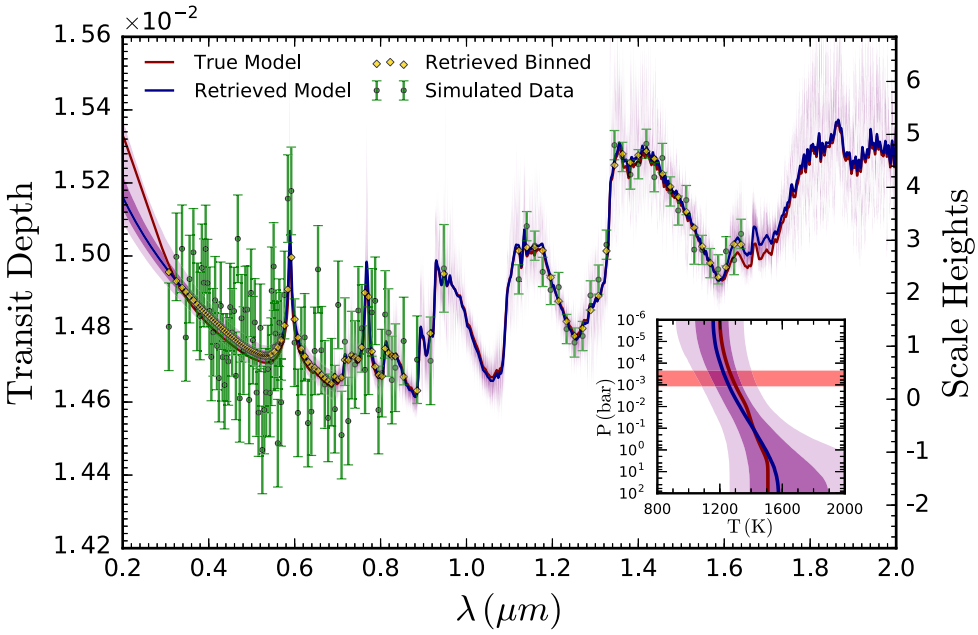



**Figure 5.** Validation of POSEIDON's ability to retrieve a spectrum and *P–T* profile from a synthetic data set. Mainfigure: A high-resolution ($R \approx 10\,000$) spectrum is generated by the forward model, convolved with the WFC3 G141 & STIS G430/G750 PSFs and integrated over the corresponding instrument functions to produce binned synthetic model points. Gaussian errors of 120 and 40 ppm in the optical and near-infrared respectively are added to create synthetic observations, shown in green. The yellow diamonds are the median binned model points resulting from the retrieval. The red and blue curves are Gaussian-smoothed representations of the true spectrum and median retrieved spectrum, respectively. The dark and light purple regions indicate $1\sigma$ and $2\sigma$ confidence regions in the transit depth at each wavelength, derived from 10 000 random samples drawn from the posterior distribution. The number of equivalent scale height above $(R_p/R_\star)^2$ is computed with respect to the median-retrieved photosphere temperature. Inset: retrieved terminator *P–T* profile. The red and blue curves are the true and median retrieved temperature profiles, respectively, the purple regions indicate $1\sigma$ and $2\sigma$ confidence regions for the temperature in each layer and the red region indicates the $1\sigma$ extent of the near-infrared photosphere ($\tau = 1$ at 1.5 μm).

which are of the order of 0.2 dex. Where absorption signatures of a molecular species are not deemed necessary to explain the data, such as with the low $NH_3$ and HCN abundances, the posterior retains the flat shape of the prior below an established upper bound. Though the mode of the cloud deck pressure is coincident with the true value, and the terminator cloud fraction is sharply localized, the haze properties remain relatively unconstrained at the noise level of the optical data. This is unsurprising, as light transmitting through the relatively small fraction of the model atmosphere above the 0.1 mbar cloud deck will be insensitive to the haze. Indeed, the small tails in the chemical abundances to higher values can be seen as stemming from a weak correlation with the possibility of strong ($a \approx 10^6$) hazes. For data where the scattering slope can be readily resolved, we do not see this tail.

### *3.2.1 Breaking cloud-composition degeneracies*

Chemical inferences are relatively independent of the values of the cloud parameters. This can be seen in Fig. 6 by the roughly horizontal correlations between the detected chemical abundances and cloud properties. This is enabled by the clear sector of the terminator, which allows POSEIDON to disentangle the reference pressure from that of the cloud deck and hence break the degeneracy between uniform clouds and chemistry. To reiterate: *non-uniform terminator cloud coverage enables precise determination of chemical abundances from transmission spectra*.

Having verified POSEIDON's ability to correctly retrieve atmospheric properties from high-precision transmission spectra, we now turn our attention towards the observed transmission spectrum of HD 209458b.

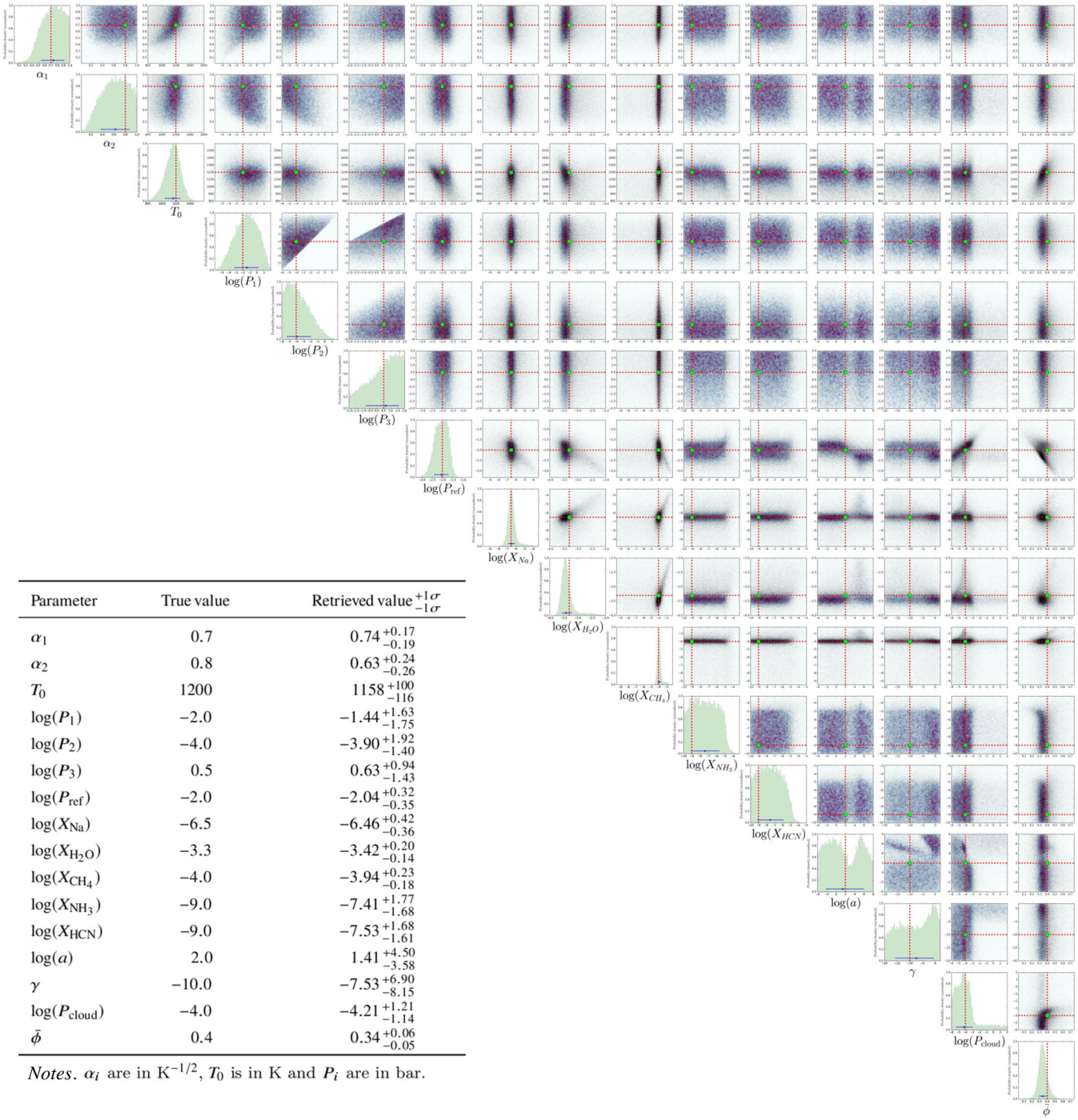
## 4 RESULTS: APPLICATION TO HD 209458B

We here report the first application of POSEIDON to an exoplanet transmission spectrum. Specifically, we consider the visible and near-infrared spectrum of HD 209458b presented in Sing et al. (2016), as this is the highest quality transmission spectrum presently available. Our retrieval, which includes observations over the range ∼ 0.3–1.7μm taken by the *HST* STIS and WFC3 instruments, is the most extensive to date, involving the computation of an unprecedented $10^8$ spectra.

We ran a comprehensive suite of retrievals spanning the model space of cloudy atmospheres. Four different cloud/haze prescriptions were considered: partial cloud coverage (both with and without hazes), uniform clouds and cloud-free atmospheres. Within each cloud prescription, we performed nested model comparisons, whereby multiple retrievals are performed with chemical species selectively removed to evaluate their detection significances. This amounts to eight independent retrievals, with $\sim 5 \times 10^6$ model computations each, for each cloud prescription, i.e. $\gtrsim 10^8$ model computations in total. We finally report constraints on the terminator chemical abundances and temperature profile for the cloud model most preferred by the data. We additionally illuminate how the assumed cloud model influences the inferred chemical abundances.

### 4.1 Cloud properties

We detect the presence of partial cloud coverage across the terminator of HD 209458b. Table 2 summarizes the results of our Bayesian model comparison, which indicates a $4.5\sigma$ preference for the patchy cloud model over the uniform cloud model. The cloud-free model




| Parameter | True value | Retrieved value $^{+1\sigma}_{-1\sigma}$ |
|---|---|---|
| $\alpha_1$ | 0.7 | $0.74^{+0.17}_{-0.19}$ |
| $\alpha_2$ | 0.8 | $0.63^{+0.24}_{-0.26}$ |
| $T_0$ | 1200 | $1158^{+100}_{-116}$ |
| $\log(P_1)$ | −2.0 | $-1.44^{+1.63}_{-1.75}$ |
| $\log(P_2)$ | −4.0 | $-3.90^{+1.92}_{-1.40}$ |
| $\log(P_3)$ | 0.5 | $0.63^{+0.94}_{-1.43}$ |
| $\log(P_{\rm ref})$ | −2.0 | $-2.04^{+0.32}_{-0.35}$ |
| $\log(X_{\rm Na})$ | −6.5 | $-6.46^{+0.42}_{-0.36}$ |
| $\log(X_{\rm H_2O})$ | −3.3 | $-3.42^{+0.20}_{-0.14}$ |
| $\log(X_{\rm CH_4})$ | −4.0 | $-3.94^{+0.23}_{-0.18}$ |
| $\log(X_{\rm NH_3})$ | −9.0 | $-7.41^{+1.77}_{-1.68}$ |
| $\log(X_{\rm HCN})$ | −9.0 | $-7.53^{+1.68}_{-1.61}$ |
| $\log(a)$ | 2.0 | $1.41^{+4.50}_{-3.58}$ |
| $\gamma$ | −10.0 | $-7.53^{+6.90}_{-8.15}$ |
| $\log(P_{\rm cloud})$ | −4.0 | $-4.21^{+1.21}_{-1.14}$ |
| $\bar{\phi}$ | 0.4 | $0.34^{+0.06}_{-0.05}$ |

*Notes.* $\alpha_i$ are in $\rm K^{-1/2}$, $T_0$ is in K and $P_i$ are in bar.

**Figure 6.** Full posterior distribution from POSEIDON's retrieval of the atmospheric properties underlying the synthetic data in Fig. 5. Main figure: corner plot depicting correlations between pairs of retrieved parameters and marginalized histograms for each parameter. The red dashed lines and green squares indicate the true values of each parameter. Table inset: the true parameter values used to generate the simulated data set and the corresponding median retrieved values. All parameters are correctly retrieved to within $1\sigma$.

is ruled out to $>5\sigma$ when compared to the patchy cloud model and to $3.4\sigma$ when compared to uniform clouds. These detection significances can be understood by examining the posterior of the patchy cloud retrieval (Fig. 7), which indicates a cloud fraction of $\bar{\phi} = 0.57^{+0.07}_{-0.12}$. Given that the cloud fraction is closer to 1 than 0, it is unsurprising that the Bayesian evidence of the uniform cloud model exceeds that of the cloud-free model. Even after marginalizing over the other parameters, the posterior probability distribution of $\bar{\phi}$ is inconsistent with values of 0 or 1 ($5\sigma$ range: 0.14–0.77); reinforcing the large penalty the Bayesian evidence suffers when forced to consider models fixed at these values. As one final assessment of the presence of partial clouds, we conducted an additional retrieval where the cloud fraction was held fixed at the best-fitting (min $\chi^2_r$) value from the patchy cloud retrieval ($\bar{\phi} \approx 0.47$). This model, the 'fixed fraction' model in Table 2, possesses the largest value of the Bayesian evidence amongst our cloud models and

**Table 2.** Bayesian model comparison of the terminator cloud distribution on HD 209458b.

| Model | Evidence $\ln(\mathcal{Z}_i)$ | Best-fitting $\chi^2_{\mathrm{r,min}}$ | Bayes factor $\mathcal{B}_{0i}$ | Detection of ref. |
|---|---|---|---|---|
| Patchy clouds | 953.16 | 1.45 | Ref. | Ref. |
| Uniform clouds | 944.91 | 1.52 | $3.8 \times 10^3$ | $4.5\sigma$ |
| Cloud-free | 940.47 | 1.62 | $3.3 \times 10^5$ | $5.4\sigma$ |
| No haze | 949.57 | 1.57 | 36 | $3.2\sigma$ |
| Fixed fraction | 953.60 | 1.44 | 0.65 | N/A |

*Notes*. The 'fixed fraction' model has $\bar{\phi}$ held at the best-fitting value from the 'patchy clouds' model (0.47). The 'no haze' model is a patchy cloud model that considers only $H_2$-Rayleigh scattering in cloudy regions. $\chi^2_{\mathrm{r,min}}$ is the minimum reduced chi-square. An $n\sigma$ detection ($n \geq 3$) indicates the degree of preference for the reference model over the alternative model.

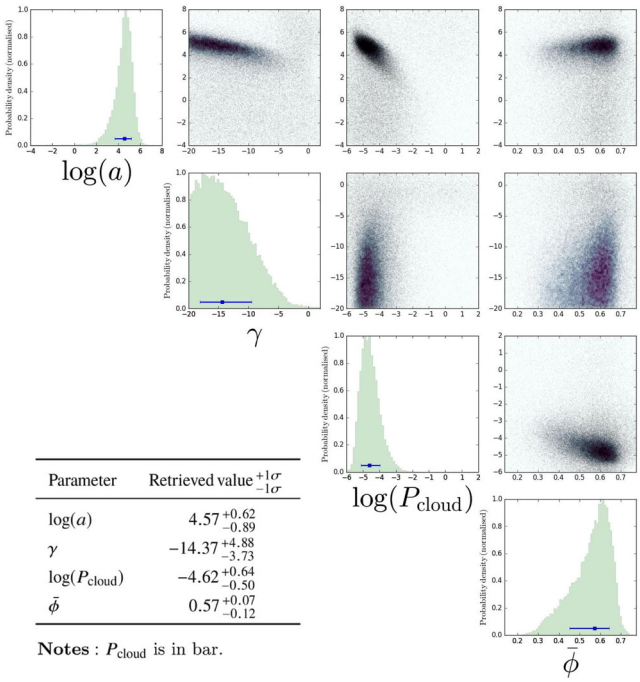



**Figure 7.** HD 209458b's terminator cloud properties, retrieved within the 'patchy clouds' model (see Table 2). Main figure: corner plot depicting correlations between pairs of derived parameters and marginalized histograms for the values of each parameter extracted by the retrieval. Table inset: median-retrieved cloud parameters and $1\sigma$ confidence levels.

hence reinforces our assertion that patchy clouds are favoured by the data.

The detection significance of patchy clouds is found to be somewhat sensitive to the lower limit of the temperature prior considered. Identical retrievals with a lower limit on $T_0$ of 400 K result in values of $\ln \mathcal{Z}$ of 949.35 and 947.17 for uniform clouds and cloud-free atmospheres, respectively. The corresponding detection significances for patchy clouds are $3.2\sigma$ and $3.8\sigma$, respectively. This effect is caused by a tendency for both cloud-free and uniform cloud models to favour lower temperature solutions (e.g. see Tsiaras et al. 2016b)

We infer a high-altitude (~0.01–0.1 mbar) cloud deck on the cloudy fraction of the terminator. Above this cloud deck, we report a moderate detection ($3.2\sigma$) of high-altitude hazes. The necessity of strongly enhanced Rayleigh scattering (~5000–100 000 × $H_2$-Rayleigh) is visibly apparent in Fig. 8 from the steep increase in the transit depth towards shorter wavelengths. In addition to the high strength of the Rayleigh enhancement factor, the scattering slope is remarkably negative, tending to prefer values towards the lower edge of the prior. The presence of such a strong scattering cross-section at these altitudes suggests two immediate possibilities: (i) incredibly light particles capable of being lofted to these altitudes by vertical mixing; and (ii) continuous replenishment of the species at altitude (e.g. by photochemical reactions).

In what follows, we select the 'fixed fraction' cloud model, unless otherwise stated. This is the model that maximizes the Bayesian evidence, and thus holds the greatest sway in light of the data. We display the full posterior of this model in Fig. 9. Note, in particular, that the values of the cloud parameters (and their associated errors) in this model remain consistent to within $1\sigma$ of the values shown in Fig. 7. This indicates that the uncertainty induced by allowing the cloud fraction to vary as a free parameter does not overly effect the inferences of the remaining cloud parameters compared to the case when it is fixed at the best-fitting value. In order to illustrate the importance of selecting the cloud model that is supported by the data, in the next section, we will demonstrate how the inferred values and constraints on the retrieved chemical abundances crucially depend on the assumed cloud model.

## 4.2 Chemistry

### *4.2.1 Detections*

We confirm previous detections of Na (Charbonneau et al. 2002; Snellen et al. 2008) and $H_2O$ (Deming et al. 2013) in the transmission spectrum of HD 209458b. Our nested model comparison establishes the presence of $H_2O$ and an alkali absorber (Na/K) at $9.1\sigma$ and $7.3\sigma$ confidence, respectively (Table 3). We do not detect $CH_4$, though we establish at $>10\sigma$ confidence that the presence of either $CH_4$ or $H_2O$ is required (due to their overlapping absorption features at 1.15 and 1.40 μm – see Benneke & Seager 2013, and Fig. 2).

We additionally detect the presence of nitrogen chemistry (in the form of $NH_3$ and/or HCN) at $3.7\sigma$ confidence. The Bayes factor of the model including nitrogen chemistry compared to the model without is 186 – indicating 'strong evidence' in favour of the presence of $NH_3$ and/or HCN on Jeffreys' scale. This detection is robust to the assumed cloud model, rising to $4.9\sigma$ and $7.7\sigma$ in uniformly cloudy and cloud-free models, respectively. In all cases, the chemical detections and confidences are insensitive to the lower limit on the temperature prior.

When considering the partial cloud model preferred by the data (Table 2), we are unable to distinguish between the presence of $NH_3$ and HCN. This is due to the effect of partial clouds in altering the slopes of absorption features (see Fig. 3). However, in both cloud-free and uniformly cloudy models, this degeneracy is lifted and $NH_3$ is detected at $4.9\sigma$ ($\mathcal{B}_{0i} = 22\,000$) and $3.3\sigma$ ($\mathcal{B}_{0i} = 58$) confidence, respectively. We do not detect HCN in the cloud-free or uniformly cloudy models, obtaining flat abundance posteriors with established upper limits of $\sim 10^{-6}$ and $10^{-4}$, respectively. The influence of nitrogen chemistry on the transmission spectrum is shown in Fig. 11 and discussed further in Section 4.2.3.

### *4.2.2 Abundance constraints*

We report our constraints on the chemical abundances on HD 209458b's terminator in Fig. 10. The abundances we report are amongst the most precise ever obtained from an exoplanet transmission spectrum (~0.3 dex for $H_2O$). This is despite our

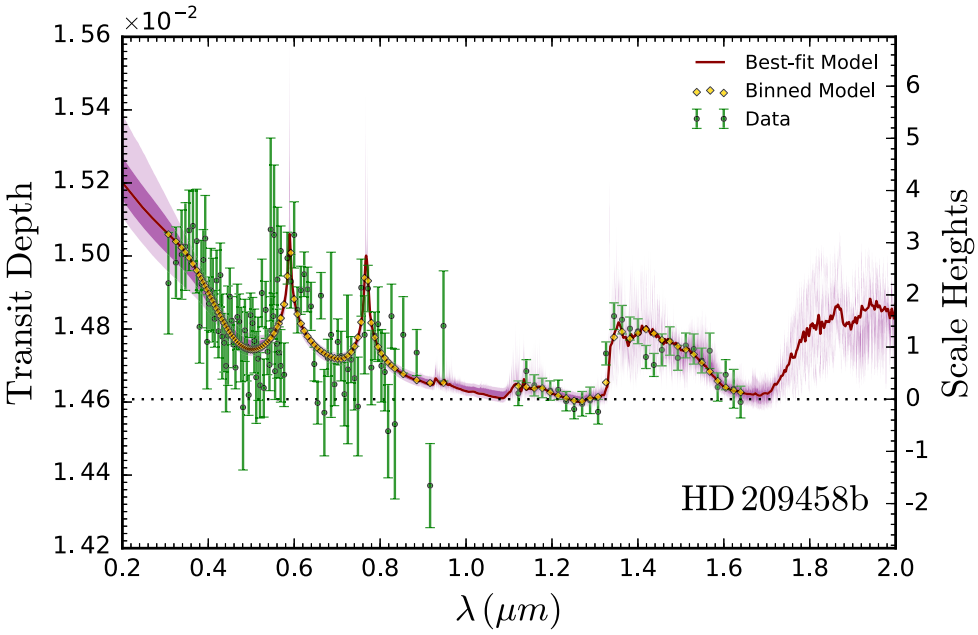



**Figure 8.** POSEIDON's retrieval of the visible and near-infrared transmission spectrum of HD 209458b. The observed transit depth is indicated by green circles with error bars. The dark red curve is a Gaussian-smoothed representation of the best-fitting retrieved spectrum. The yellow diamonds are the median binned model points produced by the retrieval. The dark and light purple regions indicate $1\sigma$ and $2\sigma$ confidence regions (at $R \approx 10\,000$) in the transit depth at each wavelength derived from 10 000 random sample draws from the posterior distribution. The black dotted line indicates the value of $(R_p/R_\star)^2$. The number of equivalent scale heights above this reference baseline is computed with respect to the median-retrieved photosphere temperature (see Section 4.3).

marginalization over two-dimensional terminator cloud coverage, due to POSEIDON's ability to break the degeneracy between clouds and chemistry (Section 3.2.1). We explore how the assumed cloud distribution affects the inferred abundances in Section 4.2.4.

The terminator of HD 209458b is inconsistent with a solar $H_2O$ abundance. This is established at $>5\sigma$ confidence for both partial cloud and cloud-free models and at $>3\sigma$ confidence for uniform clouds. The retrieved value, $\log(X_{H_2O}) = -5.24^{+0.36}_{-0.27}$, is remarkably consistent with the values reported by both Madhusudhan et al. (2014b) $\left(-5.27^{+0.65}_{-0.16}\right)$ and Barstow et al. (2017) (−5.3 to −5.0). This is unsurprising, as the observed spectrum (Fig. 8) shows the amplitude of the $H_2O$ absorption feature at 1.4 μm is only 2 scale heights – whereas a solar-composition atmosphere at a similar temperature and cloud coverage fraction would give ∼5 scale heights (Fig. 5).

We demonstrate that the Na abundance can be reasonably well constrained (∼0.6 dex), despite the ∼120 ppm errors in the optical STIS data. Whilst this serves as an important demonstration of principle, we caution against reading too much into the retrieved values. This is due to the simplicity of our treatment of the alkali cross-sections in this work (see Section 2.1.3). We will address precise alkali abundance constraints in future work.

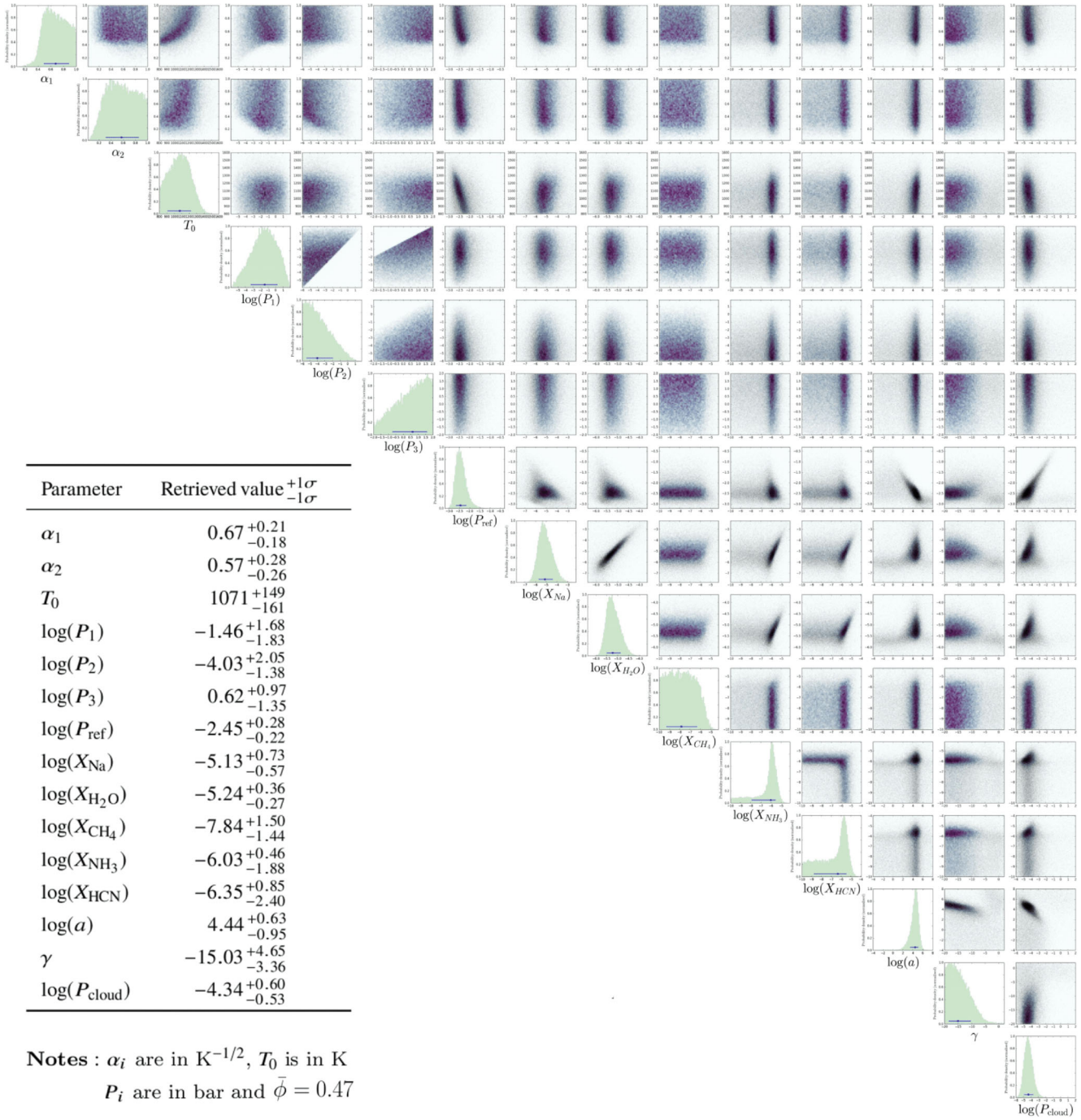
The abundances of both $NH_3$ and HCN show a sharp peak at $\sim 10^{-6}$ with a tail towards lower abundances. The abundance of ammonia is the best constrained at $\log(X_{NH_3}) = -6.03^{+0.46}_{-1.88}$ (0.01–2.7 ppm). The tails stem from the fact that either of these nitrogen-bearing species can explain the observed absorption features – if one has high abundance, the other will have low abundance and vice versa. Ultimately, it is this long tail that prevents a unique determination of the species causing the absorption. It can be seen from the lower probability density of the tail in Fig. 10 (and the higher Bayes factor in Table 3) that the presence of $NH_3$ is marginally preferred over HCN when considering partial cloud coverage. This symmetry is broken when considering cloud-free or uniformly cloudy models, both of which feature well-constrained $NH_3$ and a flat posterior for HCN – explaining the $NH_3$ detections observed in these models. Given the combination of these high $NH_3$ detection significances and the coincident peak of its abundance distribution between the cloud-free and the partial cloud models, we strongly suspect that it is this species, not HCN, that is the source of the detected nitrogen chemistry.

### 4.2.3 *Nitrogen chemistry*

We now proceed to identify the absorption features giving rise to our detection of nitrogen chemistry. In Fig. 11, we show the effect on our best-fitting spectrum (red) $\left(\log X_{[H_2O,\,CH_4,\,NH_3,\,HCN]} = [-5.21,\ -8.63,\ -5.72,\ -8.39]\right)$ of removing the $NH_3$ and HCN. Given that $NH_3$ is the dominant nitrogen-bearing molecule for this spectrum, this amounts to an assessment of the impact of ammonia on near-infrared transmission spectra.

The primary impact of $NH_3$ absorption in the WFC3 bandpass is to raise the transit depth of HD 209458b by $\sim 5 \times 10^{-5}$ compared to what would be expected from pure $H_2O$ absorption (blue) over the spectral range ∼1.45–1.7 μm. A secondary feature of magnitude $\sim 1 \times 10^{-5}$ is seen between ∼1.2–1.3 μm. These absorption features are readily identified by an examination of the $NH_3$ cross-section (Fig. 2), which is seen to dominate that of $H_2O$ over these two regions. The necessity of additional absorption is evident from the data itself, as there are four data points elevated by $2\sigma$ and one point elevated by $3\sigma$ above the model without nitrogen chemistry.

We now offer suggestions on how to distinguish between $NH_3$ and HCN in transmission spectra when nitrogen chemistry is detected. The difficulty inherent in this task is demonstrated in Fig. 11, where the black curve shows the effect of interchanging the abundances




| Parameter | Retrieved value $^{+1\sigma}_{-1\sigma}$ |
|---|---|
| $\alpha_1$ | $0.67^{+0.21}_{-0.18}$ |
| $\alpha_2$ | $0.57^{+0.28}_{-0.26}$ |
| $T_0$ | $1071^{+149}_{-161}$ |
| $\log(P_1)$ | $-1.46^{+1.68}_{-1.83}$ |
| $\log(P_2)$ | $-4.03^{+2.05}_{-1.38}$ |
| $\log(P_3)$ | $0.62^{+0.97}_{-1.35}$ |
| $\log(P_{\rm ref})$ | $-2.45^{+0.28}_{-0.22}$ |
| $\log(X_{\rm Na})$ | $-5.13^{+0.73}_{-0.57}$ |
| $\log(X_{\rm H_2O})$ | $-5.24^{+0.36}_{-0.27}$ |
| $\log(X_{\rm CH_4})$ | $-7.84^{+1.50}_{-1.44}$ |
| $\log(X_{\rm NH_3})$ | $-6.03^{+0.46}_{-1.88}$ |
| $\log(X_{\rm HCN})$ | $-6.35^{+0.85}_{-2.40}$ |
| $\log(a)$ | $4.44^{+0.63}_{-0.95}$ |
| $\gamma$ | $-15.03^{+4.65}_{-3.36}$ |
| $\log(P_{\rm cloud})$ | $-4.34^{+0.60}_{-0.53}$ |

**Notes** : $\alpha_i$ are in $\rm K^{-1/2}$, $T_0$ is in K
$P_i$ are in bar and $\bar{\phi} = 0.47$

**Figure 9.** Full posterior distribution from POSEIDON's retrieval of the transmission spectrum of HD 209458b. Main figure: corner plot depicting correlations between pairs of retrieved parameters and marginalized histograms for the values of each parameter extracted by the retrieval. The abundances of Na and $H_2O$ are tightly constrained, relatively independent of the cloud parameters – vindicating the prediction from the synthetic data retrieval shown in Fig. 6 and discussed in Section 3.2.1. Table inset: median-retrieved values and $1\sigma$ confidence levels for each parameter, following marginalization over the other 14 dimensions of the parameter space.

of $NH_3$ and HCN, such that HCN becomes the dominant nitrogen-bearing molecule. HCN causes an increase in the transit depth that is almost identical to that of $NH_3$ from $\sim$1.53−1.6 μm (Fig. 2), though it generally matches pure $H_2O$ absorption outside this range. The degeneracy with $NH_3$ may be lifted by high-resolution observations in three regions of the WFC3 bandpass: (i) 1.2–1.32 μm; (ii) 1.46–1.52 μm; and (iii) 1.6–1.7 μm. Given that the difference between $NH_3$ and HCN-dominated spectra are of the order of the error bars (35ppm), this is pushing the frontier of current observational capabilities.

There is, however, another potential avenue that may enable the unique detection of $NH_3$/HCN with WFC3 observations. Namely, the sharp $NH_3$ posterior shown in Fig. 10's cloud-free model suggests that genuinely cloud-free atmospheres may allow highly robust detections of $NH_3$ and/or HCN. We discuss this possibility further in what follows.

**Table 3.** Bayesian model comparison of the chemistry on the terminator of HD 209458b.

| Model | Evidence $\ln(\mathcal{Z}_i)$ | Best-fitting $\chi^2_{\mathrm{r,min}}$ | Bayes factor $\mathcal{B}_{0i}$ | Detection of ref. |
|---|---|---|---|---|
| Full chemistry | 953.60 | 1.44 | Ref. | Ref. |
| No $H_2O/CH_4$ | 904.62 | 2.35 | $1.9 \times 10^{21}$ | $10.1\sigma$ |
| No $H_2O$ | 914.62 | 2.14 | $8.4 \times 10^{16}$ | $9.1\sigma$ |
| No Na/K | 928.92 | 1.93 | $5.2 \times 10^{10}$ | $7.3\sigma$ |
| No $NH_3$/HCN | 948.37 | 1.53 | 186 | $3.7\sigma$ |
| No $NH_3$ | 952.80 | 1.44 | 2.2 | N/A |
| No HCN | 953.35 | 1.42 | 1.3 | N/A |
| No $CH_4$ | 954.01 | 1.42 | 0.7 | N/A |

*Notes*. The 'full chemistry' model includes opacity due to $H_2$, He, Na, K, $H_2O$, $CH_4$, $NH_3$, and HCN, and corresponds to the 'fixed fraction' cloud model given in Table 2. $\chi^2_{\mathrm{r,min}}$ is the minimum reduced chi-square. An $n\sigma$ detection ($n \geq 3$) indicates the degree of preference for the reference model over the alternative model.

#### 4.2.4 *The influence of clouds on chemical abundances*

We now proceed to quantify the extent to which the cloud model assumed by a retrieval can influence the inferred chemical abundances. We have already identified an innate challenge in distinguishing between $NH_3$ and HCN in partially cloudy transmission spectra. We further see in Section 4.2.1 and Fig. 10 that the detection significances and abundance constraints depend crucially on the cloud model employed by retrievals.

In general, our cloud-free abundances are artificially well constrained at lower values than those inferred by the preferred partial cloud model (though they remain consistent within $1\sigma$). However, this suggests that planets with genuinely low cloud coverage may permit strong detections with precise abundance constraints. This is especially evident in the case of nitrogen chemistry, where our cloud-free model clearly identifies the presence of $NH_3$ at $\approx 5\sigma$ and constrains its abundance to $\log(X_{\mathrm{NH_3}}) = -5.92^{+0.10}_{-0.11}$. We hence suggest that, even with current observational errors, precise abundance determinations of nitrogen-bearing molecules may be obtained on planets with low overall terminator cloud coverage.

In contrast, the uniform cloud model tends to overestimate chemical abundances. This is strikingly apparent in the case of Na, which favours unphysical values towards the upper limit of the prior ($10^{-2}$). The biasing of abundances to erroneously high values under the assumption of uniform clouds is a consequence of a fundamental degeneracy between clouds, hazes, the reference pressure, and the chemical abundances. For uniform terminator cloud coverage, the cloud deck pressure, $P_{\mathrm{cloud}}$, and the pressure at the radius of the planet, $P_{\mathrm{ref}}$, are equal, existing on a line in parameter space without a unique solution. Defining $P_{\mathrm{ref}}$ at a radii different to $R_{\mathrm{p}}$ merely offsets the line of degeneracy. This family of solutions determines the baseline of the spectrum (Fig. 8, black dotted line). When $P_{\mathrm{ref}}$ and $P_{\mathrm{cloud}}$ are lowered along this solution line, both the amplitude of spectral features and the Rayleigh-enhancement factor can rise to produce an identical spectrum. We have verified that this biasing to higher abundances is an artefact of the uniform cloud model itself by running simulated retrievals with solar Na abundances and confirming that the behaviour seen in Fig. 10 can be reproduced for synthetic data. We thus caution against the blind application of uniform cloud models; indeed, our results suggest that cloud-free models are a better option if one solely wishes to estimate the chemical abundances of a transiting exoplanet to within $1\sigma$.

We have shown that partial cloud coverage breaks the degeneracy between clouds and chemistry imposed artificially by the assumption of uniform clouds (section 3.2). Alternatively, this degeneracy may be broken by assuming a priori knowledge of the scattering slope in the optical (e.g Benneke & Seager 2012) or of the molecular/condensate chemistry (Benneke 2015). The attraction of partial clouds is that we do not have to make such assumptions, allowing us to directly retrieve cloud, haze, chemical and temperature properties of the atmosphere simultaneously. Though we have shown that partial clouds enable precise determination of the $H_2O$ and $CH_4$ abundances, and are preferred by the Bayesian evidence at high significance, they complicate the interpretation of nitrogen chemistry as they tend to render the slopes of molecular features more shallow (Fig. 3, bottom right-hand panel). This decreases the magnitude of the transit depth 'gap' induced by nitrogen chemistry over the range $\sim$1.45–1.70 μm (Fig. 11) by broadening the width of the pure $H_2O$ absorption feature.

Given these chemical and cloud inferences, we now proceed to present our retrieved *P*–*T* profile of the terminator of HD 209458b. A combination of all three of these properties is required to build a coherent picture of the conditions on the terminator.

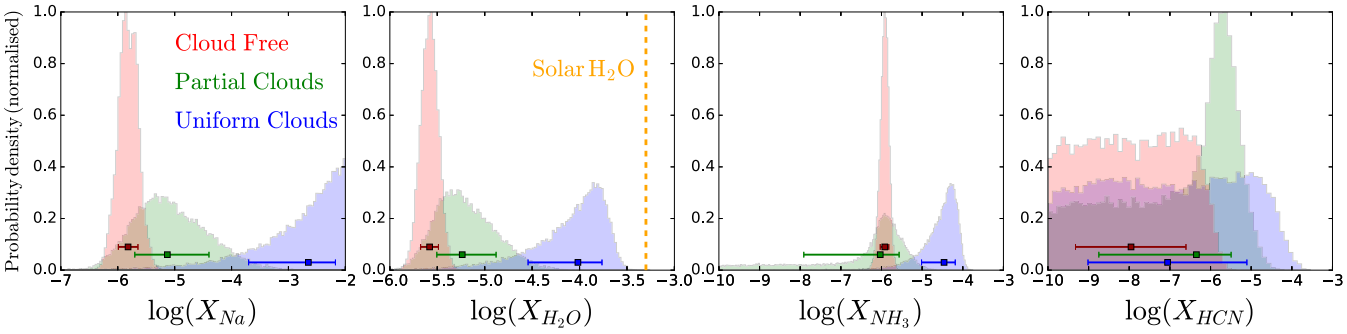



**Figure 10.** Marginalized posterior distributions of the chemical abundances on the terminator of HD 209458b. The red, green, and blue histograms are inferences from the 'cloud-free', 'fixed fraction', and 'uniform clouds' models summarized in Table 2. $H_2O$, Na/K, and $NH_3$/HCN are detected in all three models (Table 3). $NH_3$ is detected in both the cloud-free and uniform cloud models at confidence levels of $4.9\sigma$ and $3.3\sigma$, respectively. All three models are inconsistent with a solar water abundance (indicated by the dashed orange line at $\log(X_{\mathrm{H_2O}}) = -3.3$) at $>3\sigma$ confidence. The uniform cloud model is biased to higher abundances (see the text for discussion), whilst the cloud-free model is consistent with the abundances from the preferred partial cloud model. The cloud-free model underestimates the uncertainty in the derived abundances. For clarity, we do not show the posterior of the 'patchy clouds' model of Table 2 (which is almost identical to that of the 'fixed fraction' model). We also do not show $CH_4$, as it is unconstrained in all our retrievals (e.g. Fig. 9).

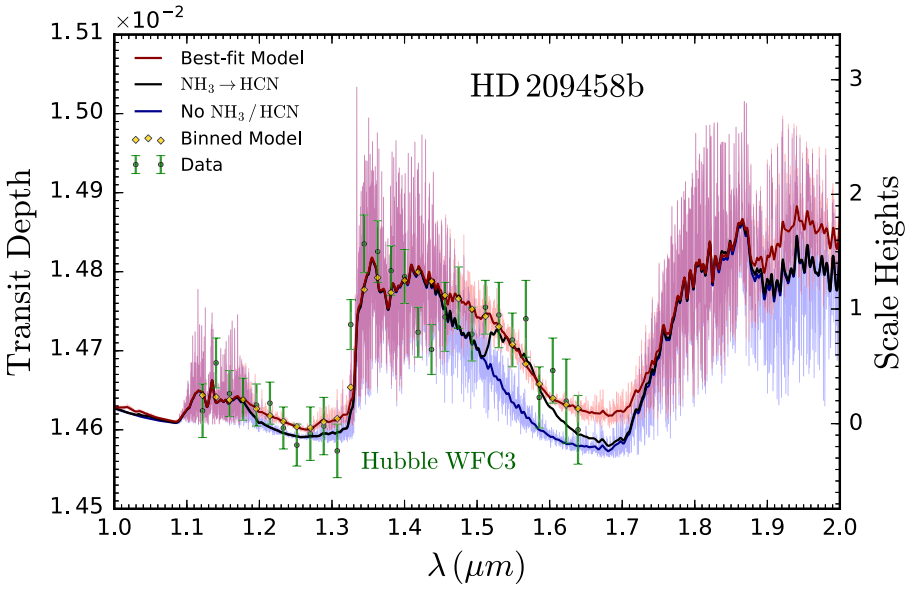



**Figure 11.** Evidence of nitrogen chemistry on the terminator of HD 209458b. The observed transit depth is indicated by green circles with error bars. The best-fitting spectrum (2ppm $NH_3$) is shown in red at a resolution of $R \approx 10\,000$. The yellow diamonds are the binned model points corresponding to the best-fitting spectrum. The black spectrum is a model identical to the best fit with the abundances of $NH_3$ and HCN interchanged (such that HCN becomes the dominant nitrogen-bearing species). The blue spectrum is a model identical to the best fit with $NH_3$ and HCN removed. The dark red, black and blue curves are Gaussian-smoothed representations of the high-resolution spectra with corresponding colours. Performing full atmospheric retrievals, exploring our entire parameter space of chemistry, temperature structure, and clouds/hazes, establishes that the model including $NH_3$ and HCN (red) is preferred by the data over models with no nitrogen chemistry at $3.7\sigma$ confidence. The primary evidence for a nitrogen-bearing species comes from the additional absorption over the spectral range ∼1.45–1.70 μm.

### 4.3 Temperature structure

The temperature structure on the terminator of HD 209458b is not isothermal. If the data supported an isothermal profile, we would expect to see $\alpha_{1,2}$ tending towards larger values, which we do not observe (see Fig. 9). Whilst unsurprising from physical arguments and GCM simulations, this point merits emphasis. It is often assumed in transmission retrieval that with currently available data (i) it is not possible to retrieve the shape of the terminator *P–T* profile; and (ii) an isothermal profile does not overly affect the inferred abundances. Here, we demonstrate the invalidity of the first assumption for high-precision data, which was also examined by Barstow et al. (2013), and for a critical examination of the second, we refer the reader to Rocchetto et al. (2016).

Fig. 12 shows our retrieved *P–T* profile. We highlight in red the near-infrared photosphere ($\tau = 1$ at 1.5μm), as this is the region predominately probed in transmission. Notably, the temperature is not constant across the photosphere, changing by ∼50 K. By assuming an isothermal profile, this behaviour, and its effect on molecular cross-sections, will not be captured. As expected intuitively, a tight constraint on the temperature is obtained at these altitudes ($T_{\rm phot} = 1221^{+131}_{-138}$ K). The confidence regions naturally expand away from the regions directly probed by the observations, particularly at pressures $\gtrsim$ 100 mbar, where we would usually expect the atmosphere to be opaque in slant geometry due to collisionally induced opacity alone. Our retrieved photospheric temperature is some 200 K colder than the planetary equilibrium temperature ($T_{\rm eq} \approx 1450$ K). This is unsurprising, given that transmission spectra probe high altitudes in the cooler terminator region. Again, we emphasize that such inferences are made possible by the high-precision transmission spectrum.

This profile represents the average terminator *P–T* profile. Strictly speaking, we expect this to be composed of two underlying profiles:

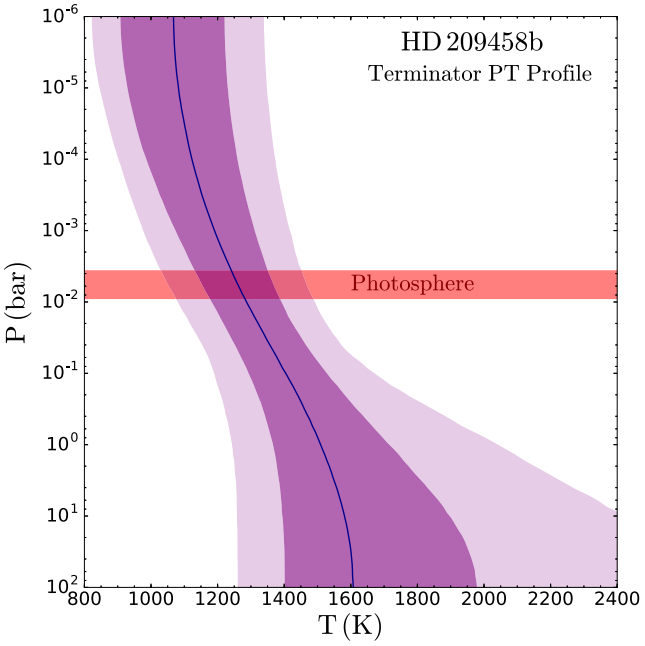



**Figure 12.** POSEIDON's retrieval of the terminator *P–T* profile on HD 209458b. The dark blue curve represents the median *P–T* profile, with the dark and light purple regions indicating the $1\sigma$ and $2\sigma$ confidence regions of the temperature in each layer (derived from 10 000 random sample draws from the posterior). The $1\sigma$ extent of the near-infrared photosphere ($\tau = 1$ at 1.5 μm) is shown by the red shaded region. The median temperature in the photosphere on the terminator, 1221 K, is approximately 200 K below the planetary equilibrium temperature, with the temperature changing by ∼50 K across the photosphere.

a cooler profile in the cloudy terminator region and a warmer profile in the clear region – as condensates tend to form in cooler regions, where the *P–T* profile may intersect additional condensation curves than in the warmer region. This averaging also explains the relatively high-altitude photosphere, as the opaque cloud deck we infer at $P_{\rm cloud} \approx 0.01$–$0.1$ mbar in the cloudy region combines with the cloud-free region to determine the $\tau = 1$ surface.

## 5 SUMMARY AND DISCUSSION

In this work, we have established a framework for retrieving properties of transiting exoplanet atmospheres with inhomogeneous clouds. As an initial demonstration, we applied our new atmospheric retrieval algorithm POSEIDON to the visible and near-infrared transmission spectrum of the hot Jupiter HD 209458b. Our major findings are as follows:

(i) We have found a potential avenue to break degeneracies between clouds and chemistry. Such degeneracies are artificially imposed by assuming one-dimensional cloud coverage and can be lifted by considering two-dimensional inhomogeneous cloud distributions. This enabled us to demonstrate precise determinations of the abundances of the prominent chemical species in a given spectral bandpass.

(ii) We report the first detection of nitrogen chemistry in an exoplanet atmosphere – established at $>3.7\sigma$ confidence. Both cloud-free and uniform cloud models identify $NH_3$ as the probable cause of nitrogen-induced absorption observed over the range 1.45–1.7 μm. The ammonia abundance is constrained to 0.01–2.7 ppm.

(iii) The $H_2O$ abundance on the terminator of HD 209458b is 30–100 × sub-solar (5–15 ppm). This is established by a fully Bayesian exploration of $\sim 10^8$ transmission spectra, including two-dimensional cloud/haze distributions.

(iv) Partially cloudy models are favoured by a Bayesian model comparison over both uniformly cloudy ($4.5\sigma$) and cloud-free ($5.4\sigma$) models. The terminator cloud fraction is constrained to $57^{+7}_{-12}$ per cent.

(v) Scattering due to high-altitude ($P < 0.05$ mbar) hazes is detected at $3.2\sigma$.

(vi) The terminator temperature structure of a transiting exoplanet can be constrained using high-precision *HST* transmission spectra. The temperature in the line-of-sight near-infrared photosphere is constrained to $1221^{+131}_{-138}$ K.

We have demonstrated that a wealth of information may be extracted from currently available high-precision *HST* transmission spectra. Our most interesting result is that non-uniform terminator properties may provide an avenue to extract *additional* information from exoplanet transmission spectra. Indeed, it appears that the consideration of inhomogeneous cloud coverage may have represented 'missing physics' that enabled us to break degeneracies in an otherwise underspecified problem. Exoplanet atmospheres are inherently multidimensional, and to treat them as one-dimensional will, at best, miss key insights and, at worse, incur unnecessary degeneracies.

### 5.1 Disequilibrium nitrogen chemistry

Given what is known from Solar system and brown dwarf studies, our detection of nitrogen chemistry in an exoplanet atmosphere should come as no surprise. In particular, ammonia is present both on Jupiter – in the form of high-altitude clouds (Sato & Hansen 1979) – and in brown dwarf atmospheres (e.g. Roellig et al. 2004; Saumon et al. 2006; Cushing et al. 2008). Recently, Line et al. (2015) demonstrated that $NH_3$ can be detected on brown dwarfs using low-resolution near-infrared data, though they noted a lack of obvious spectral features leading to their detection. We demonstrated in Figs 2 and 11 that the absorption features contributing to our $NH_3$ detection in the near-infrared are located over the ranges ∼1.45–1.7 and ∼1.2–1.3 μm. Additionally, the fact that Madhusudhan, Apai & Gandhi (2016a) reported sharp $NH_3$ abundance constraints in a population of three T-type brown dwarfs using WFC3 data alone further supports our detection. We also note that the nitrogen-bearing molecule HCN has been suggested, at low statistical significance, as a potential component in the atmosphere of the super-Earth 55 Cancri e (Tsiaras et al. 2016a). However, as we have noted in Section 4.2.1, our detection significance of nitrogen chemistry varies depending on the cloud model employed – with the lowest confidence being that for the partial cloud model at $3.7\sigma$. This sensitivity raises the possibility that additional physical mechanisms not considered in our models could produce a similar effect to that which we attribute to nitrogen chemistry. To adequately address this will require a concerted effort with both the development of increasingly sophisticated retrieval forward models alongside additional observations with sufficiently high precision to resolve the differences between models with and without nitrogen chemistry.

Inferring nitrogen chemistry and resulting abundance constraints heralds the opening of a new window into exoplanetary composition and atmospheric dynamics. The ammonia abundance we infer (≈1 ppm) using our most preferred model represents a $\gtrsim$100 times enhancement over the value expected of an atmosphere in thermochemical equilibrium with solar nitrogen abundance at our constrained temperature in the photosphere (Moses et al. 2011). This suggests that non-equilibrium processes may prove necessary in order to transport ammonia from regions where such high abundances can naturally form. One such avenue is transport-induced quenching, whereby regions where the characteristic dynamical time-scale ($\tau_{\rm dyn}$) is less than the chemical reaction time-scale ($\tau_{\rm chem}$) reflect the abundance from the 'quench' level where $\tau_{\rm dyn} = \tau_{\rm chem}$. For HD 209458b, $\tau_{\rm dyn} \sim 10^5$ s for both horizontal and vertical advection (Cooper & Showman 2006), whereas in the terminator photosphere ($P \sim 10$ mbar, $T \sim 1200$ K), $\tau_{\rm chem} \sim 10^{13}$ s for $NH_3 \rightarrow N_2$ conversion (Zahnle & Marley 2014). If horizontal quenching dominates, the $NH_3$ abundance would be expected to follow the chemical equilibrium value characteristic of the dayside, where $\tau_{\rm chem}$ is shorter (Agúndez et al. 2012). If, however, vertical quenching dominates, the $NH_3$ abundance in the observable photosphere will reflect that of the chemical equilibrium abundance at the altitude where $\tau_{\rm chem} \sim 10^5$, which occurs around pressures of ≈1 bar (Moses et al. 2011). Our abundance estimate of ≈1 ppm is remarkably consistent with that predicted at the terminator of HD 209458b by vertical quenching models (Moses et al. 2011) using nominal temperature and atmospheric mixing profiles from GCMs (Showman et al. 2009) and assuming a solar abundance of nitrogen. A wider range of parameters (e.g. mixing strengths, N abundance, etc.,) beyond those specifically assumed in the forward models of Moses et al. (2011) could also potentially explain the same. This suggests that $NH_3$ abundance constraints across a wide variety of exoplanets could provide a powerful diagnostic of the frequency and strength of non-equilibrium transport in a general context.

### 5.2 Implications for formation conditions

Recent years have seen increased interest in utilizing elemental ratios as formation diagnostics. In particular, the C/O ratio is

often invoked in attempts to constrain planetary formation and migration pathways relative to the snowlines of major condensates (e.g. Öberg et al. 2011; Madhusudhan et al. 2014c). Nitrogen chemistry offers complimentary diagnostics, as enhanced N/H ratios are anticipated for planets forming further out in protoplanetary discs (Piso, Pegues & Öberg 2016). Indeed, Piso et al. (2016) suggested that the N/O ratio for planets forming in outer discs could be significantly enhanced relative to the stellar value and to the C/O ratio.

Our present constraints on $H_2O$ and $NH_3$ suggest a scenario where the planet formed far out in the disc and migrated to its present location by dynamical scattering. The sub-solar $H_2O$ mixing ratio we find, despite the consideration of clouds/hazes, is inconsistent with an atmospheric composition of solar elemental abundances. Either a significantly sub-solar overall metallicity or supersolar C/O ratio are required to explain such low abundances (Madhusudhan et al. 2014a). On the other hand, as discussed above, the observed $NH_3$ abundance is consistent with non-equilibrium chemistry along with a nearly solar N abundance (Moses et al. 2011). Therefore, a consistent possible explanation for both the low $H_2O$ and high $NH_3$ abundance we observe is the presence of a supersolar C/O as well as a supersolar N/O ratio; the metallicity can be solar in all elements except O. This composition can be achieved if the planet formed beyond the $CO_2$ or CO snowlines, accretes mostly gas, (Öberg et al. 2011; Madhusudhan et al. 2014c; Piso et al. 2016) and migrates to its current orbit by disc-free mechanisms (Madhusudhan et al. 2014c) or formed via pebble accretion (Madhusudhan et al. 2016b). This is further supported by the fact that the host star HD 209458 is supersolar in metallicity, including O, which means that it would be infeasible to obtain such a low oxygen abundance in the planet if it migrated through the disc and accreted planetesimals (e.g. Brewer et al. 2016; Mordasini et al. 2016).

### 5.3 Solar versus sub-solar $H_2O$ estimates

Our robust demonstration that the terminator of HD 209458b is depleted in $H_2O$ relative to solar values runs contrary to the claim asserted by Sing et al. (2016). By not performing a retrieval, explicitly imposing thermochemical equilibrium, assuming isothermal *P–T* profiles and considering only global clouds/hazes, their models induce sufficient a priori biases to render their conclusions unreliable. Indeed, the inadequacy of this forward model approach has already been shown by Barstow et al. (2017), who performed a retrieval on the same data set and found a sub-solar $H_2O$ abundance (0.01–0.02) $\times$ solar in excellent agreement with ours and that of Madhusudhan et al. (2014b). Though our $H_2O$ abundances agree, the retrievals of Barstow et al. (2017) are somewhat limited by the usage of an optimal estimation algorithm, which explored only a limited volume of parameter space (3600 models versus our $10^8$) on a predefined grid of temperature profiles and cloud properties. Furthermore, the lack of marginalization over parameters or Bayesian evidence computation afforded by such an algorithm renders it impractical for parameter estimation or Bayesian model comparison, such as we have conducted here.

More sophisticated retrievals have also relied on making a priori model assumptions to break degeneracies between clouds and composition from transmission spectra. Benneke (2015), who also used a nested sampling algorithm like ours, claims a solar $H_2O$ abundance on the terminator of HD 209458b. However, their approach explicitly imposed a vast array of a priori physics: *P–T* profiles are not retrieved (radiative-convective equilibrium is assumed), C-N-O chemical pathways are enforced, and clouds, assumed to be composed of $MgSiO_3$, $MgFeSiO_4$ and SiC, are constructed using a model inspired by that of Ackerman & Marley (2001) and are uniform across the terminator. More recently, Line et al. (2016) attempted retrieval of the day-side atmospheric properties of HD 209458b using thermal emission spectra. They found a rather broad range in $H_2O$ abundance of (0.06–10) $\times$ solar, i.e. including substantially sub-solar as well as supersolar values, at $1\sigma$ confidence, on the dayside. However, as they point out, the inferences are hampered by an anomalously high $CO_2$ abundance which is strongly correlated with the $H_2O$ abundance. This is a well-recognized problem (Heng & Lyons 2016) in thermal emission retrievals, which future work needs to investigate.

More generally, the imposition of a priori assumptions has been used in retrievals of transmission spectra of several exoplanets where chemical and/or radiative equilibrium is enforced to derive elemental O and C abundances (e.g. Benneke 2015; Kreidberg et al. 2015). These approaches, more akin to forward models, undermine the ability of a retrieval to accomplish its fundamental goal: to infer the properties of an atmosphere with an absolute *minimal* set of assumptions. Succinctly put, we have shown that, in addition to clouds and hazes, a sub-solar $H_2O$ abundance at the terminator is essential to explain the low-amplitude spectral features of HD 209458b.

### 5.4 Cloud properties

Our inference of a partially cloudy atmosphere along HD 209458b's terminator compliments observations of inhomogeneous clouds in both the Solar system and brown dwarfs. On Earth and Jupiter, a banded cloud structure arises from atmospheric convection cells transporting gas parcels vertically, where clouds form upon crossing the relevant condensation curve, with the dry air carried to a different latitude where the formation of clouds is suppressed (de Pater & Lissauer 2001). A similar mechanism has been postulated to induce latitudinally inhomogeneous clouds in brown dwarfs (e.g. Marley, Saumon & Goldblatt 2010), with observational evidence also recently emerging (e.g. Buenzli et al. 2012).

Inhomogeneous cloud distributions have similarly been predicted to be common across the terminator region of hot Jupiters (Parmentier et al. 2016). The physical mechanism here is the day–night temperature contrast on tidally locked planets driving a superrotating equatorial jet, in turn, raising the temperature of the eastern terminator by hundreds of K above that of the cooler western terminator (Showman & Guillot 2002), where clouds are then more likely to form. Interestingly, we infer the properties of our cloudy region to consist of extremely high-altitude ($\approx$0.01–0.1 mbar) clouds with enhanced Rayleigh scattering above the deck. The temperatures at these altitudes are less than those at which photochemical hazes are expected to form ($\sim$1000–1100 K; see Zahnle, Marley & Fortney 2009; Moses 2014), so it is possible that the cloud deck we infer may be photochemical in origin. This possibility will require exploration by detailed photochemical models to explore its plausibility.

The clear terminator asymmetry in cloud properties on HD 209458b naturally raises the possibility of asymmetry in other observable properties. For example, Kataria et al. (2016) predicted that the eastern limb of HD 209458b should be warmer than the western limb by around 200 K and that $NH_3$ could be enhanced by an order of magnitude on the cooler western limb. The retrievals presented here do not consider such additional effects, with our present ability to disentangle the influence of clouds and chemistry contingent on both a sufficiently long spectral baseline (i.e. optical + near-infrared data) and the planet itself possessing a partially cloudy nature. Ultimately, the constraints derived by a retrieval algorithm are

specific to the framework of the assumed models, and it remains to be investigated if differences in limb *P–T* profiles and chemical abundances can be extracted using current or near-future transmission spectra.

Though future facilities, such as the *James Webb Space Telescope*, will undoubtedly revolutionize our understanding of exoplanet atmospheres, so much can still be accomplished with currently available high-precision *HST* spectra. Our results suggest that the key to obtaining precise chemical abundances from cloudy transmission spectra is rooted in the partially cloudy nature of their terminator; it is the stellar light transiting through the cloud-free region that facilitates breaking many apparent degeneracies between clouds and chemistry. Therefore, clouds need not be an insurmountable issue for sufficiently high-precision *HST* transmission spectra. Now, 8 years after the advent of exoplanet atmospheric retrieval, the time has come to move beyond one-dimensional models. As we enter the golden age of retrieval, the future is inherently two-dimensional.

## ACKNOWLEDGEMENTS

RJM would like to acknowledge financial support from the Science and Technology Facilities Council (STFC), UK, towards his doctoral programme. RM thanks Siddharth Gandhi for helpful discussions on the forward model. We acknowledge David Sing for making their spectral data publicly available, and the anonymous reviewer for their thoughtful comments on the manuscript.

## REFERENCES

Ackerman A. S., Marley M. S., 2001, ApJ, 556, 872
Agúndez M., Venot O., Iro N., Selsis F., Hersant F., Hébrard E., Dobrijevic M., 2012, A&A, 548, A73
Barman T., 2007, ApJ, 661, L191
Barstow J. K., Aigrain S., Irwin P. G. J., Bowles N., Fletcher L. N., Lee J.-M., 2013, MNRAS, 430, 1188
Barstow J. K., Aigrain S., Irwin P. G. J., Sing D. K., 2017, ApJ, 834, 50
Benneke B., 2015, ApJ, preprint (arXiv:1504.07655)
Benneke B., Seager S., 2012, ApJ, 753, 100
Benneke B., Seager S., 2013, ApJ, 778, 153
Brewer J. M., Fischer D. A., Valenti J. A., Piskunov N., 2016, ApJS, 225, 32
Buchner J. et al., 2014, A&A, 564, A125
Buenzli E. et al., 2012, ApJ, 760, L31
Charbonneau D., Brown T. M., Latham D. W., Mayor M., 2000, ApJ, 529, L45
Charbonneau D., Brown T. M. T., Noyes R. R. W. R., Gilliland R. L. R., 2002, ApJ, 568, 377
Christiansen J. L. et al., 2010, ApJ, 710, 97
Cooper C. S., Showman A. P., 2006, ApJ, 649, 1048
Crossfield I. J. M., 2015, PASP, 127, 941
Cushing M. C. et al., 2008, ApJ, 678, 1372
Dalgarno A., Williams D., 1962, ApJ, 136, 690
de Pater I., Lissauer J. J., 2001, Planetary Sciences, Cambridge Univ. Press, Cambridge
Deming D. et al., 2013, ApJ, 774, 95
Ehrenreich D. et al., 2014, A&A, 570, A89
Ehrenreich D. et al., 2015, Nature, 522, 459
Feroz F., Hobson M. P., 2008, MNRAS, 384, 449
Feroz F., Hobson M. P., Bridges M., 2009, MNRAS, 398, 1601
Feroz F., Hobson M. P., Cameron E., Pettitt A. N., 2013, preprint (arXiv:1306.2144)
Fortney J. J., 2005, MNRAS, 364, 649
Haynes K., Mandell A. M., Madhusudhan N., Deming D., Knutson H., 2015, ApJ, 806, 146
Hedges C., Madhusudhan N., 2016, MNRAS, 458, 1427
Heng K., Lyons J. R., 2016, ApJ, 817, 149
Kataria T., Sing D. K., Lewis N. K., Visscher C., Showman A. P., Fortney J. J., Marley M. S., 2016, ApJ, 821, 9
Knutson H. A., Benneke B., Deming D., Homeier D., 2014a, Nature, 505, 66
Knutson H. A. et al., 2014b, ApJ, 794, 155
Kreidberg L. et al., 2014, Nature, 505, 69
Kreidberg L. et al., 2015, ApJ, 814, 66
Lavie B. et al., 2016, preprint (arXiv:1610.03216)
Lecavelier des Etangs A., Pont F., Vidal-Madjar A., Sing D., 2008, A&A, 86, 83
Lee J. M., Fletcher L. N., Irwin P. G. J., 2012, MNRAS, 420, 170
Lee J.-M., Heng K., Irwin P. G. J., 2013, ApJ, 778, 97
Line M. R., Parmentier V., 2016, ApJ, 820, 78
Line M. R., Zhang X., Vaisht G., Chen P., Natraj V., Chen P., Yung Y. L., 2012, ApJ, 749, 93
Line M. R. et al., 2013, ApJ, 775, 137
Line M. R., Teske J., Burningham B., Fortney J., Marley M., 2015, ApJ, 807, 183
Line M. R. et al., 2016, AJ, 152, 203
Lupu R. E., Marley M. S., Lewis N., Line M., Traub W. A., Zahnle K., 2016, AJ, 152, 217
Madhusudhan N., Seager S., 2009, ApJ, 707, 24
Madhusudhan N. et al., 2011, Nature, 469, 64
Madhusudhan N., Knutson H., Fortney J., Barman T., 2014a, in Beuther H., Klessen R. S., Dullemond C. P., Henning T., eds, Protostars and Planets VI. Univ. Arizona Press, Tucson, AZ, p. 739
Madhusudhan N., Crouzet N., McCullough P. R., Deming D., Hedges C., 2014b, ApJ, 791, L9
Madhusudhan N., Amin M. A., Kennedy G. M., 2014c, ApJ, 794, L12
Madhusudhan N., Apai D., Gandhi S., 2016a, ApJ, preprint (arXiv:1612.03174)
Madhusudhan N., Bitsch B., Johansen A., Eriksson L., 2016b, MNRAS, preprint (arXiv:1611.03083)
Madhusudhan N., Agúndez M., Moses J. I., Hu Y., 2016c, Space Sci. Rev., 205, 285
Marley M. S., Saumon D., Goldblatt C., 2010, ApJ, 723, L117
Mordasini C., van Boekel R., Mollière P., Henning T., Benneke B., 2016, ApJ, 832, 41
Moses J. I., 2014, Philos. Trans. R. Soc. A, 372, 20130073
Moses J. I. et al., 2011, ApJ, 737, 15
Parmentier V., Fortney J. J., Showman A. P., Morley C., Marley M. S., 2016, ApJ, 828, 22
Piso A.-M. A., Pegues J., Öberg K. I., 2016, ApJ, 833, 203
Pont F., Sing D. K., Gibson N. P., Aigrain S., Henry G., Husnoo N., 2013, MNRAS, 432, 2917
Richard C. et al., 2012, J. Quant. Spectrosc. Radiat. Transf., 113, 1276
Rocchetto M., Waldmann I. P., Venot O., Lagage P. O., Tinetti G., 2016, ApJ, 833, 120
Roellig T. L. et al., 2004, ApJS, 154, 418
Rothman L. S. et al., 2010, J. Quant. Spectrosc. Radiat. Transf., 111, 2139
Sato M., Hansen J. E., 1979, J. Atmos. Sci., 36, 1133
Saumon D., Marley M. S., Cushing M. C., Leggett S. K., Roellig T. L., Lodders K., Freedman R. S., 2006, ApJ, 647, 552
Seager S., 2010, Exoplanet Atmospheres: Physical Processes, Princeton Univ. Press, Princeton
Seager S., Sasselov D. D., 2000, ApJ, 537, 916
Selke T., Bayarri J. J., Berger J. O., 2001, Am. Stat., 55, 62
Showman A. P., Guillot T., 2002, A&A, 385, 166
Showman A. P., Fortney J. J., Lian Y., Marley M. S., Freedman R. S., Knutson H. A., Charbonneau D., 2009, ApJ, 699, 564
Sing D. K. et al., 2016, Nature, 529, 59
Skilling J., 2004, in Fischer R., Preuss R., Toussaint U., eds, AIP Conf. Ser. Vol. 735, Bayesian Inference and Maximum Entropy Methods in Science and Engineering. Am. Inst. Phys., New York, p. 395
Snellen I. A. G., Albrecht S., de Mooij E. J. W., Le Poole R. S., 2008, A&A, 487, 357
Snellen I. A. G., de Kok R. J., de Mooij E. J. W., Albrecht S., 2010, Nature, 465, 1049

Stevenson K. B. et al., 2014, Science, 346, 838
Swain M. R. et al., 2009, ApJ, 704, 1616
Tennyson J. et al., 2016, J. Mol. Spectrosc., 327, 73
Trotta R., 2008, Contemp. Phys., 49, 71
Tsiaras A. et al., 2016a, ApJ, 820, 99
Tsiaras A., Waldmann I. P., Rocchetto M., Varley R., Morello G., Damiano M., Tinetti G., 2016b, ApJ, 832, 202
Waldmann I. P., Tinetti G., Rocchetto M., Barton E. J., Yurchenko S. N., Tennyson J., 2015a, ApJ, 802, 107
Waldmann I. P., Rocchetto M., Tinetti G., Barton E. J., Yurchenko S. N., Tennyson J., 2015b, ApJ, 813, 13
Zahnle K. J., Marley M. S., 2014, ApJ, 797, 41
Zahnle K., Marley M. S., Fortney J. J., 2009, preprint (arXiv:0911.0728)
Öberg K. I., Murray-Clay R., Bergin E. A., 2011, ApJ, 743, L16

## APPENDIX A: DERIVATION OF THE TWO-DIMENSIONAL TRANSIT DEPTH

For completeness, here we present a concise derivation of the wavelength-dependent transit depth (equation 1) for a two-dimensional atmosphere. The transit depth is defined as the fractional stellar flux difference induced as a planet transits its parent star:

$$\Delta_\lambda = \frac{F_{\lambda,\,\mathrm{out}} - F_{\lambda,\,\mathrm{in}}}{F_{\lambda,\,\mathrm{out}}}, \tag{A1}$$

where $F_{\lambda,\,\mathrm{out}}$ is the spectral flux observed outside of the transit and $F_{\lambda,\,\mathrm{in}}$ is the spectral flux observed during the transit. The fluxes are, in turn, defined as integrals of the intensity over solid angle, projected in the observer's line of sight (Seager 2010):

$$F_\lambda = \int_\Omega I_\lambda\, \hat{\boldsymbol{n}} \cdot \hat{\boldsymbol{k}}\, \mathrm{d}\Omega, \tag{A2}$$

where $\hat{\boldsymbol{n}}$ is a unit vector in the direction of propagation of a beam, $\hat{\boldsymbol{k}}$ is the unit vector in the direction of the observer, and $\Omega$ is the solid angle subtended at the observer.

The flux outside of transit is solely due to the star. From the geometry in Fig. 1, we see that $\hat{\boldsymbol{n}} \cdot \hat{\boldsymbol{k}} = 1$ and $\mathrm{d}\Omega = \mathrm{d}A/D^2 = b\,\mathrm{d}\phi\mathrm{d}b/D^2$. Thus, we have

$$F_{\lambda,\mathrm{out}} = \int_0^{2\pi}\int_0^{R_*} I_{\lambda,*}\left(\frac{b}{D^2}\right)\mathrm{d}b\mathrm{d}\phi = \frac{1}{2\pi}\int_0^{2\pi}\pi I_{\lambda,*}\left(\frac{R_*}{D}\right)^2\mathrm{d}\phi, \tag{A3}$$

where $I_{\lambda,*}$ is the intensity at the stellar photosphere. In carrying out the radial integral in the second equality, we have implicitly assumed that the stellar intensity is uniform as a function of the impact parameter. This is valid as non-uniform stellar discs, due to limb darkening, are accounted for during data reduction.

The flux during the transit consists of three components: flux from the planet itself, stellar flux transmitting through the planetary atmosphere and stellar flux from the disc of the host star surrounding the planet:

$$\begin{aligned}
F_{\lambda,\,\mathrm{in}} &= F_{\lambda,\,\mathrm{P}} + F_{\lambda,\,\mathrm{A}} + F_{\lambda,\,*},\\
F_{\lambda,\,\mathrm{P}} &= \frac{1}{2\pi}\int_0^{2\pi}\int_0^{R_\mathrm{p}} I_{\lambda,\mathrm{P}}\left(\frac{2\pi b}{D^2}\right)\mathrm{d}b\,\mathrm{d}\phi,\\
F_{\lambda,\,\mathrm{A}} &= \frac{1}{2\pi}\int_0^{2\pi}\int_0^{R_\mathrm{p}+H_\mathrm{A}} I_{\lambda,*}(\tau_\lambda(b,\phi))\left(\frac{2\pi b}{D^2}\right)\mathrm{d}b\,\mathrm{d}\phi,\\
F_{\lambda,\,*} &= \frac{1}{2\pi}\int_0^{2\pi}\int_{R_\mathrm{p}+H_\mathrm{A}}^{R_*} I_{\lambda,*}\left(\frac{2\pi b}{D^2}\right)\mathrm{d}b\,\mathrm{d}\phi.
\end{aligned} \tag{A4}$$

The first integral will evaluate to zero, as the thermal intensity from the planetary nightside, $I_{\lambda,\mathrm{P}}$, is negligible compared to the stellar flux. $I_{\lambda,*}(\tau_\lambda(b,\phi))$ is the attenuated stellar intensity at impact parameter $b$ and azimuthal angle $\phi$ following passage through a region of the atmosphere with optical depth $\tau_\lambda(b,\phi)$. Note that the lower impact parameter limit of the second integral is set to zero to account for the possibility of rays of a given wavelength passing through the planet below $R_\mathrm{p}$. Where this may occur is dictated by the local conditions (particularly clouds) at a given azimuthal angle. Regions with low opacity will possess a much deeper opaque radius. Accounting for this requires full evaluation over all impact parameters, without assuming a common lower value for all wavelengths. The third integral is readily evaluated:

$$F_{\lambda,\,*} = \frac{1}{2\pi}\int_0^{2\pi}\frac{\pi I_{\lambda,*}}{D^2}\left(R_*^2 - (R_\mathrm{p}+H_\mathrm{A})^2\right)\mathrm{d}\phi. \tag{A5}$$

To evaluate the second integral, we require an expression for $I_{\lambda,*}(\tau_\lambda(b,\phi))$. We make use of the fact that, in the case of transmission of stellar radiation, there will be no emission or scattering into the beam. This renders the solutions of the equation of radiative transfer into simple exponential attenuation, given by Beer's law: $I_\lambda(\tau_\lambda) = I_\lambda(0)\,\mathrm{e}^{-\tau_\lambda}$. Substituting this into the second integral in (A4) yields

$$F_{\lambda,\,\mathrm{A}} = \frac{1}{2\pi}\int_0^{2\pi}\int_0^{R_\mathrm{p}+H_\mathrm{A}}\frac{2\pi I_{\lambda,*}}{D^2}\left(b\,\mathrm{e}^{-\tau_\lambda(b,\phi)}\right)\mathrm{d}b\,\mathrm{d}\phi. \tag{A6}$$

Substituting this and equation (A5) into equation (A4) gives

$$\begin{aligned}
F_{\lambda,\,\mathrm{in}} = \frac{1}{2\pi}\int_0^{2\pi}\frac{\pi I_{\lambda,*}}{D^2}\Bigg(&R_*^2 - R_\mathrm{p}^2 - 2R_\mathrm{p}H_\mathrm{A} - H_\mathrm{A}^2\\
&+2\int_0^{R_\mathrm{p}+H_\mathrm{A}} b\mathrm{e}^{-\tau_\lambda(b,\phi)}\mathrm{d}b\Bigg)\mathrm{d}\phi.
\end{aligned} \tag{A7}$$

Noting that $2\int_{R_\mathrm{p}}^{R_\mathrm{p}+H_\mathrm{A}} b\,\mathrm{d}b = 2R_\mathrm{p}H_\mathrm{A} + H_\mathrm{A}^2$ and splitting the integral in equation (A7) gives

$$\begin{aligned}
F_{\lambda,\,\mathrm{in}} = \frac{1}{2\pi}\int_0^{2\pi}\frac{\pi I_{\lambda,*}}{D^2}\Bigg\{&R_*^2 - R_\mathrm{p}^2 + 2\int_0^{R_\mathrm{p}} b\mathrm{e}^{-\tau_\lambda(b,\phi)}\mathrm{d}b\\
&+2\int_{R_\mathrm{p}}^{R_\mathrm{p}+H_\mathrm{A}} b\left(\mathrm{e}^{-\tau_\lambda(b)} - 1\right)\mathrm{d}b\Bigg\}\mathrm{d}\phi.
\end{aligned} \tag{A8}$$

Finally, the flux during and outside transit can be substituted into equation (A1) to determine the transit depth

$$\Delta_\lambda = \frac{1}{2\pi}\int_0^{2\pi}\delta_\lambda(\phi)\mathrm{d}\phi, \tag{A9}$$

where $\delta_\lambda(\phi)$ is the 'one-dimensional' transmission spectrum that would result from assuming an axially symmetric atmosphere with the same properties as the two-dimensional atmosphere possesses at polar angle $\phi$:

$$\delta_\lambda = \frac{R_\mathrm{p}^2 + 2\int_{R_\mathrm{p}}^{R_\mathrm{p}+H_\mathrm{A}} b\left(1 - \mathrm{e}^{-\tau_\lambda(b,\phi)}\right)\mathrm{d}b - 2\int_0^{R_\mathrm{p}} b\mathrm{e}^{-\tau_\lambda(b,\phi)}\mathrm{d}b}{R_*^2}. \tag{A10}$$

As an additional geometric interpretation, we note that this expression for the one-dimensional transit depth is equivalent to taking a perfectly opaque disc of radius $R_\mathrm{p} + H_\mathrm{A}$ and subtracting the transmitted stellar light. The transmission is simply the integral of the relative area of successive annuli (compared to that of the disc), each weighted by $\mathrm{e}^{-\tau_\lambda}$.

This paper has been typeset from a TEX/LATEX file prepared by the author.