Open Research Online

Outstanding Challenges of Exoplanet Atmospheric Retrievals

Joanna K. Barstow · Kevin Heng

Abstract Spectral retrieval has long been a powerful tool for interpreting planetary remote sensing observations. Flexible, parameterised, agnostic models are coupled with inversion algorithms in order to infer atmospheric properties directly from observations, with minimal reliance on physical assumptions. This approach, originally developed for application to Earth satellite data and subsequently observations of other Solar System planets, has been recently and successfully applied to transit, eclipse and phase curve spectra of transiting exoplanets. In this review, we present the current state-of-the-art in terms of our ability to accurately retrieve information about atmospheric chemistry, temperature, clouds and spatial variability; we discuss the limitations of this, both in the available data and modelling strategies used; and we recommend approaches for future improvement.

1 Introduction

Remote sensing of exoplanet atmospheres is a rapidly expanding field, having progressed from the first detections of a molecular species in an atmosphere around another star (Barman 2007; Tinetti et al. 2007) to beginning to characterize complex temperature structures, clouds, and spatial heterogeneity in

J. K. Barstow
School of Physical Sciences, The Open University, Walton Hall, Kents Hill, Milton Keynes, MK7 6AA, UK
Department of Physics and Astronomy, University College London, Gower Street, London, WC1E 6BT, UK
E-mail: jo.barstow@open.ac.uk

K. Heng
Center for Space and Habitability, University of Bern, Gesellschaftsstrasse 6, CH-3012 Bern, Switzerland.

just over ten years. Retrieval methods – iteratively comparing synthetic to observed spectra in order to infer the most likely atmospheric state – were historically applied to Solar System atmospheres, and with some adaptation are now being used to analyze exoplanet spectra. Typically, retrieval codes couple a simple, parameterised, 1D radiative transfer model to a retrieval algorithm. The model parameters form the atmospheric state vector, and the output from the retrieval algorithm is a posterior probability distribution for each element in the state vector, including correlations between the model parameters.

This review paper, rather than simply presenting an overview of the current state of the art, instead discusses what we see as the major challenges facing exoplanet retrievals over the next few years, and thus the directions in which we expect development to be most rapid. In general, all of these challenges can be summarized as resolving the tension between model realism (with risks of overfitting or allowing informative priors to drive solutions) and model simplicity (with the risk that the model may be inadequate to accurately reproduce the data, or may reproduce them for the wrong reasons, and may be very far from the truth). For each challenge, we present the current status, and then provide our recommendations for future routes of exploration and improvement. The key areas which we have identified are listed below:

1. Inferring chemistry from measured molecular abundances
2. Representation of temperature structure
3. Representation of clouds and aerosols
4. Including 3D effects in 1D models

These areas will be dealt with in turn from Section 2.

1.1 Retrieval algorithms

A range of algorithms and retrieval codes have been applied to exoplanet retrievals, with each approach having different benefits. Figure 1 shows the basic structure of a retrieval code. The earliest exoplanet retrievals used either a simple grid search (e.g. Madhusudhan and Seager 2009) or Optimal Estimation (Rodgers 2000; Irwin et al. 2008). Grid searches are simple to set up, but can be inefficient (since they may involve a detailed exploration of parameter space far from the solution) and results will be highly restricted by the parameter values included within the grid. Optimal Estimation is a matrix inversion method that assumes Gaussian statistics. A Levenberg-Marquardt scheme is used to iteratively solve the inverse problem and works to minimize a cost function, which assesses both the difference between the model output and the measured spectrum and also the distance of the atmospheric state vector from a Gaussian prior state vector. Due to its imposition of Gaussianity, whilst Optimal Estimation is fast and efficient it is unable to a) effectively explore multimodal parameter spaces and b) explore a very broad parameter space, as the parameter ranges are restricted by the necessity of including a Gaussian prior constraint.

Markov-chain Monte Carlo (MCMC; see e.g. Line et al. 2013) and nested sampling algorithms have more recently become the preferred tools within the community. These Bayesian approaches both allow a more comprehensive exploration of the parameter space, as they do not restrict priors or posteriors to obeying Gaussian statistics. Of these approaches, the `MultiNest` (Feroz and Hobson 2008; Feroz et al. 2009, 2013) implementation of the nested sampling method (Skilling 2006) has proved especially popular, as this provides a relatively efficient exploration of potentially multi-modal posteriors that effectively captures multiple modes and complex degeneracies.

With standard retrieval methods, there is always tension between achieving physical and chemical realism versus completing the calculation within a reasonable period of time. In practice, this means that the forward models converting chemical abundances and opacities into transit radii and fluxes need to be simplified to enable rapid computation. Recently, machine learning approaches have been adapted to performing atmospheric retrieval, which allows the burden of computing synthetic spectra to be shifted offline. A grid of models may be computed beforehand and used as a training set for the supervised machine learning method of choice. This approach has been demonstrated using regression trees and random forests (Márquez-Neila et al. 2018). It allows model grids from different research groups to be used for atmospheric retrieval, even if the computer codes used to generate these models are proprietary and non-public. Nevertheless, paying attention to model assumptions remains a key part of the process. The unsupervised machine learning method of deep convolutional generative adversarial networks has also been implemented for atmospheric retrieval (Zingales and Waldmann 2018).

1.2 Parameterised 1D models

A general requirement for the majority of retrieval schemes is for the forward model computation to be fast. This is especially necessary for Monte Carlo and nested sampling methods, as these typically require millions of individual forward models to be computed to adequately explore the parameter space. Therefore, forward models must be relatively simple; instead of containing detailed physics, models are usually parameterised, and are generally also one-dimensional. Parameterisation must be approached with care, as simple models uncoupled from physical assumptions may be prone to converging on unrealistic solutions (e.g. atmospheres with implausible chemistry, or temperature-pressure profiles that would be unstable). However, this potential disadvantage can also be a strength in situations where our understanding of the underlying physics and chemistry is still relatively immature, as it can prevent incorrect prior assumptions from driving the solution.

The parameterised approach is especially useful in the context of exoplanet retrievals because the information content of data is continuously changing, and the complexity of parameterised models can very easily be tuned. For example, the early exoplanet retrieval model of Madhusudhan and Seager (2009)
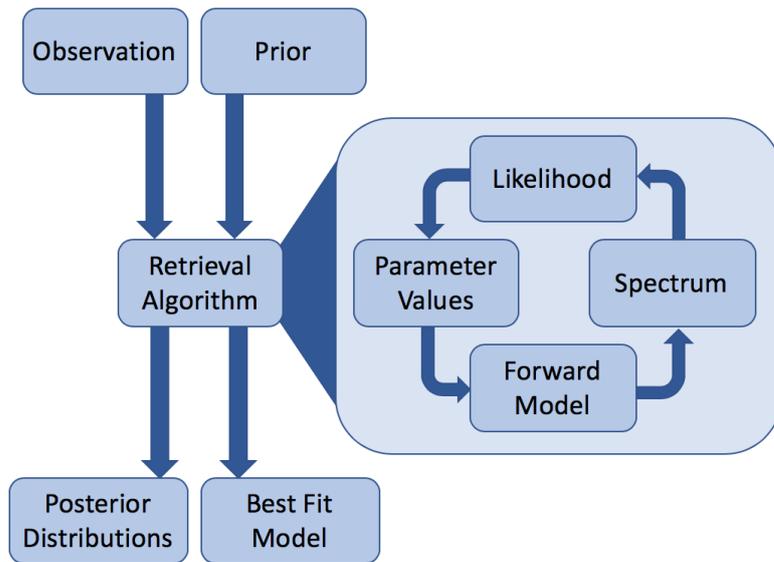
Fig. 1 Schematic showing the basic structure of a retrieval algorithm.

contained a six-parameter temperature-pressure profile, which effectively divided the atmosphere into three layers and described the temperature gradient within each layer. They also retrieved altitude-independent abundances of H_2O, CO_2, CO, CH_4 and NH_3, which were the five species they considered to be most likely to be active in the infrared. They found that they were unable to simultaneously fit the data from different instruments with the same model. A subsequent analysis by Lee et al. (2012) allowed the temperature to vary freely and smoothly as a function of pressure, which allowed a reasonable fit to be achieved to all datasets, but clearly included greater potential for model degeneracy due to the increased number of parameters. Lee et al. (2012) present correlations between the temperature-pressure profile and the abundances of the molecular species, demonstrating the extent of this degeneracy.

Parameterisation has also evolved in modelling of primary transit spectra. The different geometries of primary and secondary transit observations mean that each is sensitive to different aspects of the atmospheric state, and so different parameters are included depending on the type of observation.

The transit depth in primary transit is given by

$$\Delta_\lambda = 1 - \left(\frac{R_{p,\lambda}}{R_s} \right)^2 \tag{1}$$

where $R_{p,\lambda}$ is the radius of the planet and R_s the radius of the star. A transit spectrum is the variation in transit depth as a function of wavelength, which results from the change in atmospheric opacity due to the presence of absorbing gases and aerosols (Figure 2).
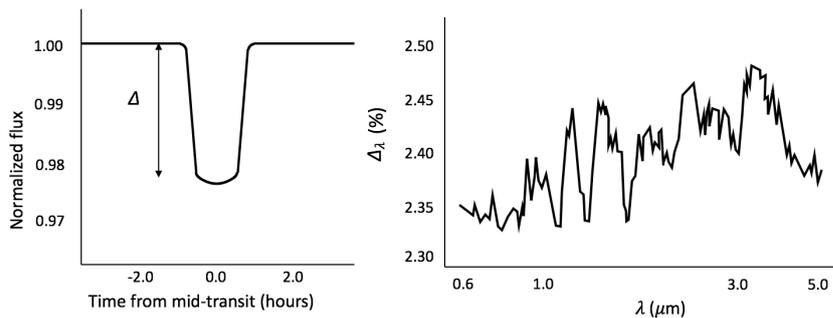
Fig. 2 Sketches showing a transit lightcurve and the corresponding transit spectrum.

By contrast, the signal in secondary eclipse is obtained by measuring the difference in flux immediately before and after the eclipse, when the dayside of the planet is visible, with the flux of the star alone when the planet is in eclipse.

Primary transit observations do not probe deeper regions of the atmosphere, as the atmosphere becomes opaque to radiation passing tangentially through the atmosphere at lower pressures compared with radiation emerging close to nadir. Transit spectra also solely measure light from the star that has passed through the atmosphere rather than thermal emission from the planet itself; therefore, primary transit spectra are less sensitive than secondary eclipse spectra to the temperature-pressure profile. Retrievals covering only a small spectral range in primary transit have therefore generally assumed an isothermal temperature-pressure profile. By contrast, primary transit spectra are extremely sensitive to the atmospheric scale height H, as it is the physical thickness of the atmosphere that determines the amplitude of the spectral features in a primary transit observation.

The pressure $p(z)$ of an atmosphere in hydrostatic equilibrium may be defined as:

$$p(z) = p(0)e^{\frac{-z}{H}} \tag{2}$$

where $p(0)$ is the surface pressure and H is the atmospheric scale height.

$$H = \frac{kT}{\mu g} \tag{3}$$

where k is the Boltzmann constant, T is temperature, μ is the mean molecular mass, and the local gravitational acceleration is

$$g = \frac{GM_P}{(R_P + z)^2}. \tag{4}$$

Here, G is the universal gravitational constant, M_P is the mass of the planet, R_P is the radius of the planet and z is the altitude above the surface. The dependence of the scale height on g means that there is significant sensitivity to the absolute radius of the planet (as opposed to the radius relative to that

of the star). This emphasises the requirement for precise and accurate radii for planet host stars.

This in turn impacts what is variously referred to as the normalisation degeneracy or the baseline issue (Benneke and Seager 2012; Griffith 2014; Heng and Kitzmann 2017); because the planetary radius quoted in the literature is derived from the white light transit, the pressure that this represents is dependent on the atmospheric properties. Either the pressure at some given radius, or the radius at some given pressure, must therefore also be free parameters in the retrieval. Because the scale height is then proportional to both temperature and the square of the radius, these two quantities are degenerate and are inversely correlated in retrievals. Fisher and Heng (2018) demonstrated that the normalisation degeneracy may be partially broken using low-resolution transmission spectra measured by *Hubble*-WFC3 alone, because information on temperature and chemical abundances are encoded in the shape of the transmission spectrum. However, this degeneracy and others are more easily broken by including broad wavelength coverage data, as discussed in Section 2.1.

Primary transit spectra are also affected by the presence of clouds. Effects can be dramatic to the point of cloud obscuring all molecular and atomic features in the spectrum (e.g. Kreidberg et al. 2014). In less extreme cases, the amplitudes of gas absorption features are reduced in the presence of cloud because the atmosphere becomes opaque below the cloud top, so only the centres of molecular bands are observed. This effect can be difficult to distinguish from either a) low abundances of the molecular species in question or b) a high mean molecular mass (and therefore low scale height) atmosphere.

Simple 1D forward models for retrieval codes need to include parameterisations of these effects. Cloud is often treated as a completely grey, opaque layer with a variable top pressure (e.g. Kreidberg et al. 2014). This approach has the advantage of introducing only a single parameter, but is also not very representative of a real cloud, which is likely to have a wavelength-dependent optical depth and to be partially transparent at some wavelengths. We discuss cloud parameterisation in more detail in Section 4.1. The mean molecular mass may be specified as a separate free parameter, or may be calculated after the fact based on the retrieved abundances of the modelled gases; this approach is computationally simpler, but risks misinterpretation should large abundances of a spectrally inactive, heavy gas such as N_2 are present. It also relies on a complete range of molecular species being included in the model.

2 Chemistry

In this section, we discuss the challenges of inferring information about chemistry, and thence planetary formation and origin scenarios, from the retrieved abundances of individual gases. We begin by summarising the current state of the art. Whilst there is a wealth of literature available dealing with detailed studies of individual planets, here we find it is more instructive to focus on works that analyse multiple planets, as this provides a more general indica-

tion of the degree to which atmospheric properties can be constrained with currently available data.

2.1 State of the art: chemistry

Hot Jupiters observed in primary transit are ideal targets for molecular species detection and constraint, as these planets have large scale heights and therefore large feature amplitudes in primary transit (in the absence of clouds). Several comparative retrieval studies of hot Jupiters with *Hubble Space Telescope* and *Spitzer* observations have been recently performed, following on from the presentation by Sing et al. (2016) of near-infrared spectra of ten hot Jupiters with consistent data reduction.

Hubble Wide Field Camera 3 (WFC3) data are now available for several tens of exoplanets. Many of these also have photometry from the *Spitzer* InfraRed Array Camera (IRAC) and spectra from the *Hubble* Space Telescope Imaging Spectrograph (STIS). As WFC3 spectra are the most widely available, studies such as Tsiaras et al. (2018) and Fisher and Heng (2018) focus on this dataset only. Because WFC3 has a relatively narrow wavelength range, between 0.8 and 1.6 µm, only a subset of interesting molecular species can be constrained. The 1.4 µm H_2O band dominates the spectral shape in this region, although features from TiO, VO and FeH may be discernable at the shorter wavelength end if present, along with CH_4, HCN and NH_3 longwards of 1 µm.

Tsiaras et al. (2018) use a 10-parameter model to study 30 hot and warm gaseous planets, including volume mixing ratios of H_2O, CO_2, CO, CH_4 and NH_3; isothermal temperature; planet radius; and three cloud parameters (discussed further in Section 4.1. For planets hotter than 1400 K they also include TiO and VO abundances. They define an atmospheric detectability index (ADI) which is the Bayes factor between the nominal atmospheric model and a straight line (featureless) spectrum, and they class any planet with ADI >3 as having a detectable atmosphere. They find that 16 of the 30 planets studied fulfil this criterion; H_2O is found to be present on all of these planets, with abundances typically constrained to ± an order of magnitude. No constraints are obtained for CO_2, CO, CH_4 or NH_3 on any planet, but for two (WASP-76b and WASP-121b) there is evidence that TiO and VO are present; a subsequent analysis including STIS data for WASP-121b by Evans et al. (2018) corroborates the presence of VO but not of TiO.

Fisher and Heng (2018) examine a similar dataset of 38 WFC3 transmission spectra, although their analysis extends to smaller and temperate planets, such as the warm mini-Neptune GJ 1214b and the likely rocky earth-sized planets TRAPPIST 1d–g. Unlike Tsiaras et al. (2018), they only consider volume mixing ratios of H_2O, NH_3 and HCN in their model. They include a slightly more complex cloud parameterisation (see Section 4.1) and allow for a non-isothermal temperature profile. They retrieve a reference pressure rather than a reference radius for the planet. Fisher and Heng (2018) find no evidence that

the region of the atmosphere probed during transit deviates from an isothermal profile, and they conclude that most of these spectra may be explained by an isothermal transit chord containing only water and grey clouds.

Two further studies, Barstow et al. (2017) and Pinhas et al. (2019), consider a smaller number of planets but take into account data from *Hubble*/STIS and *Spitzer*/IRAC. A broader wavelength range allows degeneracies between cloud properties and gas abundances to be broken, but this coverage is not available for as many planets, and the inclusion of spectral segments obtained at different times introduces the issue of stitching together non-contemporaneous spectra that may have been affected by instrumental and astrophysical systematics in different ways. For this reason, the datasets used are those provided by Sing et al. (2016), in which spectra were consistently reduced in an attempt to minimise this issue.

Barstow et al. (2017) uses a hybrid approach, combining the fast but prior-restricted optimal estimation retrieval method with a grid search to ensure exploration of a wide parameter space. Gases included in the retrieval are H_2O, CO_2, CO, CH_4, but there is no evidence for the presence of any gas except H_2O. Constraints on H_2O abundance are obtained for all planets except WASP-12b, which has poor quality WFC3 data in the Sing et al. (2016) paper, and WASP-6b and WASP-39b, for which no WFC3 data were available at the time. H_2O abundances are constrained only to within an order of magnitude, but show a clear trend towards subsolar abundances. This trend was also found by Pinhas et al. (2019), who performed a nested sampling retrieval of the same dataset; Pinhas et al. (2019) used new WFC3 data for WASP-12b and WASP-39b, which allowed constraints on H_2O abundance for these planets also.

Although Barstow et al. (2017) and Pinhas et al. (2019) use different cloud parameterisations, the H_2O abundance results are consistent with each other where the same data are used. The differing results for the cloud properties are discussed further in Section 4.1.

In Figure 3, we present a comparison of the retrieved H_2O abundances for each of the studies described above. The values shown for Barstow et al. (2017) are taken from the range of values from the best-fitting models for each planet; the central value shown is just the average of the minimum and maximum. All other values are obtained directly from the marginalised retrieval solution in each case. In general, retrievals accounting for *Hubble*/STIS and *Spitzer*/IRAC data converge on lower H_2O abundances, whereas solutions from just *Hubble*/WFC3 have higher H_2O abundances. Pinhas et al. (2019) conclude that H_2O abundances are generally subsolar. Error-weighted averages are shown, calculated over all available planets except for WASP-6b, for which no WFC3 data are available. The very low H_2O volume mixing ratio retrieved for WASP-6b from Pinhas et al. (2019) is likely to be a result of a substantial drop in transit depth between the STIS spectrum and the the IRAC points in the infrared, which forces a scenario in which the spectrum is characterised by opaque haze and low gas abundances. Averages between Barstow et al. (2017) and Fisher and Heng (2018) differ by more than two orders of magnitude, indicating that the inclusion of STIS and IRAC data is influential on the solution.
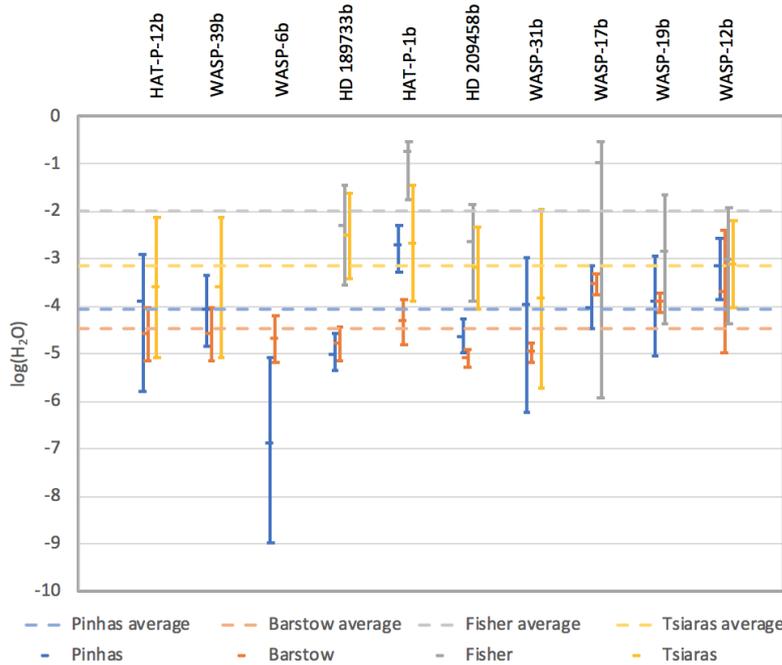
Fig. 3 A comparison of H_2O volume mixing ratios retrieved using four different retrieval algorithms. Pinhas et al. (2019) and Barstow et al. (2017) use spectra combining *Hubble*/STIS and WFC3, and *Spitzer*/IRAC, whereas Fisher and Heng (2018) and Tsiaras et al. (2018) use only *Hubble*/WFC3. The dashed lines represent the error-weighted average abundances, excluding WASP-6b for which no WFC3 data is available.

The difference between the retrieval results with and without STIS and IRAC is most apparent for HD 189733b and HD 209458b.

The likely reason for these differences when broader wavelength coverage data are added is that these data provide more constraints on cloud characteristics than WFC3 does by itself; muted H_2O features can either be the result of a low abundance of H_2O, or the presence of cloud. The detection of absorption features due to multiple gases can also break degeneracies between temperature and gas abundance, and low-amplitude features can also be a result of low temperatures. Following this logic, we expect to see substantial improvements with the launch of *JWST*, which will provide extremely broad wavelength coverage (although it cannot cover the full spectral range simultaneously).

Primary transit observations are generally preferred for obtaining constraints on molecular species abundance, but Line et al. (2014) completed a comparative study of 9 hot Jupiters in emission and present retrieval results for molecular species. The spectral coverage and resolving power is extremely variable across the 9 objects, with HD 189733b having data from *Hubble*/Near Infrared Camera and Multi-Object Spectrograph (NICMOS), *Spitzer*/IRAC

and *Spitzer*/InfraRed Spectrograph (IRS), whereas the majority are restricted to photometric observations only. Good constraints on molecular abundances (beyond upper and lower limits) are generally only obtained for cases with spectroscopic data. In this case, volume mixing ratios for H_2O, CO_2 and CH_4 are constrained to within an order of magnitude for HD 189733b, and H_2O is similarly constrained for TRES-3b, but no further strong constraints are obtained for any of the other planets in the sample.

Subsequent publications looking at single planets in emission have obtained some constraints on molecular abundances. Stevenson et al. (2014a) analyse the dayside spectrum of ultra-hot Jupiter WASP-12b. They test oxygen-rich (C:O ~ 0.5) and carbon-rich (C:O $>\sim 1.0$) atmospheric models, and find that the carbon-rich model is preferred, although their best-fit solution has what the authors consider to be implausibly high abundances of CH_4 and CO_2, and very low abundances of H_2O. An analysis of the same dataset by Oreshenko et al. (2017) shows that the solution is highly dependent on prior assumptions made about the chemistry. Heng and Lyons (2016) point out that it is nearly impossible to have CO_2 be more abundant than CO in H_2-dominated atmospheres unless the metallicity exceeds solar by about 3 orders of magnitude, and this constraint should be used to rule out chemically implausible retrieval solutions. Some evidence for the presence of TiO and VO has also been reported from secondary eclipse observations, of WASP-33b (by Haynes et al. 2015) and WASP-121b (VO only, by Evans et al. 2017).

Molecular abundance information from secondary eclipse spectra lags behind that available from transits, as secondary eclipse contrast improves at wavelengths beyond the reach of *Hubble*, and the lack of cryogenic cooling for *Spitzer* means that currently precise secondary eclipse spectra are hard to come by. This situation is expected to improve enormously once *JWST* has launched. Despite significant advances in spectral quality for both primary transit and secondary eclipse over the last decade, precise abundance constraints are only reliably available for H_2O, and even this is not universally possible. The main barrier to molecular species constraint is the typically narrow wavelength range accessible for most planets; wavelengths beyond the red end of the *Hubble*/WFC3 G141 grism ($>\sim 1.6$ μm) are required to constrain most molecular species apart from H_2O and metal oxides/hydrides, and spectral data in this range is currently unavailable. This situation will be vastly improved once *JWST* has launched, as it will improve signal-to-noise and resolving power by at least a factor of 10, and push spectral coverage further into the infrared. Several predictive studies exist that indicate *JWST* spectra will provide excellent opportunities for retrieval constraints on molecular abundances from both primary transit and secondary eclipse spectra of hot Jupiters (e.g. Barstow et al. 2015; Greene et al. 2016) and also allow the characterization of smaller, terrestrial worlds (e.g. Barstow and Irwin 2016, Krissansen-Totton et al. 2018).

2.2 Recovery of underlying chemical trends

A key part of the planetary formation/evolution puzzle is the bulk C:O ratio of a planet. It has been postulated that this is an indicator of where in the disc a planet has formed (Öberg et al. 2011) as the location of the planet relative to the snowlines could affect the composition of the accreted material. Oreshenko et al. (2017) attempted this exercise for WASP-12b using an emission spectrum constructed from *Hubble*-WFC3 and *Spitzer*-IRAC, and suggested that WASP-12b experienced disk-free migration during its formation history. Determining the bulk C:O ratio from spectroscopy has already been attempted in exoplanet retrievals (e.g. Line et al. 2014; Kreidberg et al. 2015), although so far this is hampered by a lack of access to regions of the spectrum containing features of carbon species. Observations by *JWST* will alleviate this aspect of the problem, but the question remains to what degree of precision underlying chemical trends such as the C:O ratio can be recovered. This is particularly important in the context of future missions such as *ARIEL*, which aims to provide the first exoplanet atmosphere population study.

Kreidberg et al. (2015) compare retrievals with free chemistry (where each gas is retrieved individually) and retrievals of metallicity and C:O ratio under the assumption of equilibrium chemistry for the *Hubble*/WFC3 spectrum of WASP-12b. The results for each case are in agreement in terms of the retrieved temperature and H_2O abundance, where H_2O is the only gas that can be constrained. Based on the assumptions within the chemical equilibrium model, Kreidberg et al. (2015) reject a carbon-rich atmosphere scenario at $> 3\sigma$ confidence, as the retrieved H_2O abundance is higher than predicted for a carbon-rich model. However, this result is dependent on the assumptions within the chemical model used, so is somewhat less agnostic than a free-chemistry retrieval would be; there is a trade off between obtaining a tighter constraint and relying on a potentially flawed chemical model.

The only way to reliably demonstrate recoverability of underlying chemical trends is to conduct blind tests of retrieval algorithms on synthetic observations with known chemistry. There are two distinct facets to this challenge; 1) can the correct atmospheric C:O ratio be recovered for the constituents present within the observable atmosphere of the planet? and 2) can the correct planet bulk C:O ratio be recovered from the atmospheric C:O? The first issue simply relies on the ability of a retrieval algorithm to accurately determine the abundances of molecular and atomic species within a planet's atmosphere, whereas the second encompasses scenarios in which the bulk planet chemistry is not reflected in the molecular make up of the atmosphere, for example because substantial amounts of some elements are present in the form of clouds deep in the atmosphere. An illustration of this difficulty is the challenge of determining the H_2O volume mixing ratio in Jupiter's atmosphere; see e.g. Li et al. (2020). Simple tests can be performed to answer question 1) with 1D forward models containing some parameterised chemistry, but for question 2) more complex models following through from planet formation to the eventual atmospheric composition will be required.

In the short term, studies testing the ability to accurately recover chemical trends in atmospheric composition should be undertaken. Efforts in this direction are already underway in preparation for the *ARIEL* mission, but similar studies are required for other datasets as the information content of spectra is highly dependent on the precise details of resolving power and wavelength coverage.

Recommended action: conduct retrievals of simulated datasets with known atmospheric chemistry, for a range of planetary temperatures and metallicities, as observed by a variety of instruments. This will allow us to determine observational requirements for precise constraints on C:O ratio, and other trends of interest e.g. N:O ratio.

3 Temperature structure

Whilst detailed information about temperature structure is difficult to obtain from primary transit observations due to the relatively narrow pressure range that is probed, temperature-pressure profiles have been retrieved from secondary eclipse and phase curve spectra. Whilst very broad spectral coverage, such as that available for HD 189733b (e.g. Lee et al. 2012; Line et al. 2014), probes a sufficient range of atmospheric pressures to allow a smoothed, free retrieval of temperature as a function of pressure, the majority of secondary eclipse spectra cover a smaller range and parameterisation is necessary to extrapolate the atmospheric structure beyond the region that is directly constrained.

3.1 State of the art: temperature-pressure profiles

The simplest approach to retrieving temperature is to make the crude assumption that the temperature profile is isothermal. This has often been the approach taken when analysing primary transit spectra; however, Rocchetto et al. (2016) show in their synthetic retrieval study for the *James Webb Space Telescope* that this assumption can result in errors of more than an order of magnitude in the retrieved gas abundances for some cases. The isothermal approximation is therefore clearly inadequate, and approaches that capture the broad shape of the temperature structure must be explored.

There are two parameterisation approaches favoured by retrieval groups, the simpler of the two being the Guillot profile (Guillot 2010) which has 5 free parameters and was first implemented by Line et al. (2012), and the other being the approach advocated by Madhusudhan and Seager (2009), which we will call the Madhusudhan profile, and has 6 free parameters. The original Guillot profile assumes that no scattering occurs in the atmosphere; Heng et al. (2012) and Heng et al. (2014) respectively generalised the Guillot profile to include isotropic scattering (by either aerosols or atoms and molecules), and non-isotropic scattering (large particles).

The Guillot profile is based on a three-channel approximation for an atmosphere in thermal equilibrium and is described by the following equation,

$$T^4(\tau) = \frac{3T_{\text{int}}^4}{4}\left(\frac{2}{3} + \tau\right) + \frac{3T_{\text{irr}}^4}{4}(1 - \alpha)\xi_{\gamma_1}(\tau) + \frac{3T_{\text{irr}}^4}{4}(\alpha)\xi_{\gamma_2}(\tau) \quad (5)$$

where

$$\xi_{\gamma_i} = \frac{2}{3} + \frac{2}{3\gamma_i}\left[1 + \left(\frac{\gamma_i\tau}{2} - 1\right)e^{-\gamma_i\tau}\right] + \frac{2\gamma_i}{3}\left(1 - \frac{\tau^2}{2}\right)\text{E}_2(\gamma_i\tau) \quad (6)$$

and the irradiation temperature is

$$T_{\text{irr}} = \beta\left(\frac{R_\star}{2a}\right)^{1/2}T_\star \quad (7)$$

The 5 free parameters are κ_{IR}, the infrared opacity; $\gamma_1 = \kappa_{\text{v1}}/\kappa_{\text{IR}}$ and $\gamma_2 = \kappa_{\text{v2}}/\kappa_{\text{IR}}$, the ratio of two-band visible opacities to the IR opacity; α, the ratio of the flux between the two visible streams; and β is a measure of the recirculation efficiency of the atmosphere. T_{int} is the planet's internal temperature, and T_{irr} is the temperature calculated from irradiation by the parent star. $R_\star$ and $T_\star$ are the radius and temperature of the parent star, and a is the orbital semi-major axis. $\tau = \kappa_{\text{IR}}p/g$ is the infrared optical depth of the atmosphere, where p is atmospheric pressure and g is gravitational acceleration. E_2 is the second order exponential integral function.

The Madhusudhan profile divides the atmosphere into three layers. Layer 1 is the uppermost and is bounded at the base by pressure P_1. Layer 2 extends from pressure P_1 to P_3, and Layer 3 extends downwards from P_3. The temperature in each layer is defined as follows:

$P_0 < P < P_1 \quad P = P_0 e^{\alpha_1(T-T_0)^{\beta_1}}$

$P_1 < P < P_3 \quad P = P_2 e^{\alpha_2(T-T_2)^{\beta_2}}$

$P > P_3 \quad T = T_3$

In all cases, $P_0 < P_1 < P_3$. If the temperature profile is inverted, $P_1 < P_2 < P_3$; if not, $P_1 \geq P_2$. This can be simplified to only 6 free parameters by setting P_0 equal to the pressure at the top of the atmosphere; empirically setting $\beta_1 = \beta_2 = 0.5$. Finally, the temperature profile is forced to be continuous at the boundaries between the layers where $P = P_1$ and $P = P_3$, which leaves 6 free parameters: P_1, P_2, P_3, α_1, α_2 and T_3.

The advantage of the Guillot profile is that the shape is physically motivated by the assumption of radiative equilibrium, whilst still being a fairly simple parameterisation. It does however contain a bias in that it produces isothermal profiles at low pressures, which may not be an accurate reflection of a real atmosphere. This isothermal behavior is a subtle artefact of using mean opacities, where "mean" in this case is ill-defined. Specifically, in order for the solution to be analytically tractable, the derivation assumes that the absorption, flux and Planck mean opacities are equal. The Madhusudhan profile allows more flexibility of shape, particularly with regards to resolving temperature inversions, at the expense of an additional free parameter. Blecic

et al. (2017) investigate the ability of such 1D temperature parameterisations to recover the temperature structure from synthetic eclipse spectra generated from 3D atmospheric circulation models. They find that the Madhusudhan profile provides a better match to the temperature structure in the middle atmosphere as it is more capable of producing an inversion; however, it does not match the deep temperature structure. We discuss the reliability of fitting a 1D temperature model to a dataset generated from a 3D circulation model in Section 5.1.

A key science question relating to T-p profile retrievals is the presence or absence of a temperature inversion in hot Jupiter atmospheres. Inversions were predicted to occur in planets with incident flux of greater than 10^9 erg s^{-1} cm^{-2}, due to the presence of optical absorbers TiO and VO in their atmospheres (Fortney et al. 2008). This category includes several well-studied hot Jupiters such as HD 209458b, but so far only a handful of planets show evidence for thermal inversions in their dayside spectra. These include WASP-33b (Haynes et al. 2015; fit using Madhusudhan profile); WASP-121b (Evans et al. 2017; fit using Guillot profile); and WASP-18b (Sheppard et al. 2017; fit using Madhusudhan profile). All of these are ultra-hot Jupiters with equilibrium temperatures of over 2000 K, suggesting that the cut-off irradiation for thermal inversions is somewhat higher than originally predicted.

Reliably retrieving the dayside temperature structure is further complicated by the presence of solution degeneracy with gas abundance retrievals. Stevenson et al. (2014a) retrieve the dayside atmospheric state for WASP-12b and test two models which force either carbon-rich or oxygen-rich chemistry; the retrieved temperature profiles differ by several hundred K at low pressures. Similarly, Barstow et al. (2014) test the effect of varying gas abundance priors on a continuous Optimal Estimation retrieval of temperature from HD 189733b emission spectra, and find that the precise shape of the profile is dependent on the gas abundance prior chosen.

3.2 Future challenges for temperature parameterisation

Investigations are underway into the most appropriate temperature parameterisations in the *JWST* era and beyond. Rocchetto et al. (2016) simulate several *JWST* hot Jupiter transmission spectra for model atmospheres with varying C:O ratios, and they demonstrate that oversimplified parameterisations in temperature structure retrieval can introduce significant bias in other retrieved properties. Assuming that the temperature profile is isothermal can result in, for example, retrieved CO abundances over an order of magnitude too high. A Guillot temperature-pressure profile, whilst it increases the uncertainty on the retrieved properties, results in a more accurate retrieval of the gas abundances. However, it is important to note that the input temperature-pressure profile is close to the typical shape predicted by the Guillot parameterisation, so the ability of the Guillot profile to achieve a good fit may be serendipitous.

It is clear, therefore, that accurate chemistry retrievals are dependent on the suitability of the temperature parameterisation.

This issue is likely to only become more complex as the information content of the spectrum increases. The key difficulty in transmission will still be the relatively small pressure range (when compared with eclipse spectra) probed by the observation, and the degeneracy between the effects on temperature, mean molecular weight and gravity on the scale height. Further investigations of the kind presented by Rocchetto et al. (2016) are likely to be a critical aspect of model development. Ultimately, the ideal for eclipse spectra would be to explicitly retrieve temperature at each level in the model atmosphere, subject to some correlation length to ensure smoothness, but this is likely to only be possible for the very highest signal-to-noise observations.

Recommendation: conduct retrievals of simulated datasets with a variety of temperature structures and chemistry, to investigate regions of parameter space where the temperature profile parameterisation introduces most bias. Investigate alternative approaches to those currently in the literature for intractable cases.

4 Clouds

Initial attempts to characterise the atmospheres of hot exoplanets via retrieval were conducted without reference to clouds, due to the erroneous belief that the extreme temperatures would make it impossible for clouds to exist. The inclusion of clouds also inevitably complicates the retrieval process, as it introduces further parameters into what is already an underconstrained retrieval problem. Clouds are complex, potentially spatially variable, structures that provide broadband absorption and scattering, and as such affect spectra in ways that can be difficult to identify. They can also have the effect of muting molecular absorption features.

4.1 State of the art: clouds

So far, retrieval efforts have used simple parameterisations to try and capture the cloud properties that produce the most significant effects on spectra. The different geometries of exoplanet observations require different treatment; in primary transit, due to the long path length through the atmosphere what is often referred to as the cloud top pressure is especially important because the atmospheric opacity rapidly increases below the cloud top. Conversely, the cloud top pressure is less critical if the planet is being directly imaged in the infrared, as the measured radiation is emerging from the planet beneath the cloud top.

Cloud top pressure is not in reality a well-defined pressure above which cloud ceases to exist, although it can be treated as such in simple parameterisations. It represents the pressure level at which the cloud optical depth
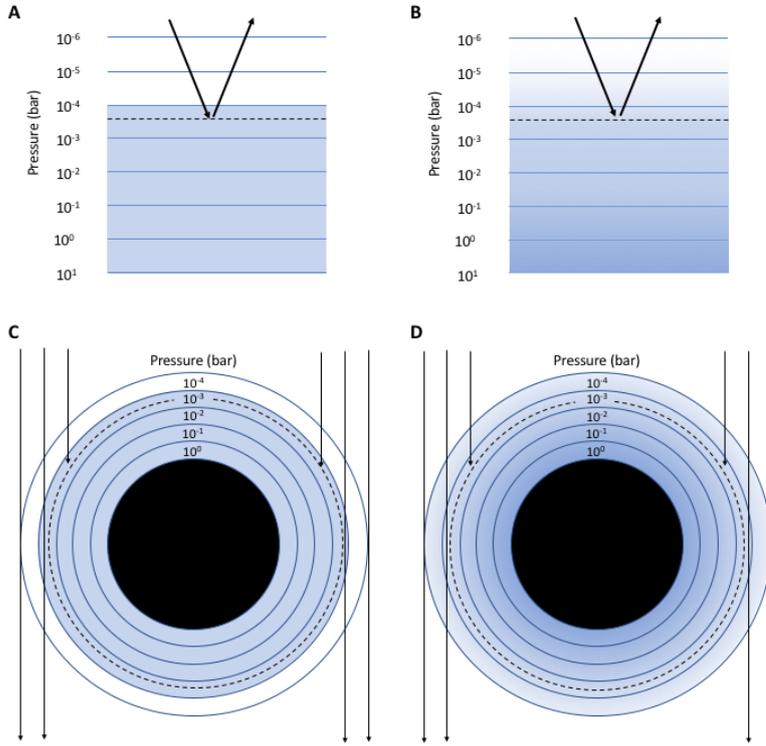
Fig. 4 Effective cloud top pressure for different models in both reflection (nadir; panels A and B) and transit (limb; panels C and D) geometry. In A and C, the cloud has a uniform specific density below a cloud top pressure of 10^{-4} bar; in B and D, the cloud has a specific density that decreases with decreasing pressure, with the cloud extended throughout the atmosphere. The level in the cloud at which the optical depth is unity (dashed line) is the same for both cloud models in each geometry, even though the vertical distribution of aerosol is very different.

is unity, which is highly dependent on the observation geometry - the cloud optical depth reaches unity at a higher altitude in limb geometry compared with nadir. The effective cloud top pressure can be altered in a simple model by setting a physical cloud top, or by varying the opacity of a cloud that is not confined to any particular pressure range. These two approaches are not exactly equivalent, so two different models with different predicted spectra could have the same effective cloud top pressure (Figure 4).

The simplest primary transit studies have assumed that the atmosphere is completely opaque at all wavelengths, for pressures higher than the cloud top pressure. This is suitable over relatively small wavelength ranges, and for planets with spectra that are flat over a wide wavelength range (e.g. GJ 1214b, Kreidberg et al. 2014; Fisher and Heng 2018). However, in general this would only be representative of a cloud made of large particles with a broad

size distribution, and fails to account for scenarios where aerosols may more closely resemble small-particle haze. Slightly more complex parameterisations allow for the possibility of optically thin clouds, and a simple power law for extinction as a function of wavelength (e.g. Barstow et al. 2017; Pinhas et al. 2019.)

Whilst the pressure at the cloud top and the extinction slope are the most important parameters for primary transit, the vertical distribution of the cloud below the cloud top may also be important, depending on the cloud optical thickness. It is also possible that the cloud consists of multiple components - for example, an optically thin, small particle haze layer overlying an optically thick cloud (MacDonald and Madhusudhan 2017; Pinhas et al. 2019). This has led to a range of different parameterisation options even just within primary transit retrievals, which can produce different and apparently contradictory results when applied to the same dataset. For example, retrievals of the same HD 189733b dataset by Barstow et al. (2017) and Pinhas et al. (2019) give consistent values for the H_2O abundance of $\sim 10^{-5}$, but the retrieved cloud properties appear dramatically different at first glance. Barstow et al. (2017) characterize the HD 189733b cloud layer as a vertically thin, high Rayleigh scattering haze layer, whilst Pinhas et al. (2019) retrieve a cloud top deep in the atmosphere. However, this retrieved cloud top is the top of an opaque, grey cloud, which is coupled to a scattering haze layer for $P < P_{\text{top}}$. Therefore, results from both parameterisations are in agreement that there is no visible grey cloud layer, and are consistent with the presence of scattering, small particle haze higher in the atmosphere.

More complex parameterisations that include some information about composition have also been tested. Kitzmann and Heng (2018) develop a parameterisation based on analytical fits to expected extinction cross-section curves of potential cloud species, such as e.g. $MgSiO_3$. The extinction efficiency κ as a function of wavelength is parameterised as follows:

$$\kappa_{\text{cloud}} = \frac{\kappa_0}{Q_0 x^{-a} + x^{0.2}} \tag{8}$$

where κ_0 is a scaling factor, Q_0 determines the wavelength at which the extinction efficiency peaks and is related to the cloud composition, a is a scattering slope index and x is the particle size parameter, given by

$$x = \frac{2\pi r}{\lambda} \tag{9}$$

This parameterisation is more easily related to real physical characteristics of cloud, such as particle size and composition. So far, it has been applied to *Hubble*/WFC3 data by Fisher and Heng (2018), which provides relatively little constraint on cloud properties; it has not yet been applied to data spanning a broader wavelength range.

The limited information available from current spectra, and the range of possible ways cloud can be represented, makes interpretation of these retrievals

very difficult. Without a good understanding of the precise effects of different parameterisations on the spectrum, erroneous conclusions can be drawn.

Attempts have also been made to consider cloud for secondary transit and directly imaged spectra. Barstow et al. (2014) consider the effect of clouds on the HD 189733b reflection spectrum observed by Evans et al. (2013), but due to the requirement to include multiple scattering for reflection spectra only a simple grid search was performed. The cloud properties showed substantial degeneracy with the sodium abundance in the visible part of the spectrum.

4.2 Future challenges for clouds

Current exoplanet retrieval efforts are already demonstrating that the details of parameterisation for cloud properties have the potential to bias results. In the case of cloud properties, gas abundance retrievals seem to be somewhat immune to the differences in cloud treatment, but the conclusions drawn about the clouds themselves can vary widely, as discussed in Section 4.1. The main challenge we face here is to tune complexity of parameterisation to the information content of the data, whilst avoiding where possible introducing bias into the retrieval. Again, the only way to guard against this is to conduct rigorous simulation tests of retrieval parameterisations.

Recent work has been undertaken to combine cloud microphysics models with 3D circulation models, and to use this to predict emergent spectra (Lines et al. 2018). Whilst we do no expect these simulations to perfectly predict real cloud and haze in exoplanet atmospheres, the ability of the retrieval scheme to recover key parameters from these synthetic spectra is an important test of the cloud parameterisation used. It provides an opportunity to check whether the parameterisation is sufficient to represent the spectral effect of complex cloud structure, and ensure that it does not introduce bias into the retrieval. Several different approaches to modelling cloud microphysics (e.g. Helling et al. 2008; Ackerman and Marley 2001) and including cloud in GCMs (e.g. Lee et al. 2017; Parmentier et al. 2016; Mendonça et al. 2018) are available; the ideal would be a parameterised model that can recover key cloud properties from this range of available cloud models, whilst also accurately retrieving other atmospheric properties.

Recommended action: conduct retrievals of simulated datasets based on more detailed, physically motivated, 3D cloudy atmosphere models. Test a variety of simple cloud parameterisations, for a range of observational geometries, and compare results.

5 Phase curves and 3D effects

For a handful of the most favourable targets, spectroscopic phase curves have been obtained which have allowed phase-resolved retrievals to be undertaken. The first example of this is the Stevenson et al. (2014b) phase curve retrieval

for WASP-43b, obtained using *Hubble*/WFC3. The limited wavelength coverage means there is only sensitivity to temperature structure over a small pressure range, and some information about the H_2O abundance. Difficulties of interpretation are compounded because the pressure of weighting function peak varies with phase, so comparison between phases is not straightforward. Phase curve observations with broader spectral coverage would resolve these difficulties and are planned for *JWST*. Mendonça et al. (2018) re-analyzed the Spitzer data of WASP-43b and ran cloudy GCMs to jointly analyze the Hubble and Spitzer phase-resolved emission spectra. They find that the dayside is consistent with being cloudfree, with clouds confined to the nightside, and tentative evidence for elevated levels of carbon dioxide.

Whilst phase curves provide some direct information about spatial variations in the thermal emission from the planet (and in some cases the reflected light), spatial variation in the atmospheric properties can also affect transmission spectra, albeit in a more subtle way. Evidence from observed phase curves and GCMs suggests that one terminator is likely to be hotter than the other for hot Jupiters, which will in turn impact the terminator chemistry and cloud coverage. For example, Mendonça et al. (2018) ran GCMs with disequilibrium chemistry (using a method known as "chemical relaxation") and demonstrated that the coupling between atmospheric dynamics and chemistry produces spatial inhomogeneities across latitude, longitude and pressure for molecules such as water, and cannot be neglected if one wishes to accurately model phase-resolved spectra or wavelength-dependent phase curves. The challenge in interpretation is that transmission spectra are averaged over the whole terminator region, so observations are implicitly 1D. Similarly, for planets not sufficiently favourable for us to have phase curve observations, secondary eclipse spectra are also 1D integrations over a non-uniform (and asymmetric) disc; the structure and chemistry retrieved using a 1D model will represent some sort of disc average, but it is unclear exactly what this corresponds to (Figure 5).

Likewise, the transit spectroscopy technique relies on the stellar disc being uniform once stellar limb darkening is corrected for, since it makes the implicit assumption that the planet transits a region of the stellar disc that is representative of the whole. This is of course not the case; stellar surfaces are highly non-uniform, with time-variable coverage of features such as spots and faculae. Spots and faculae have different spectral characteristics compared with the rest of the stellar disc, so unknown spot/faculae coverage fractions could lead to misinterpretation of transit spectra (e.g. Rackham et al. 2018).

Directly imaged planets are relatively free of these issues, since they are not highly irradiated and their observations do not depend on the uniformity of the star's behaviour; we expect them to more closely resemble the Solar System giant planets in terms of their dynamics. However, we cannot rule out spatial asymmetry on these objects; whilst the dynamical regimes of hot Jupiters result in strong longitudinal gradients in temperature, the Solar System giants display latitudinal variation in chemistry and cloud properties (see e.g. PH_3 abundance on Jupiter and Saturn, Fletcher et al. 2009; in addition to the
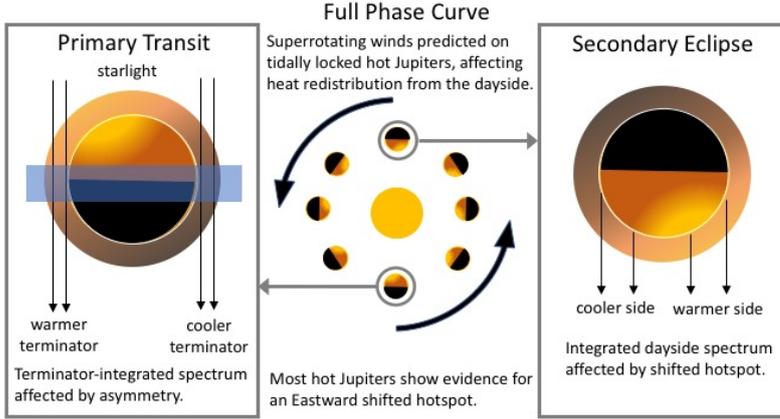
Fig. 5 Strong superrotation on hot Jupiters, coupled with extreme irradiation, results in significant variation in temperature around the terminator region as observed in transit, and an asymmetric pattern on variation on the dayside as observed in eclipse.

equator-pole differences observed on Jupiter, Saturn also has strong north-south seasonal asymmetry due to its axial tilt of 26.7°).

5.1 State of the art: 3D effects

5.1.1 Phase curve retrievals

For WASP-43b, a moderately hot Jupiter, it has been possible to obtain a spectroscopic phase curve using the *Hubble*/WFC3 instrument. This allows retrievals to be performed as a function of phase, allowing longitudinal variations in chemistry and temperature structure to be mapped. Stevenson et al. (2014b) use the CHIMERA retrieval algorithm to analyse temperature structure at 16 different phases. The model includes 6 molecular absorbers, but only H_2O has a significant influence on the spectral characteristics. The temperature structure is modelled using 5 free parameters, after the method presented by Parmentier and Guillot (2014). The retrieved upper atmosphere temperatures vary by 1000 K between the dayside and nightside, implying inefficient recirculation.

So far, this is the only planet for which a full spectroscopic phase curve exists, so further exploration of phase curve retrievals is hindered by a lack of available data. Retrieval algorithms have not been applied to single- and multi-channel photometric phase curves that exist for other planets, presumably because the problem would be highly degenerate. However, spectroscopic phase curve observations are likely to be a priority for *JWST*. WASP-43b is particularly well-suited to such observations as it has a very short period of only 19.52 hours; this planet will be re-observed at longer wavelengths with

the Mid-InfraRed Instrument (MIRI) during the *JWST* Early Release Science programme (Batalha et al. 2017), which will provide stronger constraints on the variation in atmospheric properties with phase. A phase curve for WASP-43b will also be obtained with the shorter wavelength NIRSpec instrument as part of the Guaranteed Time Observation for the instrument team (Birkmann et al. 2017).

5.1.2 3D cloud effects in transmission

Work is already in progress to account for terminator asymmetry in retrieval models (e.g. Line and Parmentier 2016; MacDonald and Madhusudhan 2017), although so far it is restricted to cloud coverage, which ignores the fact that temperature structure, and likely the chemistry too, will also vary. Line and Parmentier (2016) demonstrate that, over narrow wavelength ranges such as those probed by *Hubble*/WFC3 only, partial terminator cloud cover is degenerate with cloud-free, high mean molecular weight atmosphere scenarios. Over a wider wavelength range, this degeneracy can be broken. Pinhas et al. (2019) include terminator cloud fraction in their retrieval of *Hubble*/STIS + WFC + *Spitzer*/IRAC spectra, and they recover a range of values between ∼0.2 and ∼0.8 for the 10 planets in their sample. They find no correlation between cloud fraction and any other key parameters in the study. Line and Parmentier (2016) analyse WFC3 data only for HD 189733b, and find a cloud fraction that is comparable with the result from Pinhas et al. (2019).

5.1.3 3D temperature structure from eclipse spectra

Blecic et al. (2017) investigate the ability of a 1D parameterised model to recover an average temperature structure from a simulated dayside spectrum generated from a 3D model atmosphere. They test both the Guillot and Madhusudhan temperature parameterisations discussed previously in Section 3.1. Both parameterisations produce a retrieved temperature profile close to the arithmetic mean of the circulation model temperature profiles across the dayside, which is somewhat odd; the amount of radiation detected from different regions of the dayside is weighted by the cosines of the latitude and longitude, so it should follow that the hemisphere-integrated temperature-pressure profile should be a weighted average rather than a straightforward arithmetic mean (Figure 6). This is an indication that further development of 1D retrieval models, and an investigation into surprising results such as this one, are required to reliably interpret these hemispherically averaged spectra.

Feng et al. (2016) test the impact of using two temperature-pressure profiles to represent the hotter/colder regions of a planetary disc. They apply this to the first-quarter observation of WASP-43b, for which the visible portion of the planet is half in daylight and half in shadow, maximising the expected contrast. They also test simulated spectra for both current state-of-the-art observational scenarios (*Hubble*+*Spitzer*) and future observations with *JWST*. They find that there is insufficient evidence with current data to favour a more complex
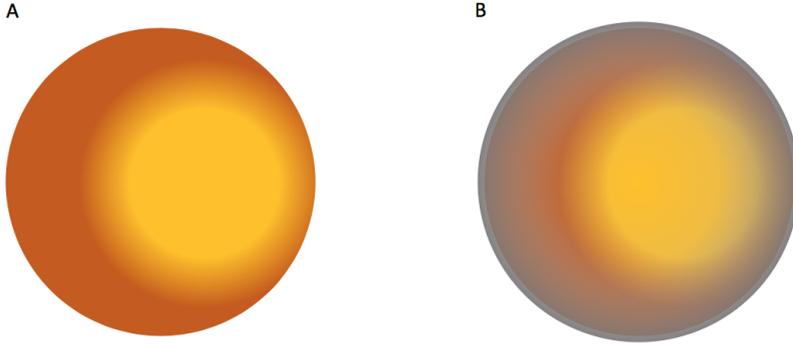
Fig. 6 Panel A illustrates the contributions from each part of the dayside disc where the emission angle is not taken into account, whereas panel B shows which parts of the planet would dominate the signal when emission angle is accounted for.

model, but that for *JWST* simulations significant biases in gas abundances are introduced when only a single temperature-pressure profile is used to represent the temperature structure. This approach is shown to work well where the temperature variation is adequately represented by two temperature-pressure profiles of equal weight, but it remains to be seen whether this is appropriate in the context of secondary eclipse, where the hotspot is likely to dominate.

5.1.4 Stellar heterogeneity in retrievals

Parameterisation of the effects of starspots and faculae is now starting to be included within exoplanet retrieval frameworks. Initial results for super Earth GJ 1214b are presented by Rackham et al. (2017). *Magellan* telescope observations are fit using the CPAT absorber model coupled with a Markov-Chain Monte Carlo algorithm. The CPAT model for describing stellar heterogeneity divides the stellar disc into occulted and unocculted fractions. The wavelength-dependent transit depth, instead of being simply given by

$$\Delta_\lambda = 1 - \left(\frac{R_{\mathrm{p},\lambda}}{R_\mathrm{s}} \right)^2 \qquad (10)$$

where $R_{\mathrm{p},\lambda}$ is the radius of the planet and R_s the radius of the star, is instead given by

$$\Delta_\lambda = 1 - \frac{(R_{\mathrm{p},\lambda}/R_\mathrm{s})^2 S_\mathrm{o}}{(1-F)S_\mathrm{o} + F S_\mathrm{u}} \qquad (11)$$

where S_o is the spectrum of the star in the occulted region, S_u is the spectrum of the star in the unocculted region, and F is the fraction of the disc that is unocculted. Rackham et al. (2017) test PHOENIX (Husser et al. 2013) model spectra with different metallicities, and different temperatures as a proxy for

varying levels of absorption across the stellar disc. They retrieve metallicity/temperature contrast between the occulted and unocculted regions and a constant offset in $R_{\rm p}/R_{\rm s}$, finding that the observed optical spectrum can be described by a case where 3.2 % of the unocculted disc is 350 K hotter than the rest of the disc. This may be explained by starspot or facula contrast.

Pinhas et al. (2018) perform a retrieval analysis of nine hot Jupiters (the sample from Sing et al. (2016) excluding HD 189733b) using the same model as that presented in Pinhas et al. (2019) but also including stellar hetereogeneity. This is parameterised by the temperature of the heterogeneous regions (with the star's measured average photospheric temperature fixed) and the fractional coverage of any heterogeneities. Pinhas et al. (2018) do not discuss these values in detail, but instead present the model evidence for inclusion of stellar effects. They find substantial evidence of stellar heterogeneity for WASP-6 and WASP-39; whilst WASP-6 is one of the two most active stars in the sample based on $\log R_{\rm HK}$ index, WASP-39 is less active, and for the most active star (WASP-19) the evidence is substantially against there being any stellar heterogeneity. This would indicate that the $\log R_{\rm HK}$ index is an unreliable estimator of the importance of stellar heterogeneity effects on transit spectra.

5.2 Future challenges for recovering planetary and stellar spatial information

Two key resources for exploration of our ability to recover 3D information about planets are Global Circulation Models (GCMs; e.g. Selsis et al. 2011; Rauscher and Menou 2012; Charnay et al. 2015; Amundsen et al. 2016; Lee et al. 2016; Parmentier et al. 2016; Mendonça et al. 2018), and the Solar System planets. GCMs are based on our current best understanding of the physical processes on hot Jupiters and young directly-imaged planets, and should be able to predict the broad characteristics of spatial variability on these planets. However, there are limits to the predictive power of GCMs due to the inability to accurately specify and represent all sources of dissipation in the atmosphere, e.g, Goodman (2009); Heng et al. (2011); Fromang et al. (2016); on Earth, these uncertainties can be mitigated by empirically calibrating the sources of dissipation in the GCM using in-situ data, an approach that is impossible for exoplanets.

The Solar System giant planets on the other hand, whilst they exist in a very different temperature/dynamics regime to the majority of well-studied exoplanets, have the advantage that we can directly compare spatially resolved datasets with the information that we would be able to recover if the planet was treated as a point source. This allows us to investigate for real objects how much information about large scale atmospheric ability and asymmetry persists in disc-integrated observations.

Models will also be key for understanding the impact of stellar heterogeneities on transmission spectra. As shown by Rackham et al. (2017), whilst monitoring of target stars can provide an indication of the amplitude of variation in spot coverage, this does not provide information about the baseline

level or the relative contributions of spots and faculae, both of which are important for transmission spectra. Understanding typical distributions and sizes of spots/faculae on different types of star will be extremely important for future observations.

Recommended actions: use simulated datasets from GCMs/stellar atmosphere models to test the ability of parameterised retrieval models to recover 3D information about the planet and the star. Investigate how the information content of spatially resolved observations of Solar System giants compares with that of the same observation degraded to a point source.

6 Conclusions

We have presented a summary of the current and imminent future challenges surrounding atmospheric retrievals of exoplanets. In general, the obstacles faced result from the lack of available ground truth for exoplanet observations, and, especially in the near future, a rapid increase in the information content of observations which requires modelling strategies to constantly evolve.

A common theme for solutions to these challenges is the use of physically based climate and circulation models to provide simulated datasets. Whilst we cannot yet be sure that the outputs from these models are accurate representations of real exoplanet atmospheres, they do allow us to perform important tests of how well simple parameterised models capture more complex atmospheric characteristics. In the case of stellar heterogeneity, it is likely that we will have to rely to some extent on ab initio stellar atmosphere models if we want to correct for spectral contamination of starspots and faculae.

Another key attribute required for retrieval models is flexibility; since the data quality is, and is likely to remain, variable across different planets, it is important that models can be easily tuned to maximally exploit the information content of a given observation. Oversimplification has been demonstrated to introduce bias - for example, assuming an isothermal temperature structure for broad wavelength coverage observations can significantly bias the retrieved chemistry - but equally overfitting can also produce problems. Explicit calculation of information content, such as that featured by Howe et al. (2017), may prove useful both for observation planning and also for tailoring retrieval models.

There are of course several aspects of exoplanet spectral inversion that we have not touched on. Perhaps one of the most significant is the completeness and accuracy of the gas absorption information that is included in retrieval schemes. Tennyson and Yurchenko (2018) provide a summary of the ExoMol project, which is one of the current community efforts to ensure that gas absorption data are as accurate as possible. In addition, the processing of these data for inclusion in retrieval models is also an important step that can be a potential source of error.

Finally, there are other methods for extracting spectral information of exoplanet atmospheres which we have not discussed here, as they are beyond the scope of this paper. These include high-spectral-resolution observations, which can also be used to recover information about exoplanet chemistry and atmospheres (e.g. de Mooij and Snellen 2009; Schwarz et al. 2015; Hoeijmakers et al. 2018); their use in retrieval scenarios is currently being explored (Brogi and Line 2019). We have also focused on transiting exoplanets in this work; with the launch of *JWST*, and first-light for next generation ground-based telescopes such as the *Extremely Large Telescope* fast approaching, significant advances in direct spectral imaging of exoplanets may also be expected over the current state-of-the-art (represented by e.g. Macintosh et al. (2015); Bonnefoy et al. (2016); Gravity Collaboration et al. (2019)), opening up further opportunities to characterise non-transiting worlds.

References

A.S. Ackerman, M.S. Marley, Precipitating Condensation Clouds in Substellar Atmospheres. Astrophys. J. **556**, 872–884 (2001). doi:10.1086/321540

D.S. Amundsen, N.J. Mayne, I. Baraffe, J. Manners, P. Tremblin, B. Drummond, C. Smith, D.M. Acreman, D. Homeier, The UK Met Office global circulation model with a sophisticated radiation scheme applied to the hot Jupiter HD 209458b. Astron. Astrophys. **595**, 36 (2016). doi:10.1051/0004-6361/201629183

T. Barman, Identification of absorption features in an extrasolar planet atmosphere. The Astrophysical Journal **661**(2), 191–194 (2007). doi:10.1086/518736

J.K. Barstow, P.G.J. Irwin, Habitable worlds with JWST: transit spectroscopy of the TRAPPIST-1 system? Mon. Not. R. Astron. Soc. **461**, 92–96 (2016). doi:10.1093/mnrasl/slw109

J.K. Barstow, S. Aigrain, P.G.J. Irwin, T. Hackler, L.N. Fletcher, J.M. Lee, N.P. Gibson, Clouds on the Hot Jupiter HD189733b: Constraints from the Reflection Spectrum. Astrophys. J. **786**, 154 (2014). doi:10.1088/0004-637X/786/2/154

J.K. Barstow, S. Aigrain, P.G.J. Irwin, S. Kendrew, L.N. Fletcher, Transit spectroscopy with James Webb Space Telescope: systematics, starspots and stitching. Mon. Not. R. Astron. Soc. **448**(3), 2546–2561 (2015). doi:10.1093/mnras/stv186

J.K. Barstow, S. Aigrain, P.G.J. Irwin, D.K. Sing, A Consistent Retrieval Analysis of 10 Hot Jupiters Observed in Transmission. Astrophys. J. **834**, 50 (2017). doi:10.3847/1538-4357/834/1/50

N. Batalha, J. Bean, K. Stevenson, M. Alam, N. Batalha, B. Benneke, Z. Berta-Thompson, J. Blecic, G. Bruno, A. Carter, J. Chapman, I. Crossfield, N. Crouzet, L. Decin, B. Demory, J. Desert, D. Dragomir, T. Evans, J. Fortney, J. Fraine, P. Gao, A. Garcia Munoz, N. Gibson, J. Goyal, J. Harrington, K. Heng, R. Hu, E. Kempton, S. Kendrew, B. Kilpatrick, H. Knutson, L. Kreidberg, J. Krick, P. Lagage, M. Lendl, M. Line, M. Lopez-Morales, T. Louden, N. Madhusudhan, A. Mandell, M. Mansfield, E. May, G. Morello, C. Morley, J. Moses, N. Nikolov, V. Parmentier, S. Redfield, J. Roberts, E. Schlawin, A. Showman, D. Sing, J. Spake, M. Swain, K. Todorov, A. Tsiaras, O. Venot, W. Waalkes, H. Wakeford, P. Wheatley, R. Zellem, *The Transiting Exoplanet Community*

Early Release Science Program, JWST Proposal ID 1366. Cycle 0 Early Release Science, 2017

B. Benneke, S. Seager, Atmospheric Retrieval for Super-Earths: Uniquely Constraining the Atmospheric Composition with Transmission Spectroscopy. Astrophys. J. **753**(2), 100 (2012). doi:10.1088/0004-637X/753/2/100

S. Birkmann, P. Ferruit, J.A. Valenti, *Transiting exoplanet characterization with JWST/NIRSPEC*, JWST Proposal ID 1224. Cycle 1 GTO, 2017

J. Blecic, I. Dobbs-Dixon, T. Greene, The Implications of 3D Thermal Structure on 1D Atmospheric Retrieval. Astrophys. J. **848**, 127 (2017). doi:10.3847/1538-4357/aa8171

M. Bonnefoy, A. Zurlo, J.L. Baudino, P. Lucas, D. Mesa, A.-L. Maire, A. Vigan, R. Galicher, D. Homeier, F. Marocco, First light of the VLT planet finder SPHERE. IV. Physical and chemical properties of the planets around HR8799. Astron. Astrophys. **587**, 58 (2016). doi:10.1051/0004-6361/201526906

M. Brogi, M.R. Line, Retrieving Temperatures and Abundances of Exoplanet Atmospheres with High-resolution Cross-correlation Spectroscopy. Astron. J. **157**(3), 114 (2019). doi:10.3847/1538-3881/aaffd3

B. Charnay, V. Meadows, J. Leconte, 3D Modeling of GJ1214b's Atmosphere: Vertical Mixing Driven by an Anti-Hadley Circulation. Astrophys. J. **813**(1), 15 (2015). doi:10.1088/0004-637X/813/1/15

E.J.W. de Mooij, I.A.G. Snellen, Ground-based K-band detection of thermal emission from the exoplanet TrES-3b. Astron. Astrophys. **493**(2), 35–38 (2009). doi:10.1051/0004-6361:200811239

T.M. Evans, F. Pont, D.K. Sing, S. Aigrain, J.K. Barstow, J.-M. Désert, N. Gibson, K. Heng, H.A. Knutson, A. Lecavelier des Etangs, The Deep Blue Color of HD 189733b: Albedo Measurements with Hubble Space Telescope/Space Telescope Imaging Spectrograph at Visible Wavelengths. Astrophys. J. Lett. **772**, 16 (2013). doi:10.1088/2041-8205/772/2/L16

T.M. Evans, D.K. Sing, T. Kataria, J. Goyal, N. Nikolov, H.R. Wakeford, D. Deming, M.S. Marley, D.S. Amundsen, G.E. Ballester, J.K. Barstow, L. Ben-Jaffel, V. Bourrier, L.A. Buchhave, O. Cohen, D. Ehrenreich, A. García Muñoz, G.W. Henry, H. Knutson, P. Lavvas, A. Lecavelier Des Etangs, N.K. Lewis, M. López-Morales, A.M. Mandell, J. Sanz-Forcada, P. Tremblin, R. Lupu, An ultrahot gas-giant exoplanet with a stratosphere. Nature **548**, 58–61 (2017). doi:10.1038/nature23266

T.M. Evans, D.K. Sing, J.M. Goyal, N. Nikolov, M.S. Marley, K. Zahnle, G.W. Henry, J.K. Barstow, M.K. Alam, J. Sanz-Forcada, T. Kataria, N.K. Lewis, P. Lavvas, G.E. Ballester, L. Ben-Jaffel, S.D. Blumenthal, V. Bourrier, B. Drummond, A. García Muñoz, M. López-Morales, P. Tremblin, D. Ehrenreich, H.R. Wakeford, L.A. Buchhave, A. Lecavelier des Etangs, É. Hébrard, M.H. Williamson, An Optical Transmission Spectrum for the Ultra-hot Jupiter WASP-121b Measured with the Hubble Space Telescope. Astron. J. **156**, 283 (2018). doi:10.3847/1538-3881/aaebff

Y.K. Feng, M.R. Line, J.J. Fortney, K.B. Stevenson, J. Bean, L. Kreidberg, V. Parmentier, The Impact of Non-uniform Thermal Structure on the Interpretation of Exoplanet Emission Spectra. Astrophys. J. **829**(1), 52 (2016). doi:10.3847/0004-637X/829/1/52

F. Feroz, M.P. Hobson, Multimodal nested sampling: an efficient and robust alternative to Markov Chain Monte Carlo methods for astronomical data analyses. Mon. Not. R. Astron. Soc. **384**, 449–463 (2008). doi:10.1111/j.1365-2966.2007.12353.x

F. Feroz, M.P. Hobson, M. Bridges, MULTINEST: an efficient and robust Bayesian inference tool for cosmology and particle physics. Mon. Not. R. Astron. Soc. **398**, 1601–1614 (2009). doi:10.1111/j.1365-2966.2009.14548.x

F. Feroz, M.P. Hobson, E. Cameron, A.N. Pettitt, Importance Nested Sampling and the MultiNest Algorithm. ArXiv e-prints (2013)

C. Fisher, K. Heng, Retrieval analysis of 38 WFC3 transmission spectra and resolution of the normalization degeneracy. Mon. Not. R. Astron. Soc. **481**, 4698–4727 (2018). doi:10.1093/mnras/sty2550

L.N. Fletcher, G.S. Orton, N.A. Teanby, P.G.J. Irwin, Phosphine on Jupiter and Saturn from Cassini/CIRS. Icarus **202**, 543–564 (2009). doi:10.1016/j.icarus.2009.03.023

J.J. Fortney, K. Lodders, M.S. Marley, R.S. Freedman, A Unified Theory for the Atmospheres of the Hot and Very Hot Jupiters: Two Classes of Irradiated Atmospheres.

Astrophys. J. **678**(2), 1419–1435 (2008). doi:10.1086/528370

S. Fromang, J. Leconte, K. Heng, Shear-driven instabilities and shocks in the atmospheres of hot Jupiters. Astron. Astrophys. **591**, 144 (2016). doi:10.1051/0004-6361/201527600

J. Goodman, Thermodynamics of Atmospheric Circulation on Hot Jupiters. Astrophys. J. **693**(2), 1645–1649 (2009). doi:10.1088/0004-637X/693/2/1645

Gravity Collaboration, S. Lacour, M. Nowak, J. Wang, O. Pfuhl, F. Eisenhauer, R. Abuter, A. Amorim, N. Anugu, M. Benisty, First direct detection of an exoplanet by optical interferometry. Astrometry and K-band spectroscopy of HR 8799 e. Astron. Astrophys. **623**, 11 (2019). doi:10.1051/0004-6361/201935253

T.P. Greene, M.R. Line, C. Montero, J.J. Fortney, J. Lustig-Yaeger, K. Luther, Characterizing Transiting Exoplanet Atmospheres with JWST. Astrophys. J. **817**, 17 (2016). doi:10.3847/0004-637X/817/1/17

C.A. Griffith, Disentangling degenerate solutions from primary transit and secondary eclipse spectroscopy of exoplanets. Philosophical Transactions of the Royal Society of London Series A **372**(2014), 20130086–20130086 (2014). doi:10.1098/rsta.2013.0086

T. Guillot, On the radiative equilibrium of irradiated planetary atmospheres. Astron. Astrophys. **520**, 27 (2010). doi:10.1051/0004-6361/200913396

K. Haynes, A.M. Mandell, N. Madhusudhan, D. Deming, H. Knutson, Spectroscopic Evidence for a Temperature Inversion in the Dayside Atmosphere of Hot Jupiter WASP-33b. Astrophys. J. **806**(2), 146 (2015). doi:10.1088/0004-637X/806/2/146

C. Helling, P. Woitke, W.-F. Thi, Dust in brown dwarfs and extra-solar planets. I. Chemical composition and spectral appearance of quasi-static cloud layers. Astron. Astrophys. **485**, 547–560 (2008). doi:10.1051/0004-6361:20078220

K. Heng, D. Kitzmann, The theory of transmission spectra revisited: a semi-analytical method for interpreting WFC3 data and an unresolved challenge. Mon. Not. R. Astron. Soc. **470**(3), 2972–2981 (2017). doi:10.1093/mnras/stx1453

K. Heng, J.R. Lyons, Carbon Dioxide in Exoplanetary Atmospheres: Rarely Dominant Compared to Carbon Monoxide and Water in Hot, Hydrogen-dominated Atmospheres. Astrophys. J. **817**(2), 149 (2016). doi:10.3847/0004-637X/817/2/149

K. Heng, J.M. Mendonça, J.-M. Lee, Analytical Models of Exoplanetary Atmospheres. II. Radiative Transfer via the Two-stream Approximation. Astrophys. J. Supp. **215**(1), 4 (2014). doi:10.1088/0067-0049/215/1/4

K. Heng, K. Menou, P.J. Phillipps, Atmospheric circulation of tidally locked exoplanets: a suite of benchmark tests for dynamical solvers. Mon. Not. R. Astron. Soc. **413**(4), 2380–2402 (2011). doi:10.1111/j.1365-2966.2011.18315.x

K. Heng, W. Hayek, F. Pont, D.K. Sing, On the effects of clouds and hazes in the atmospheres of hot Jupiters: semi-analytical temperature-pressure profiles. Mon. Not. R. Astron. Soc. **420**(1), 20–36 (2012). doi:10.1111/j.1365-2966.2011.19943.x

H.J. Hoeijmakers, D. Ehrenreich, K. Heng, D. Kitzmann, S.L. Grimm, R. Allart, R. Deitrick, A. Wyttenbach, M. Oreshenko, L. Pino, Atomic iron and titanium in the atmosphere of the exoplanet KELT-9b. Nature **560**(7719), 453–455 (2018). doi:10.1038/s41586-018-0401-y

A.R. Howe, A. Burrows, D. Deming, An Information-theoretic Approach to Optimize JWST Observations and Retrievals of Transiting Exoplanet Atmospheres. Astrophys. J. **835**(1), 96 (2017). doi:10.3847/1538-4357/835/1/96

T.-O. Husser, S. Wende-von Berg, S. Dreizler, D. Homeier, A. Reiners, T. Barman, P.H. Hauschildt, A new extensive library of PHOENIX stellar atmospheres and synthetic spectra. Astron. Astrophys. **553**, 6 (2013). doi:10.1051/0004-6361/201219058

P.G.J. Irwin, N.A. Teanby, R. de Kok, L.N. Fletcher, C.J.A. Howett, C.C.C. Tsang, C.F. Wilson, S.B. Calcutt, C.A. Nixon, P.D. Parrish, The NEMESIS planetary atmosphere radiative transfer and retrieval tool. JQSRT **109**, 1136–1150 (2008). doi:10.1016/j.jqsrt.2007.11.006

D. Kitzmann, K. Heng, Optical properties of potential condensates in exoplanetary atmospheres. Mon. Not. R. Astron. Soc. **475**(1), 94–107 (2018). doi:10.1093/mnras/stx3141

L. Kreidberg, J.L. Bean, J.-M. Désert, B. Benneke, D. Deming, K.B. Stevenson, S. Seager, Z. Berta-Thompson, A. Seifahrt, D. Homeier, Clouds in the atmosphere of the super-Earth exoplanet GJ1214b. Nature **505**, 69–72 (2014). doi:10.1038/nature12888

L. Kreidberg, M.R. Line, J.L. Bean, K.B. Stevenson, J.-M. Désert, N. Madhusudhan, J.J.

Fortney, J.K. Barstow, G.W. Henry, M.H. Williamson, A Detection of Water in the Transmission Spectrum of the Hot Jupiter WASP-12b and Implications for Its Atmospheric Composition. Astrophys. J. **814**(1), 66 (2015). doi:10.1088/0004-637X/814/1/66

J. Krissansen-Totton, R. Garland, P. Irwin, D.C. Catling, Detectability of Biosignatures in Anoxic Atmospheres with the James Webb Space Telescope: A TRAPPIST-1e Case Study. Astron. J. **156**, 114 (2018). doi:10.3847/1538-3881/aad564

G.K.H. Lee, K. Wood, I. Dobbs-Dixon, A. Rice, C. Helling, Dynamic mineral clouds on HD 189733b. II. Monte Carlo radiative transfer for 3D cloudy exoplanet atmospheres: combining scattering and emission spectra. Astron. Astrophys. **601**, 22 (2017). doi:10.1051/0004-6361/201629804

G. Lee, I. Dobbs-Dixon, C. Helling, K. Bognar, P. Woitke, Dynamic mineral clouds on HD 189733b. I. 3D RHD with kinetic, non-equilibrium cloud formation. Astron. Astrophys. **594**, 48 (2016). doi:10.1051/0004-6361/201628606

J.-M. Lee, L.N. Fletcher, P.G.J. Irwin, Optimal estimation retrievals of the atmospheric structure and composition of HD 189733b from secondary eclipse spectroscopy. Mon. Not. R. Astron. Soc. **420**, 170–182 (2012). doi:10.1111/j.1365-2966.2011.20013.x

C. Li, A. Ingersoll, S. Bolton, S. Levin, M. Janssen, S. Atreya, J. Lunine, P. Steffes, S. Brown, T. Guillot, M. Allison, J. Arballo, A. Bellotti, V. Adumitroaie, S. Gulkis, A. Hodges, L. Li, S. Misra, G. Orton, F. Oyafuso, D. Santos-Costa, H. Waite, Z. Zhang, The water abundance in Jupiter's equatorial zone. Nature Astronomy (2020). doi:10.1038/s41550-020-1009-3

M.R. Line, V. Parmentier, The Influence of Nonuniform Cloud Cover on Transit Transmission Spectra. Astrophys. J. **820**, 78 (2016). doi:10.3847/0004-637X/820/1/78

M.R. Line, X. Zhang, G. Vasisht, V. Natraj, P. Chen, Y.L. Yung, Information Content of Exoplanetary Transit Spectra: An Initial Look. Astrophys. J. **749**, 93 (2012). doi:10.1088/0004-637X/749/1/93

M.R. Line, A.S. Wolf, X. Zhang, H. Knutson, J.A. Kammer, E. Ellison, P. Deroo, D. Crisp, Y.L. Yung, A Systematic Retrieval Analysis of Secondary Eclipse Spectra. I. A Comparison of Atmospheric Retrieval Techniques. ApJ **775**, 137 (2013). doi:10.1088/0004-637X/775/2/137

M.R. Line, H. Knutson, A.S. Wolf, Y.L. Yung, A Systematic Retrieval Analysis of Secondary Eclipse Spectra. II. A Uniform Analysis of Nine Planets and their C to O Ratios. Astrophys. J. **783**, 70 (2014). doi:10.1088/0004-637X/783/2/70

S. Lines, J. Manners, N.J. Mayne, J. Goyal, A.L. Carter, I.A. Boutle, G.K.H. Lee, C. Helling, B. Drummond, D.M. Acreman, D.K. Sing, Exonephology: transmission spectra from a 3D simulated cloudy atmosphere of HD 209458b. Mon. Not. R. Astron. Soc. **481**, 194–205 (2018). doi:10.1093/mnras/sty2275

R.J. MacDonald, N. Madhusudhan, HD 209458b in new light: evidence of nitrogen chemistry, patchy clouds and sub-solar water. Mon. Not. R. Astron. Soc. **469**(2), 1979–1996 (2017). doi:10.1093/mnras/stx804

B. Macintosh, J.R. Graham, T. Barman, R.J. De Rosa, Q. Konopacky, M.S. Marley, C. Marois, E.L. Nielsen, L. Pueyo, A. Rajan, Discovery and spectroscopy of the young jovian planet 51 Eri b with the Gemini Planet Imager. Science **350**(6256), 64–67 (2015). doi:10.1126/science.aac5891

N. Madhusudhan, S. Seager, A temperature and abundance retrieval method for exoplanet atmospheres. The Astrophysical Journal **707**(1), 24 (2009). http://stacks.iop.org/0004-637X/707/i=1/a=24

P. Márquez-Neila, C. Fisher, R. Sznitman, K. Heng, Supervised machine learning for analysing spectra of exoplanetary atmospheres. Nature Astronomy **2**, 719–724 (2018). doi:10.1038/s41550-018-0504-2

J.M. Mendonça, M. Malik, B.-O. Demory, K. Heng, Revisiting the Phase Curves of WASP-43b: Confronting Re-analyzed Spitzer Data with Cloudy Atmospheres. Astron. J. **155**, 150 (2018). doi:10.3847/1538-3881/aaaebc

K.I. Öberg, R. Murray-Clay, E.A. Bergin, The Effects of Snowlines on C/O in Planetary Atmospheres. Astrophys. J. Lett. **743**, 16 (2011). doi:10.1088/2041-8205/743/1/L16

M. Oreshenko, B. Lavie, S.L. Grimm, S.-M. Tsai, M. Malik, B.-O. Demory, C. Mordasini, Y. Alibert, W. Benz, S.P. Quanz, R. Trotta, K. Heng, Retrieval Analysis of the Emission Spectrum of WASP-12b: Sensitivity of Outcomes to Prior Assumptions and Implications

for Formation History. Astrophys. J. Lett. **847**, 3 (2017). doi:10.3847/2041-8213/aa8acf

V. Parmentier, T. Guillot, A non-grey analytical model for irradiated atmospheres. I. Derivation. Astron. Astrophys. **562**, 133 (2014). doi:10.1051/0004-6361/201322342

V. Parmentier, J.J. Fortney, A.P. Showman, C. Morley, M.S. Marley, Transitions in the Cloud Composition of Hot Jupiters. Astrophys. J. **828**(1), 22 (2016). doi:10.3847/0004-637X/828/1/22

A. Pinhas, N. Madhusudhan, S. Gandhi, R. MacDonald, H_2O abundances and cloud properties in ten hot giant exoplanets. Mon. Not. R. Astron. Soc. **482**, 1485–1498 (2019). doi:10.1093/mnras/sty2544

A. Pinhas, B.V. Rackham, N. Madhusudhan, D. Apai, Retrieval of planetary and stellar properties in transmission spectroscopy with AURA. Mon. Not. R. Astron. Soc. **480**(4), 5314–5331 (2018). doi:10.1093/mnras/sty2209

B.V. Rackham, D. Apai, M.S. Giampapa, The Transit Light Source Effect: False Spectral Features and Incorrect Densities for M-dwarf Transiting Planets. Astrophys. J. **853**, 122 (2018). doi:10.3847/1538-4357/aaa08c

B. Rackham, N. Espinoza, D. Apai, M. López-Morales, A. Jordán, D.J. Osip, N.K. Lewis, F. Rodler, J.D. Fraine, C.V. Morley, J.J. Fortney, ACCESS I: An Optical Transmission Spectrum of GJ 1214b Reveals a Heterogeneous Stellar Photosphere. Astrophys. J. **834**(2), 151 (2017). doi:10.3847/1538-4357/aa4f6c

E. Rauscher, K. Menou, A General Circulation Model for Gaseous Exoplanets with Double-gray Radiative Transfer. Astrophys. J. **750**(2), 96 (2012). doi:10.1088/0004-637X/750/2/96

M. Rocchetto, I.P. Waldmann, O. Venot, P.-O. Lagage, G. Tinetti, Exploring Biases of Atmospheric Retrievals in Simulated JWST Transmission Spectra of Hot Jupiters. Astrophys. J. **833**, 120 (2016). doi:10.3847/1538-4357/833/1/120

C.D. Rodgers, *Inverse Methods for Atmospheric Sounding* (World Scientific, ???, 2000)

H. Schwarz, M. Brogi, R. de Kok, J. Birkby, I. Snellen, Evidence against a strong thermal inversion in HD 209458b from high-dispersion spectroscopy. Astron. Astrophys. **576**, 111 (2015). doi:10.1051/0004-6361/201425170

F. Selsis, R.D. Wordsworth, F. Forget, Thermal phase curves of nontransiting terrestrial exoplanets. I. Characterizing atmospheres. Astron. Astrophys. **532**, 1 (2011). doi:10.1051/0004-6361/201116654

K.B. Sheppard, A.M. Mandell, P. Tamburo, S. Gand hi, A. Pinhas, N. Madhusudhan, D. Deming, Evidence for a Dayside Thermal Inversion and High Metallicity for the Hot Jupiter WASP-18b. Astrophys. J. **850**(2), 32 (2017). doi:10.3847/2041-8213/aa9ae9

D.K. Sing, J.J. Fortney, N. Nikolov, H.R. Wakeford, T. Kataria, T.M. Evans, S. Aigrain, G.E. Ballester, A.S. Burrows, D. Deming, J.-M. Désert, N.P. Gibson, G.W. Henry, C.M. Huitson, H.A. Knutson, A. Lecavelier Des Etangs, F. Pont, A.P. Showman, A. Vidal-Madjar, M.H. Williamson, P.A. Wilson, A continuum from clear to cloudy hot-Jupiter exoplanets without primordial water depletion. Nature **529**, 59–62 (2016). doi:10.1038/nature16068

J. Skilling, Nested sampling for general bayesian computation. Bayesian Anal. **1**(4), 833–859 (2006). doi:10.1214/06-BA127. https://doi.org/10.1214/06-BA127

K.B. Stevenson, J.L. Bean, N. Madhusudhan, J. Harrington, Deciphering the Atmospheric Composition of WASP-12b: A Comprehensive Analysis of its Dayside Emission. Astrophys. J. **791**, 36 (2014a). doi:10.1088/0004-637X/791/1/36

K.B. Stevenson, J.-M. Désert, M.R. Line, J.L. Bean, J.J. Fortney, A.P. Showman, T. Kataria, L. Kreidberg, P.R. McCullough, G.W. Henry, D. Charbonneau, A. Burrows, S. Seager, N. Madhusudhan, M.H. Williamson, D. Homeier, Thermal structure of an exoplanet atmosphere from phase-resolved emission spectroscopy. Science **346**, 838–841 (2014b). doi:10.1126/science.1256758

J. Tennyson, S. Yurchenko, The ExoMol Atlas of Molecular Opacities. Atoms **6**(2), 26 (2018). doi:10.3390/atoms6020026

G. Tinetti, A. Vidal-Madjar, M.-C. Liang, J.-P. Beaulieu, Y. Yung, S. Carey, R.J. Barber, J. Tennyson, I. Ribas, N. Allard, G.E. Ballester, D.K. Sing, F. Selsis, Water vapour in the atmosphere of a transiting extrasolar planet. Nature **448**, 169–171 (2007). doi:10.1038/nature06002

A. Tsiaras, I.P. Waldmann, T. Zingales, M. Rocchetto, G. Morello, M. Damiano, K. Karpouzas, G. Tinetti, L.K. McKemmish, J. Tennyson, S.N. Yurchenko, A Population Study

of Gaseous Exoplanets. Astron. J. **155**, 156 (2018). doi:10.3847/1538-3881/aaaf75

T. Zingales, I.P. Waldmann, ExoGAN: Retrieving Exoplanetary Atmospheres Using Deep Convolutional Generative Adversarial Networks. Astron. J. **156**(6), 268 (2018). doi:10.3847/1538-3881/aae77c